Efficient PDE based numerical estimation of credit and liquidity risk measures for realistic derivative portfolios

degraaf, C.S.L.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Efficient PDE based numerical estimation of credit and liquidity risk measures for realistic derivative portfolios

\[
\frac{\partial u}{\partial \tau} + \frac{1}{2} x^2 \frac{\partial^2 u}{\partial x^2} + \frac{1}{2} \gamma^2 v \frac{\partial^2 u}{\partial v^2} + \frac{1}{2} \eta^2 \frac{\partial^2 u}{\partial (r^d)^2} + \frac{1}{2} \eta^2 \frac{\partial^2 u}{\partial (r^f)^2} \\
+ (r^d - r^f) v \frac{\partial u}{\partial x} + \kappa (v - v) \frac{\partial u}{\partial v} + \lambda_d (\theta_d (T - \tau) - r^d) \frac{\partial u}{\partial r^d} \\
+ \lambda_f (\theta_f (T - \tau) - r^f - \rho_1 \eta f \sqrt{v}) \frac{\partial u}{\partial r^f} \\
+ \rho_1 \gamma x v \frac{\partial^2 u}{\partial x \partial v} + \rho_1 \eta \eta_d x \sqrt{v} \frac{\partial^2 u}{\partial x \partial r^d} + \rho_1 \eta f x \sqrt{v} \frac{\partial^2 u}{\partial x \partial r^f} \\
+ \rho_2 \gamma \eta d \sqrt{v} \frac{\partial^2 u}{\partial v \partial r^d} + \rho_2 \gamma \eta f \sqrt{v} \frac{\partial^2 u}{\partial v \partial r^f} + \rho_3 \eta \eta d \frac{\partial u}{\partial r^d} + \rho_3 \eta \eta f \frac{\partial u}{\partial r^f} - r^d \frac{\partial u}{\partial d}
\]

Kees de Graaf
Efficient PDE based numerical estimation of credit and liquidity risk measures for realistic derivative portfolios

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie, in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 13 december 2016, te 10:00 uur

door

Cornelis Simon Leendert de Graaf

geboren te 's Gravenhage
The work described in this thesis was carried out in the Computational Science research group of the University of Amsterdam within the context of the project named ‘Advanced Estimation of Credit Valuation Adjustment’. This research is supported by the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO) and partly funded by the Ministry of Economic Affairs (project number 12214).
Summary

Samenvatting

1 Introduction 1
  1.1 Derivatives ........................................ 1
  1.1.1 The Black-Scholes pricing formula ............... 2
  1.1.2 Limitations of Black-Scholes ..................... 3
  1.2 Credit and liquidity risk ............................ 5
  1.3 New pricing and risk framework ....................... 6
  1.3.1 Credit Value Adjustment (CVA) ..................... 6
  1.3.2 Debt Value Adjustment (DVA) ....................... 8
  1.3.3 Funding Value Adjustment (FVA) ..................... 8
  1.3.4 Capital Value Adjustment (KVA) .................... 9
  1.4 Complexity of exposure estimation .................. 9
  1.4.1 Computational methods for option pricing ......... 10
  1.4.2 Computational methods for exposure ............... 11
  1.5 Research questions and outline ..................... 12

2 Preliminaries 13
  2.1 The Black-Scholes model ............................ 13
  2.2 CVA .................................................. 16
  2.3 Other risk measures .................................. 18
  2.3.1 Exposures .......................................... 18
  2.3.2 Quantiles .......................................... 19
  2.4 Finite difference technique ........................ 20
  2.4.1 Space discretization ................................ 20
  2.4.2 Finite difference approximations in space ......... 21
  2.4.3 ADI method ........................................ 22
3 Efficient computation of exposure profiles for counterparty credit risk 23
  3.1 Introduction ........................................ 24
  3.2 Problem Formulation ................................. 25
    3.2.1 CVA and Exposure of Bermudan options under Heston’s model ......................... 25
  3.3 Numerical Methods to Compute Expected Exposure ............ 26
    3.3.1 General pricing approach ......................... 26
    3.3.2 The Finite-Difference-Monte-Carlo method .......... 27
    3.3.3 Stochastic-Grid-Bundling method .................. 30
    3.3.4 The COS-Monte-Carlo method ....................... 32
  3.4 Numerical Results .................................... 35
    3.4.1 Comparison of Black-Scholes to Heston to assess impact of stochastic volatility on exposure .......... 36
    3.4.2 Error FDMC ................................... 39
    3.4.3 Error SGBM ................................... 40
  3.5 Conclusions .......................................... 43

4 Efficient estimation of sensitivities for counterparty credit risk with the finite difference Monte-Carlo method 45
  4.1 Introduction ........................................ 46
  4.2 Problem Formulation .................................. 46
    4.2.1 CVA under the Heston Hull-White Model .......... 46
  4.3 Computation of CCR and sensitivities .................... 48
    4.3.1 The FDMC Method ................................ 48
    4.3.2 The finite-difference method ..................... 49
    4.3.3 Computing CVA and its Sensitivities .............. 53
    4.3.4 Pricing a Portfolio .............................. 58
  4.4 Numerical Results .................................... 60
    4.4.1 Heston Model .................................. 61
    4.4.2 Heston Hull-White model ......................... 69
  4.5 Conclusion .......................................... 70

5 Efficient exposure computation by risk factor decomposition 73
  5.1 Introduction ........................................ 74
  5.2 Problem formulation .................................. 76
  5.3 Approximation by risk factor decomposition ............... 78
    5.3.1 An anchored-ANOVA-type approximation .......... 79
    5.3.2 A control variate ............................... 82
    5.3.3 Application to derivative portfolios .......... 84
  5.4 Numerical approximation of the Kolmogorov PDEs .......... 84
CONTENTS

5.4.1 Backward equations ........................................ 84
5.4.2 Forward equations .......................................... 85
5.5 Case studies .................................................... 87
5.5.1 Driving risk factors ......................................... 87
5.5.2 Derivatives .................................................. 88
5.5.3 Market and calibration ...................................... 90
5.6 Results .......................................................... 92
5.6.1 Case A: Single EURUSD CCYS ............................ 93
5.6.2 Case B: Three CCYS (EURUSD, EURGBP and EURJPY) 95
5.6.3 Case C: Case B with an additional IRS in EUR ........ 97
5.6.4 Case D: Case C with an additional EURUSD FX call option 97
5.6.5 Variance reduction .......................................... 98
5.6.6 Other base risk factors ...................................... 100
5.7 Conclusion ....................................................... 103

6 Smile and default: the role of stochastic volatility and interest rates in counterparty credit risk 105
6.1 Introduction ...................................................... 106
6.2 Problem Formulation .......................................... 107
6.2.1 Models ..................................................... 107
6.2.2 Derivatives .................................................. 109
6.2.3 Algorithm ................................................... 110
6.3 Market data and model calibration .......................... 112
6.3.1 Data ........................................................ 112
6.3.2 Calibration .................................................. 113
6.4 Case studies ...................................................... 116
6.4.1 Cross-currency swaps ..................................... 116
6.4.2 Barrier Options ............................................. 122
6.4.3 American options .......................................... 125
6.5 Discussion ....................................................... 127

7 Discussion .......................................................... 131
7.1 Contributions of individual Chapters ....................... 132
7.2 Future research on value adjustments ....................... 134
7.3 Future challenges ............................................... 135

A CVA and sensitivities for barrier options ...................... 137

B Market data ......................................................... 141
B.1 Implied volatility bootstrapping ............................. 141
B.2 Calibrated parameters ......................................... 142
B.3 Drift over time .................................................. 142
CONTENTS

B.4 Model parameters ........................................ 143
B.5 Numerical parameters .................................... 144
B.6 Regression-based Monte Carlo algorithm .............. 145
B.7 Finite difference errors and individual terms for Case B .... 145

Acknowledgements ............................................. 159
Publications .................................................. 161
Summary

Efficient PDE based numerical estimation of credit and liquidity risk measures for realistic derivative portfolios

In the Basel III accords in 2013, it was stated that financial institutions should charge Credit Value Adjustment (CVA) to their counterparties for (previously under-regulated) Over-The-Counter (OTC) trades. This CVA can be used to hedge a possible default of the counterparty. One important ingredient of CVA is the calculation of the future exposure of the portfolio on which CVA has been charged. This future exposure is also used to determine more recent value adjustments like Debt Value Adjustment (DVA) and Capital Value Adjustments (KVA), and can be calculated from a future distribution of the portfolio value. This distribution can also be used to compute quantiles which are so-called ‘worst case’ scenarios, and are therefore relevant for risk management. Computing the distribution of future portfolio values requires simulating the future states of the risk factors, and then evaluating the portfolio in all these future states. As the number of risk drivers in a typically traded portfolio is high, computing exposure for portfolios is a numerical challenge. One of the key contributions of this thesis is the Finite Difference Monte Carlo (FDMC) method for an efficient computation of future exposure.

The method is validated by comparing it with two other computational techniques, namely a semi-analytic method and the regression based, Stochastic Grid Bundling (SGBM) method. The comparison is made by considering exposures, quantiles and sensitivities of exposures of exotic options driven by one, two or three risk factors.

For portfolios driven by more than four risk factors, the exposure can be approximated by using a dimension reduction technique. We show that by decomposing the problem, exposure profiles for portfolios consisting of multiple derivatives driven by even 7 different risk factors can be computed by solving only one, two and three-dimensional PDEs.

By using the FDMC method, it is possible to incorporate stochastic volatility
and stochastic interest rate in exposure calculations of Foreign Exchange derivatives. For two real market scenarios, we have shown that these factors have to be taken into account, and an industry standard piecewise constant volatility model is not sufficient whenever considering expected or worst case scenario outcomes in Counterparty Credit Risk (CCR).
Samenvatting

Numerieke benaderingen met behulp van partiële differentiaalvergelijking voor krediet en liquiditeitsrisico van portefeuilles van financiële derivaten

De Basel III regelgeving uit 2013 schrijft voor dat financiële instellingen ‘Credit Value Adjustment’ (CVA) in rekening moeten brengen bij partijen waarmee ze derivaten verhandelen. De prijs van een derivaat wordt dan aangepast op basis van de kredietwaardigheid van de tegenpartij. Als deze kredietwaardigheid slecht is, is de CVA groot en vice versa. Deze prijsaanpassing kan worden gebruikt om bescherming in te kopen tegen een mogelijk faillissement van de betreffende tegenpartij. Een belangrijke stap in het kwantificeren van CVA is de berekening van de toekomstige waarde van het betreffende derivaat. Deze waarde, ook wel exposure genoemd, wordt ook gebruikt voor het berekenen van andere recent geïntroduceerde prijsaanpassingen zoals ‘Debt Value Adjustment’ (DVA) en ‘Capital Value Adjustment’ (KVA) en kan worden afgelezen van de toekomstige distributie van de waarde van het derivaat. Deze distributie kan ook worden gebruikt om kwantielen te berekenen. Deze kwantielen kunnen fungeren als ‘worst case’ scenario’s, en zijn daarom relevant voor risicomanagement. Prijsaanpassingen zijn vooral belangrijk bij portefeuilles die verschillende derivaten bevatten die gedreven worden door meerdere risicofactoren en daardoor een hoge waarde kunnen hebben. Het berekenen van de toekomstige distributies van portefeuillewaardes vereist simulatie van de relevante risicofactoren en het kunnen waarderen van deze portefeuille voor alle mogelijke toestanden. Het aantal relevante risicofactoren in typisch verhandelde portefeuilles is zo groot dat het berekenen van dit exposure een numerieke uitdaging is. De belangrijkste bijdrage van dit proefschrift is de ontwikkeling van de ‘Finite Difference Monte Carlo’ (FDMC) -methode, een numerieke methode om de distributie van portefeuillewaardes te berekenen.

De methode is gevalideerd door het te vergelijken met twee andere methodes, een semi-analytische methode en een methode gebaseerd op regressie. Voor exposures, de bijbehorende gevoeligheden en kwantielen gedreven door een, twee of
drie risicofactoren, is de FDMC-methode nauwkeurig en computationeel efficiënt.

Voor portefeuilles die gedreven worden door meer dan vier risicofactoren kan exposure berekend worden met behulp van dimensiereductie. Ik laat zien dat door een hoogdimensionaal probleem op te splitsen, de exposure voor portefeuilles van verschillende derivaten kan worden benaderd door alleen een-, twee- of drie-dimensionale partiële differentiaalvergelijkingen op te lossen. Met behulp van de FDMC-methode kunnen stochastische volatiliteit en stochastische rente worden meegenomen in exposureberekeningen voor derivaten op wisselkoersen. Voor twee realistische marktscenario’s heb ik laten zien dat deze stochastische factoren niet buiten beschouwing kunnen worden gelaten en dat het veel gebruikte constante volatiliteitsmodel niet voldoende is om ‘worst case’ scenario’s te berekenen.
1

Introduction

“In my view, derivatives are financial weapons of mass destruction, carrying dangers that, while now latent, are potentially lethal”

Warren Buffett (Quote from Berkshire Hathaway annual report 2002.)

1.1 Derivatives

It is often said that modern finance started in Amsterdam where the Verenigde Oost-Indische Compagnie established the Amsterdam Stock Exchange (AEX) [75]. Derivatives are contracts whose value depends on (or can be derived from) one or more underlyings. Already in the nineteenth century, these contracts were used, for example, to help sell crops that still needed to be grown. In the same century, the so-called derivative exchange was established in Chicago. The exchanges were used as markets where farmers and merchants could trade standardized contracts. Nowadays, there are many different users of derivatives. Sovereigns, central banks, regional/local authorities, hedge funds, asset managers, pension
funds, insurance companies, corporates, and even individual investors, all use derivatives. They do this as part of their investment strategy, or to hedge risk that is stemming from their business activities [48].

Options are derivatives which can be exercised by the holder depending on the value of the underlying entity, e.g., a stock or FX rate. A call option on a stock with a strike value $K$ gives the holder the right, but not the obligation, to buy the underlying stock for price $K$ at a future time $T$. The option market expanded with the introduction of listed stock options on the Chicago Board Option Exchange (CBOE) in 1973 [77] and today, derivatives are traded all over the world on numerous exchanges and so-called Over-The-Counter (OTC) markets. OTC means that the contract is bilateral and there is no exchange that functions as an intermediary.

For these trades, specific agreements can be made to make the trade more secure. For example, in a loan, a borrower can offer assets to a lender to secure the loan. This is called collateral and it is often posted in the form of cash. Another way of securing OTC derivatives is by central clearing, where the derivatives are traded via a Central Counterparty Clearing House (CCP). The CCPs guarantee the transactions of a group of participants (also called the clearing members) by demanding collateral from its members.

According to Gregory [48], 91% of the derivatives traded in 2014 are traded over the counter. Of this 91%, 60% is centrally cleared and from the remaining 40%, collateral is posted in 80% of the cases. So in total only 7%\(^1\) is traded without any form of collateral or central clearing. Note that, as the notional outstanding in OTC and exchange traded derivatives is around 800 trillion [48], this relatively small part still accounts for trillions of notional yearly.

There are different asset classes involved in derivatives. Over the years it can be seen that interest rates represent the biggest part in terms of notional, see the first column in Table 1.1. This, however, gives a somewhat misleading view. For interest rate derivatives, coupon payments are made over time and these net payments are often only a small part of the notional. In this case the sum of the coupons is the so-called market value of the trade: the amount that is possibly lost in case of a default. Therefore, if we look at the market value of the derivatives per asset class as percentage of notional in Table 1.1, we see that the impact of interest rates is still significant, but relatively smaller.

1.1.1 The Black-Scholes pricing formula

The issuer of a derivative would like to reduce the risk of losing significant amounts of money due to price movements. This can be done by offsetting the position and this is called hedging. The value of the derivative should therefore be

\(^1\)We arrive at 7% by taking 20% of 40% of the initial 91% OTC trades.
1.1. DERIVATIVES

<table>
<thead>
<tr>
<th>Notional outstanding</th>
<th>Market value</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest rate</td>
<td>384.03</td>
<td>10.15</td>
</tr>
<tr>
<td>Foreign exchange</td>
<td>70.45</td>
<td>2.58</td>
</tr>
<tr>
<td>Credit default swaps</td>
<td>12.29</td>
<td>0.42</td>
</tr>
<tr>
<td>Equity</td>
<td>7.14</td>
<td>0.50</td>
</tr>
<tr>
<td>Commodities</td>
<td>1.32</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 1.1: Notional amounts outstanding per asset class compared to market value in trillions of US dollars as of December 2015. Source: http://stats.bis.org/stats/srs/table/d5.1 (date accessed: August 15 2016).

equal to the cost of hedging. This principle was given a mathematical basis by the development of the Black-Scholes pricing framework in 1973 [13], long before the 2008 credit and liquidity crisis. The model is described in more detail in Chapter 2. An important assumption underlying the Black-Scholes model is the absence of arbitrage, which is defined as the possibility of a guaranteed profit without any risk taken. For example, a trading strategy which has zero initial costs and a strictly positive payoff with probability 1 would be an arbitrage opportunity. By making additional assumptions about the dynamics of the underlying asset, the Black-Scholes model can be solved analytically [58]. However, history has proven that these assumptions do not always hold, as we will illustrate below.

1.1.2 Limitations of Black-Scholes

Although the Black-Scholes formula gave market participants a mathematical foundation to their trading activities, this new knowledge could not guarantee stability of the financial system. This became evident when Long Term Capital Management (LTCM) was bailed out in September 1998. This hedge fund, involving Robert Merton and Myron Scholes, the founders of the Black-Scholes theory, tried to buy options that were underpriced and sell options that were overpriced. LTCM identified which options to trade by comparing option quotes with the prices computed by the formula, which initially proved to be very lucrative. However, in 1998, Russia defaulted on its debt and this came as a surprise to many investors, who in a reaction, searched for more solid investments. This drove up the prices of securities that LTCM was short on and therefore it got in trouble and had to eliminate many investments [74].

The return over a time period of a traded asset is defined as the gain or loss of this asset over that period. The Black-Scholes model implies that the log of the returns follow a normal distribution. However, when markets are in stress, this assumption does not always hold, and this can be seen by looking at the performance of a number of stocks, so-called stock indices, during crisis periods.
For example, on October 19, 1987, a date often referred to as “Black Monday”, the Dow Jones Industrial Average (DJIA), which is an important stock index, fell by more than 20%. This event also elucidated the danger of new technologies, as computers were used to generate the trades, but the exchange systems were overloaded causing delays and making index arbitrage possible [58].

An important flaw of the Black-Scholes model is known as the implied volatility smile phenomenon. The implied volatility of an option is defined as the value of volatility needed to replicate the market prices. The Black-Scholes formula to price a call option requires as input the volatility of the underlying stock. Together with option specific parameters as the strike and maturity of the option, a theoretical price of the option can be derived. The volatility value that produces a theoretical value which is exactly equal to the market price is called the implied volatility. This implied volatility, plotted as a function of the strike price, is typically not a straight line. The curvature is known as the volatility smile (see Figure 1.1). The smile gives rise to an implied distribution of the underlying asset which has fatter tails than a log-normal distribution.

![Figure 1.1: EURUSD implied volatilities as a function of moneyness (which is defined as strike as percentage of the underlying value at time $t = t_0$) for different maturities observed on September 9, 2008.](image)

A popular extension of the Black-Scholes model is the use of stochastic volatility [57] or local volatility [72] models. The computation of option prices then becomes more challenging as under these models analytic expressions of option prices might not be available for all option types. In addition to that, calculation of quantities needed for hedging, can be more complex.

The hedging argument, as originally proposed in the Black-Scholes model, concerns only market risk. In a world where financial institutions could not
1.2. CREDIT AND LIQUIDITY RISK

default and there is always enough liquidity this hedge would be perfect. In reality, unfortunately this is not the case. Due to several crises, we have learned that risk management involves more than offsetting positions to hedge out the market risk. The risk of counterparty default or the risk that there is not enough liquidity to hedge, are just two examples which make risk management of prime importance for financial institutions.

1.2 Credit and liquidity risk

From a market participant’s point of view, proper risk management helps avoiding a downturn as a consequence of unexpected market behavior. However, from the perspective of sovereign entities, the possible monetary effects of a financial crisis is of great concern \[11\]. This was further emphasized by the recent credit crisis, which completely changed the world of derivative pricing \[48\].

The credit crisis culminated in 2008 when governments bailed out large banks, namely Royal Bank of Scotland in the UK and AIG in the US. For these bailouts taxpayers’ money was used to cover the losses that occurred due to excessive risk-taking \[48\]. In September 2008 Lehman Brothers, the fourth largest US investment bank with more than 100 years of history, declared bankruptcy. This demonstrated that there is no such thing as “too big to fail”.

One triggering event was the sharp rise in interest rates in the US, which slowed down the housing market. The sub-prime loans were given out to people with a poor or no credit history to buy houses and when house prices were rising that was fine, but when the house prices dropped, these people started defaulting on their mortgages.

Big banks were trading on these mortgages by packaging them together. These packages of sub-prime mortgages were popular investments. The return on the investment of these packages is backed by the value of these mortgages and therefore the investments are called Mortgage Backed Securities (MBSs). MBSs could get a good credit rating by combining the sub-prime loans together with more solid loans. Because of the good rating, many financial institutions invested in these securities, but had to report big losses in September 2007 when initial defaults set in.

At the same time, financial institutions traded a class of instruments, to protect themselves against a possible default of these trades. This was done via Credit Default Swaps (CDSs). In return for a premium, the issuer of a CDS pays out a predefined amount in case of a default to the buyer. This instrument offers a protection against a default of the underlying name. There was an estimated amount of $400 billion of CDS insurance written on Lehman Brothers. When the bank defaulted, this triggered massive pay-outs, such that the counterparties that provided the protection suffered big losses and also got into financial problems.
CHAPTER 1. INTRODUCTION

This cascading effect was the on-set of the systemic credit and liquidity crisis.

The premium paid for this protection provides information about the markets’ expectation of the likelihood of the default of the counterparty. This information can be used to compute a market implied default probability. The existence of the CDS market offers an opportunity to hedge default risk, and this concept formed the basis of the first value adjustment that was proposed by regulators after the credit crisis.

1.3 New pricing and risk framework

“People are dying, we need guns!”
______________________________

Because of the crisis, financial regulators had more momentum and succeeded in signing the Dodd-Frank act in 2010, and agreed upon the Basel III accords in 2013 (only 3 years after the Basel II accords that did not provide the desired stability). In these accords, it was stated that financial institutions should charge Credit Value Adjustment (CVA) to their counterparties for the (previously under-regulated) OTC trades. This sparked the beginning of the value adjustments, which posed a new computational challenge for financial institutions.

1.3.1 Credit Value Adjustment (CVA)

The credit crisis has shown that not taking into account the possibility of a counterparty default has serious consequences. Counterparty risk is defined as the risk that a counterparty will fail to deliver contractual obligations, thus any undelivered cashflow in the contract is considered a default. By definition CVA is the difference between the risk-free value of a trade and the risky value which takes into account this default:

\[
\text{Risky value} = \text{Risk-free value} - \text{CVA}. \tag{1.1}
\]

This expression assumes that any derivative value can be separated in a risk-free part and a risky part. Clearly CVA thus depends on three factors, namely:

1. The Probability of Default (PD).
2. The Loss Given Default (LGD), which is the fraction of the exposure that is lost in the case of a default event.
3. The expected value of the portfolio that an institution has traded with the counterparty. This is the so-called Expected Positive Exposure (EPE).
1.3. NEW PRICING AND RISK FRAMEWORK

The probability of default

By charging CVA, an investor can theoretically hedge away default risk. This is why CVA is computed under the risk neutral measure and as such, risk neutral (market induced) default probabilities are needed [48]. This information can be extracted from the market by looking at CDS quotes. A commonly used approximation [48] is:

\[
PD(t_{i-1}, t_i) \approx \exp\left(\frac{s_{t_{i-1}} t_{i-1}}{LGD}\right) - \exp\left(-\frac{s_t t_i}{LGD}\right),
\]

where PD\((t_{i-1}, t_i)\) is the probability of default in time interval \((t_{i-1}, t_i)\), \(s_t\) is the credit spread at time \(t\). Note that the CDS market is an illiquid market and, as such, not all default probabilities can be approximated in this fashion. This leads to additional model and liquidity risk.

Loss Given Default (LGD)

The LGD depends on the Recovery (R) of the exposure after a default. It is highly uncertain and varies significantly depending on the seniority of the claim, economic state and sector. As is shown in the previous subsection, this LGD is also used in the computation of the probability of default. A higher LGD directly results in an increased CVA, but implicitly, as a result of a higher LGD, also the probability of default decreases and that results in a lower CVA, see Equation (1.2). Because of this cancellation effect the LGD is often assumed to be constant [48]. There is, however, some work on modeling the recovery see e.g. Unal et al. [99] and Schläfer and Uhrig-Homburg [86]. Throughout this thesis, we assume the LGD to be constant.

Expected Positive Exposure (EPE)

The expected positive exposure can be seen as the future value of the position. As these future values can be non-trivial option prices, this quantity can be hard to calculate. For example, valuing exotic derivatives at time zero is already a challenge, so calculating multiple future values is even harder. In addition to that, one is typically interested in the CVA of a portfolio that consists of multiple derivatives that are driven by multiple risk factors, which are all correlated with each other and thus should not be modeled independently. From a computational perspective, this is a demanding task. Because of this, EPE is the most computational challenging building block of CVA. Furthermore, as we will see later, it plays an important role in other value adjustments. In this work we therefore focus on the efficient computation of this key quantity using numerical techniques.
By combining the EPE, LGD and the probability of default, the commonly used formula for CVA [103] is defined by:

\[
CVA = -(1 - \delta) \sum_{i=1}^{N_T} \text{EPE}^*(t_i) \cdot PD(t_{i-1}, t_i),
\]

(1.3)

were \(\delta\) is the recovery rate, \(\text{EPE}^*(t)\) the discounted expected positive exposure at time \(t\), \(PD(t_{i-1}, t_i)\) the probability of default in \([t_{i-1}, t_i]\) and \(N_T\) the number of time points.

### 1.3.2 Debt Value Adjustment (DVA)

In bilateral contracts, there is credit risk on both sides of the trades. CVA accounts for the default probability of the issuer of the contract, whereas Debt Value Adjustment (DVA) covers the default of the buyer of the contract. If a party considers both its own default as well as the default of the counterparty, one speaks of Bilateral CVA (BCVA):

\[
BCVA = CVA + DVA,
\]

(1.4)

where DVA is defined as:

\[
DVA = (1 - \delta) \sum_{i=1}^{N_T} \text{ENE}^*(t_i) \cdot PD(t_{i-1}, t_i),
\]

(1.5)

were \(\text{ENE}^*(t)\) is the expected negative exposure. Here, the recovery rate and default probability is that of the issuer instead of the counterparty as in the CVA case. There has been quite some debate (see, e.g. Gregory [48] and Kenyon [70]), if DVA should be included in the valuation. Widening of the credit spread of a bank increases DVA and therefore will result in a profit for the bank. This is counterintuitive, since it suggests that deterioration of a banks creditworthiness could lead to a profit.

### 1.3.3 Funding Value Adjustment (FVA)

Although more and more derivatives are traded with collateral, still a big fraction of OTC derivatives are traded without [48]. Not all market participants (such as corporates and sovereigns) have the liquidity or capacity to post collateral frequently. These uncollateralized trades still need to be hedged by banks and that is where funding comes into play.

Consider an uncollateralized trade of a bank with a party, which is hedged by setting up an interbank trade. For this hedge, the bank needs to post collateral...
and if the trade moves in the money, the hedge moves into the opposite direction
and so, more collateral needs to be posted. The reverse is also true, when the
trade moves out of the money, collateral will be received. The bank may need
funding for this trade and that is an additional cost or benefit. An intuitive FVA
formula is equal to:

\[
FVA = - \sum_{t=1}^{N_T} (EPE(t_i) - ENE(t_i)) FS(t_i)(t_i - t_{i-1})SP(t_i),
\]

(1.6)

where FS(t_i) is the funding spread, i.e. the costs for lending or borrowing the col-
lateral that needs to be posted or received, and SP(t_i) is the survival probability
of both the parties concerned. Note that again the exposure is needed.

1.3.4 Capital Value Adjustment (KVA)

Regulators have imposed strict governance on the amount of capital that a bank
must hold to absorb unexpected losses when markets are in stress. The amount
of capital required depends on the trading book of the bank, the more risky
OTC trades a bank has, the more capital is required. To account for this cost of
capital, dealers have introduced the so-called Capital Value Adjustment (KVA).
The family of all these value adjustment is now referred to as xVA and they
have in common that they all depend on future values of portfolios, or in other
words expected exposure. This exposure is thus extremely relevant for financial
institutions and efficient computation of this quantity is vital.

1.4 Complexity of exposure estimation

As the value of derivatives (and in general investments) depends on markets
expectation, getting information fast has always been important. In the 19th
century, carrier pigeons were used to deliver messages. Soon after, with advances
in technology, investors even laid a physical cable between the United States
of America and Great Britain to transfer messages faster with Morse code [53].
Today, in the computer age, trading on the basis of acquiring information faster
than others is very relevant in High Frequency Trading (HFT). This demand
for speed is also present in derivative pricing and motivates the need for high
performance computing. Equally important in this context are the availability
of efficient numerical techniques. These techniques are used for option prices,
but, as we will discuss in this section, are vital for the computation of value
adjustments as well. In the remainder of this section we will elaborate on the
complexity of exposure calculations. We will first give a brief summary of the
main numerical techniques used in the field of option pricing.
CHAPTER 1. INTRODUCTION

1.4.1 Computational methods for option pricing

As discussed before, pricing financial derivatives is a challenging task. The risk factors underlying the contract are modeled by Stochastic Differential Equations (SDEs) and often these SDEs in combination with the option payoff, do not allow for closed-form solutions of the corresponding derivatives. Due to the increased complexity of the derivatives market over time, the number of challenging problems has even grown in the past decades. Next to that, not only prices, but also the sensitivities of these prices with respect to market variables are needed for hedging purposes. Therefore one heavily relies on numerical approximations. In general, there are two commonly used numerical methods for the approximation of option values namely: Monte Carlo and PDE based methods.

Monte Carlo methods

The Monte Carlo method is a sampling technique where scenarios are generated by simulating paths. For simple options, the payoff along these paths can be evaluated and from these values, the distribution of the future value can be determined. By taking the mean of these future payoffs and discounting it back to time $t = t_0$, we estimate the option price. This method scales linearly in the number of risk factors and is intuitively easy to understand. This is why this method is widely used in industry. However, there are important drawbacks. The forward nature of the method makes it non-trivial to compute values of path-dependent options. For this, dynamic programming techniques, involving regression, can be used. In this case, this method no longer scales linear in the number of dimensions unless approximations are used. Next to that the sensitivities are non-trivial to compute [23].

PDE methods

Another popular numerical technique relies on approximating the solution of the partial differential equation (PDE) underlying the derivative value. An example is the Black-Scholes PDE which we will discuss in the next chapter. Solving this PDE can be done in a number of ways, e.g. via semi-analytic Fourier techniques (to be discussed briefly in Chapter 3) or via finite difference methods. In this approach, the PDE is discretized on a predefined computational grid and local finite difference approximations of the partial derivative terms in the PDE are evaluated. By doing this, a system of ODEs is obtained which can be solved by using a time stepping procedure. This involves matrix vector operations. Typically this method is very useful for the computation of path dependent options and sensitivities. Nevertheless it is only applicable to low-dimensional problems.
1.4. COMPLEXITY OF EXPOSURE ESTIMATION

In the next Chapter, we will discuss these methods in more detail, but first, we will illustrate how these methods can be used to compute exposures.

1.4.2 Computational methods for exposure

For the estimation of exposures we need the value of a portfolio along the life of all its trades. In the Monte Carlo framework, we estimate this by simulating the future state of all underlying risk factors, and evaluating the value of the portfolio on that state, as illustrated in figure 1.2. By doing this for many paths and time points, we obtain the future exposure distribution. Although this idea works for simple options, in the case of complex path-dependent payoffs, obtaining the option values for each future state in principle requires another nested Monte Carlo simulation. Efficient solutions to circumvent this problem can be found in Feng and Oosterlee [42] and Karlsson et al. [69].

The use of PDEs for exposure calculations is, to the best of our knowledge, very limited. It should be noted, that solving partial differential equations on a grid gives us future option values for a discrete set of states for free. It is therefore very interesting to investigate if one can leverage from this available information and circumvent the issues related to the nested Monte Carlo. Although interesting, this approach has not been extensively studied in the literature.
1.5 Research questions and outline

As argued before, in the context of exposure calculations, the Monte Carlo technique is industry standard. However, extending the PDE framework to compute exposure is very promising since the information on option prices for future time points and states is already present when the time zero option premium is calculated. Although interesting, there are several methodological challenges that need to be addressed here, namely:

i. How do we combine the information available on the PDE grid with the simulated Monte Carlo states? Is a straightforward interpolation sufficient, not only for the valuation, but also for the first and higher-order sensitivities?

ii. Can we estimate the distribution of the future states in a PDE framework by solving a forward PDE for the transition densities and couple this with future option values obtained from the backward PDE solution. This will result in a full PDE framework and thus allows us to eliminate the Monte Carlo noise and interpolation error.

iii. Since exposure is intrinsically a portfolio level calculation, for the PDE method to be useful, it should be applicable for high-dimensional problems. Can we therefore break the curse of dimensionality by using risk factor decomposition techniques?

In Chapter 3, 4 and 5 respectively, these challenges will be systematically addressed, resulting in a novel and efficient PDE based framework for the estimation of exposure of realistic portfolios. This framework will be systematically validated by a rigorous comparison with other methods for a wide range of test problems.

In Chapter 6, we will apply this framework to address a complex and urgent problem in derivative pricing, namely:

iv. What is the role of the volatility smile and stochastic rates in default risk?

This problem combines two key difficulties: the fat-tailed distributions of the underlying risk factors, which already in the context of option pricing poses challenges, and the intrinsic multi-dimensional hybrid nature of future portfolio distributions. For this investigation, a high dimensional model which includes stochastic volatility and interest rates needs to be used. In the literature, little is reported on this problem, mainly due to the computational challenges mentioned above. We demonstrate that our PDE framework can handle these difficulties. The real case studies considered, where actual market data corresponding to silent and stressed periods are employed, show that the role of stochastic volatility and rates is highly significant for default risk.
In this part we will first introduce the Black-Scholes model and derive the associated PDE. Next we derive the general CVA formula by analyzing the future value of the risk-free payoff, where a counterparty cannot default, and the risky payoff where a counterparty can default. Then, in the following section, we will discuss the other risk measures that are computed in this thesis. Finally, we will discuss the finite difference method to solve a PDE in detail and present all the schemes that are used throughout the thesis. Additional background knowledge on these subjects is provided in this chapter since this essential to understand the other chapters, and as such it is considered common knowledge throughout the rest of this thesis. We assume the reader to be familiar with partial- as well as stochastic differential equations and the notions of Brownian motions and martingales. Specifically, the concepts of risk neutral and real world measures are assumed to be known. Next to that we will use Itô’s lemma to arrive at a Partial Differential equation. More details on these basic concepts in mathematical finance can be found in [12].

2.1 The Black-Scholes model

The first assumption of the Black-Scholes model is that the value of the underlying entity follows a geometric Brownian motion. By definition, $X_t$ is said to follow a geometric Brownian motion with drift $\mu$ and volatility $\sigma$ if, for all $t, s > 0$ the
random variable $\frac{X_{t+s}}{X_t}$ is independent of all values up to time $t$ and if $\log \left( \frac{X_{t+s}}{X_t} \right)$ is a normal random variable with mean $\left( \mu - \frac{\sigma^2}{2} \right)s$ and variance $\sigma^2 s$. This process can be represented by a Stochastic Differential Equation (SDE):

$$dX_t = \mu X_t dt + \sigma X_t dW_t$$

(2.1)

where the value $dW_t$ is the change in the driving Brownian motion which is normally distributed with mean 0 and variance $dt$. By looking at historical time series of stock prices, this assumption seems plausible\(^1\). In figure 2.1, we see a path simulated by a geometric Brownian motion and the realized Apple stock values from 2011 to 2016. For this SDE we can derive an expectation with the help of Itô calculus. First we derive the distribution of $G_t = \log X_t$ by Itô’s lemma [12]:

$$dG_t = \left( \mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t.$$  

(2.2)

Next, it is assumed that:

- The market is free of arbitrage.
- Short selling is allowed.

\(^1\)Nevertheless, detailed quantitative analysis shows that returns do follow fat-tailed distributions.
2.1. THE BLACK-SCHOLES MODEL

- There are no transaction costs.
- The underlying does not pay dividend.
- There is continuous security trading.
- There is a risk free interest rate $r$ that is constant over time.

Some of these assumptions can be relaxed. In the following paragraph we show the key result of the Nobel prize winning research by Black and Scholes.

The Black-Scholes PDE

Now suppose $V(t, X)$ to be the price of the option on underlying $X$ at time $t$, by Itô’s lemma, we have:

$$
\frac{dV(t, X)}{dt} = \left( \frac{\partial V}{\partial X} \mu X + \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} \sigma^2 X^2 \right) dt + \frac{\partial V}{\partial X} \sigma X dW_t, \quad (2.3)
$$

Now a portfolio with value $\Pi_t$ is constructed which mimics the value of this option. The portfolio consists of the underlying and the derivative itself: sell (short) one derivative and buy $\Delta$ amount of shares:

$$
\Pi_t = -V(t, X) + \Delta X \quad (2.4)
$$

The change in value over the small time interval $dt$ satisfies the so-called self financing condition:

$$
\frac{d\Pi_t}{dt} = -\frac{dV(t, X)}{dt} + \Delta dX. \quad (2.5)
$$

Substituting equations 2.1 and 2.3 yields:

$$
\frac{d\Pi_t}{dt} = -\left( \frac{\partial V}{\partial X} \mu X + \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} \sigma^2 X^2 \right) dt - \frac{\partial V}{\partial X} \sigma X dW_t \\
+ \Delta \left( \mu X dt + \sigma X dW_t \right)
$$

$$
= -\left( \frac{\partial V}{\partial X} \mu X + \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} \sigma^2 X^2 - \Delta \mu X \right) dt - \left( \frac{\partial V}{\partial X} \sigma X - \Delta \sigma X \right) dW_t, \quad (2.6)
$$

and risk can be eliminated by choosing $\Delta = \frac{\partial V}{\partial X}$, such that:

$$
\frac{d\Pi_t}{dt} = -\left( \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial X^2} \sigma^2 X^2 \right) dt. \quad (2.7)
$$

Because of the absence of risk during time $dt$, this portfolio must have the same rate of return as any other risk-free investment, otherwise arbitrage would be
possible. Another important assumption of the Black-Scholes model is the existence of a risk-free interest rate \( r \) such that over a small time interval \( dt \) our portfolio will grow in value:

\[
d \Pi_t = r \Pi_t dt
\]

by substituting (2.5) and (2.7) into (2.8), we get

\[
\frac{\partial V}{\partial t} + rX \frac{\partial V}{\partial X} + \frac{1}{2} \sigma^2 X^2 \frac{\partial^2 V}{\partial X^2} - rV = 0.
\]

which is known as the Black-Scholes Partial Differential Equation (PDE). This PDE can be solved analytically by transforming it to the well-known heat equation, or approximated with the help of finite difference schemes. By choosing \( \Delta = \frac{\partial V}{\partial X} \), the replicating portfolio can be set up that can hedge out the risk. This quantity is also called the sensitivity of the option with respect to the underlying value. In general, sensitivities of option prices to movements of specific market parameters are of prime importance for risk management and are often referred to as the Greeks.

**Risk neutrality** Note that in the Black Scholes PDE, the initial drift \( \mu \) is no longer present. This absence of \( \mu \) in the final Black Scholes PDE comes from the fact that the equation does not involve any variable that is affected by the risk preferences of investors. The final formula only depends on the current stock price, time, the risk-free interest rate and the volatility of the stock. Because the risk preferences do not play a role, the assumption can be made that all investors are risk neutral. This is a very helpful assumption, because now the present value of any cashflow is equal to its expected value discounted by the risk-free rate. Looking at expectations of stochastic processes under the risk neutral world involves going from a objective probability measure \( P \) to a risk adjusted, or martingale measure \( Q \). Throughout this thesis we will work under the risk neutral measure \( Q \).

### 2.2 CVA

For the computation of CVA we first assume that a derivative value can be separated in a risk-free part and a risky part, as is shown in equation (1.1). By using this principle, in [48] an expression is derived for the risky value\(^2\). Let \( V(t,T) \)\(^3\) denote the risk-free value of an uncollateralized portfolio of trades with

\(^2\)The following derivation follows Appendix 14A of [48]

\(^3\)Here we slightly abuse notation and declare the value \( V \) to be a function of only time. Later on we will define \( V \) to be a function of risk factors.
2.2. CVA

one counterparty which has default time \( \tau \) and let \( \tilde{V}(t, T) \) denote the risky value. There are two possible scenarios:

1. The counterparty does not default before \( T (\tau > T) \)

In this case, the payoff is equal to the risk-free payoff and can be denoted as:

\[
1_{(\tau > T)} V(t, T),
\]

(2.10)

where \( 1_{(\tau > T)} \) is the indicator function which equals to 1 if \( \tau > T \) and zero otherwise.

2. The counterparty does default before \( T (\tau \in [t, T]) \)

In this case, the payoff consists of three terms, namely the amount already paid up to time \( \tau \) \( (V(t, \tau)) \) and the payoff at default time \( \tau \). Then, if the nominal value at time \( \tau \) is positive, a so-called recovery fraction \( \delta \) of the value can be paid out \( (\delta V(\tau, T)) \), however when it is negative, the full amount needs to be paid \( (V(\tau, T)) \). The payoff thus equals

\[
1_{(\tau \leq T)} (V(t, \tau) + \delta V(\tau, T)^+ + V(\tau, T)^-),
\]

(2.11)

where, \( x^+ = \max(x, 0) \) and \( x^- = \min(x, 0) \). With these two possible outcomes, the risk-neutral expected value of risky asset \( \tilde{V}(t, T) \) can now be defined as:

\[
\tilde{V}(t, T) = E^Q_t \left[ 1_{(\tau > T)} V(t, T) + 1_{(\tau \leq T)} (V(t, \tau) + \delta V(\tau, T)^+ + V(\tau, T)^-) \right],
\]

(2.12)

which can be simplified as follows:

\[
\tilde{V}(t, T) = E^Q_t \left[ 1_{(\tau > T)} V(t, T) + 1_{(\tau \leq T)} (V(t, \tau) + \delta V(\tau, T)^+ + V(\tau, T)^-) \right]
\]

\[
= E^Q_t \left[ 1_{(\tau > T)} V(t, T) + 1_{(\tau \leq T)} (V(t, \tau) + (\delta - 1) V(\tau, T)^+ + V(\tau, T)^-) \right]
\]

\[
= E^Q_t \left[ 1_{(\tau > T)} V(t, T) + 1_{(\tau \leq T)} (V(t, \tau) + (\delta - 1) V(\tau, T)^+) \right]
\]

\[
= E^Q_t \left[ V(t, T) + 1_{(\tau \leq T)} (\delta - 1) V(\tau, T)^+ \right]
\]

\[
= V(t, T) - E^Q_t \left[ 1_{(\tau \leq T)} (1 - \delta) V(\tau, T)^+ \right].
\]

(2.13)

Thus, we have a risk neutral value of CVA:

\[
CVA = -(1 - \delta) E^Q_t \left[ 1_{(\tau \leq T)} V(\tau, T)^+ \right].
\]

(2.14)

The one minus the recovery fraction \( (1 - \delta) \) is also called the Loss Given Default (LGD) and note that we made the assumption that \( V(\tau, T) \), includes discounting and it is the value knowing that default occurs at time \( \tau \). As default can happen
at any time before maturity, we can integrate over all possible default times and by dropping the assumption that $V(\tau, T)$ is discounted one has:

$$\text{CVA} = -(1 - \delta) \mathbb{E}^Q_t \left[ \int_t^T B(t, \tau) V(\tau, T)^+ dF(t, \tau) \right],$$

(2.15)

where $B(t, \tau)$ is the discount factor and $F(t, \tau)$ is the cumulative default probability between time $t$ and $\tau$ [103]. When the default probability is assumed to be independent of exposure\footnote{This assumption implies also independence between default probability and exposure which is called Wrong Way Risk (WWR) and will be treated in more detail later.}, the expectation and integration can be interchanged and approximated such that:

$$\text{CVA} \approx -(1 - \delta) \mathbb{E}^Q_t \int_t^T B(t, \tau) V(\tau, T)^+ dF(t, \tau)$$

(2.16)

2.3 Other risk measures

As described before, expectations of the future liabilities are needed for risk management. For derivatives or portfolios this boils down to knowledge of the future derivative or portfolio value distribution. From this distribution, for example the mean can be extracted, but one can also look at the quantiles to get an insight in the worst case scenario.

As shown in the introduction, all value adjustments of the xVA family depend on some sort of exposure.

2.3.1 Exposures

The future value $V$ of a derivative or portfolio can depend on multiple time dependent risk factors. Let these risk factors be represented by vector $\vec{\theta}_t$. From here on, we denote the value of the derivative driven by $\vec{\theta}_t$ as $V(\vec{\theta}_t, t)$. The present Expected Exposure (EE) at a future time $t_0 < T$ is given by:

$$\text{EE}(t_0, t) := \mathbb{E} \left[ V(\vec{\theta}_t, t) | \mathcal{F}_t \right],$$

(2.17)
2.3. OTHER RISK MEASURES

where $\mathcal{F}_t$ is the filtration at time $t$ \[12\], which can be interpreted as representing all historical, but not future information at time $t$. The discounted version of EE is computed as:

$$EE^*(t_0, t) := E \left[ B(t_0, t) V(\tilde{\theta}_t, t) | \mathcal{F}_t \right]. \tag{2.18}$$

where $B(t_0, t)$ is the discount factor. Throughout this thesis, the expectation is calculated under a risk-neutral measure $Q$.\footnote{Typically, the future states can also be modeled under real-world measure. This is possible when a simulation approach is used, but as this research focuses on the numerical aspect, the risk-neutral measure $Q$ is assumed.}

For risk management, the holder of a derivative contract needs to know the possible amount he can lose. From this point of view, it is obvious that one is interested in the case that a loss is positive (a negative loss may be a profit), therefore the exposure of an option at a future time $t < T$ is defined as:

$$E^+(t) := \max (V(\tilde{\theta}_t, t), 0). \tag{2.19}$$

Negative exposure $E^-(t)$ is defined as the minimum instead of the maximum of $V(\tilde{\theta}_t, t)$ and zero. The Expected Positive Exposure (EPE) and Expected Negative Exposure (ENE) at a future time $t < T$ are then given by:

$$\text{EPE}(t_0, t) := E \left[ E^+(t) | \mathcal{F}_t \right], \tag{2.20a}$$

$$\text{ENE}(t_0, t) := E \left[ E^-(t) | \mathcal{F}_t \right], \tag{2.20b}$$

and similarly as in the case of EE, the discounted versions are given by:

$$\text{EPE}^*(t_0, t) := E \left[ B(t_0, t) E^+(t) | \mathcal{F}_t \right], \tag{2.21a}$$

$$\text{ENE}^*(t_0, t) := E \left[ B(t_0, t) E^-(t) | \mathcal{F}_t \right]. \tag{2.21b}$$

In the case of a long position in an option, the price $V(\tilde{\theta}_t, t)$ is always positive and thus the EPE (2.20a) is equal to the EE.

2.3.2 Quantiles

The exposures (EE/EPE/ENE) provide insight into the mean properties of the contract value distribution which are mostly used to attach price tags to risk. However, from the point of risk, one is also interested in the worst-case scenarios which in our work are analyzed by looking at the quantiles of the future contract value distribution. The quantile functions $Q_{97.5\%}(t)$ ($Q_{2.5\%}(t)$) represent the best (worst) case scenario the buyer can face at a future time $t$ and is defined as:

$$Q_\theta(t) = \inf \{x : P(V(\tilde{\theta}_t, t) \leq x) \geq \theta \}. \tag{2.22}$$
CHAPTER 2. PRELIMINARIES

Note that if the quantile is simulated under the real-world measure, we speak of Potential Future Exposure (PFE). In that case the future scenarios needs to be generated by historically calibrated models [96]. In this thesis, when we don’t use fully calibrated models, we do not distinguish between real or risk-neutral measure and thus speak of PFE. However, when we do calibrate the model and compute the quantile under the risk neutral measure, we speak of quantiles.

2.4 Finite difference technique

For the finite difference method to solve a time-dependent partial differential equation, first a finite region in space on which a solution to the partial differential equation is needed is discretized, which results in a finite discrete space grid. We impose initial and boundary conditions. At the initial state, partial derivatives in space are approximated by using finite differences such that we have a finite number of ordinary differential equations. Subsequently, time is discretized and the ODE can be solved via a time stepping procedure. In this section we step by step explain the methods used.

2.4.1 Space discretization

Because we can only evaluate a finite number of grid points, the solution space is bounded, and boundary conditions are needed. On these boundaries, approximations are made and to limit the impact of these approximations, these boundaries are placed far from the region of interest. For example if we approximate an option on underlying asset $X_0$, the boundaries $X_{\text{min}}$ and $X_{\text{max}}$ are taken far away from $X_0$. The most straightforward discretization is a uniform discretization, where we divide the space in $m$ equidistant grid points. The accuracy of the finite difference approximation depends on the mesh: the more grid points are used, the more accurate the finite difference approximations. As the computational domain is stretched, a good way to gain more accurate approximations without adding extra points is by using non-uniform grids. These grids distribute more points in the region were high accuracy is required (e.g. close to $X_0$ in the example) and less points in less important regions (e.g. close to the boundaries). In any case, a discrete grid needs to be generated that is increasing, such that $X_{\text{min}} = x_0 < x_1, \ldots, x_m = X_{\text{max}}$. On this grid, option values $v$ can then be calculated at maturity by using the payoff function. For convenience, solutions on this grid are represented by a solution vector:

$$\vec{v} = (v_0, v_1, \ldots, v_{m-1}, v_m).$$ (2.23)

In the higher dimensional case, this solution vector needs to be ordered in a convenient way. In this thesis we use a lexicographic ordering. When the number
of grid points in the first dimension is equal to \( m_1 \) and the second dimension equal to \( m_2 \), we have:

\[
\vec{v} = (v_{0,0}, \ldots, v_{0,m_2}, \ldots, v_{k,0}, \ldots, v_{k,m_2}, \ldots, v_{m_1,0}, \ldots, v_{m_1,m_2}).
\]

Note that in this case, \( \vec{v} \) is a vector of size \( 1 \times (m_1 + 1)(m_2 + 1) \). In general adding a dimension with \( m \) grid points will make the solution vector \( m \) times as large.

### 2.4.2 Finite difference approximations in space

If we want to calculate the value \( V \) of a derivative on an underlying \( X \) driven by the Black-Scholes dynamics, we have to solve the Black-Scholes PDE (2.9). Therefore we approximate the partial derivatives in the PDE by finite differences. On the internal mesh points of the discrete space, the partial derivatives are approximated using neighboring points. For this, we use forward, central or backward schemes. When the grid is non uniform, the finite differences are also non uniform and we define \( \Delta x_i = x_i - x_{i-1} \) for the local differences. As in [62] we define the following well known three different difference formulas for the first derivative:

\[
\begin{align*}
\frac{\partial V}{\partial x}(x_i) &\approx \alpha_{-2} v_{i-2} + \alpha_{-1} v_{i-1} + \alpha_0 v_i, \\
\frac{\partial V}{\partial x}(x_i) &\approx \beta_{-1} v_{i-1} + \beta_0 v_i + \beta_{+1} v_{i+1}, \\
\frac{\partial V}{\partial x}(x_i) &\approx \gamma_0 v_i + \gamma_{+1} v_{i+1} + \gamma_{+2} v_{i+2},
\end{align*}
\]

where,

\[
\begin{align*}
\alpha_{-2} &= \frac{\Delta x_i}{\Delta x_{i-1}(\Delta x_{i-1} + \Delta x_i)}, \quad \alpha_{-1} = \frac{\Delta x_{i-1} - \Delta x_i}{\Delta x_{i-1} \Delta x_i}, \quad \alpha_0 = \frac{\Delta x_{i-1} + 2 \Delta x_i}{\Delta x_i (\Delta x_i + \Delta x_{i-1})}, \\
\beta_{-1} &= \frac{-\Delta x_{i+1}}{\Delta x_i (\Delta x_i + \Delta x_{i+1})}, \quad \beta_0 = \frac{-\Delta x_i + \Delta x_{i+1}}{\Delta x_{i+1} \Delta x_i}, \quad \beta_{+1} = \frac{-\Delta x_i}{\Delta x_{i+1} (\Delta x_i + \Delta x_{i+1})}, \\
\gamma_0 &= \frac{-2 \Delta x_{i+1} - \Delta x_{i+2}}{\Delta x_{i+1} (\Delta x_{i+1} + \Delta x_{i+2})}, \quad \gamma_{+1} = \frac{\Delta x_{i+1} + \Delta x_{i+2}}{\Delta x_{i+1} \Delta x_{i+2}}, \quad \gamma_{+2} = \frac{-\Delta x_{i+1}}{\Delta x_{i+2} (\Delta x_{i+1} + \Delta x_{i+2})}.
\end{align*}
\]

For all the internal grid points, the central scheme can be used and, when a derivative is needed at the boundaries, the forward or backward schemes can be used. For the second derivative with respect to \( x \), we use the central scheme:

\[
\frac{\partial^2 V}{\partial x^2}(x_i) \approx \eta_{-1} v_{i-1} + \eta_0 v_i + \eta_{+1} v_{i+1},
\]

(2.27)
where in this case we have:

\[ \eta_{-1} = \frac{2}{\Delta x_i(\Delta x_i + \Delta x_{i+1})}, \quad -\eta_0 = \frac{-2}{\Delta x_i \Delta x_{i+1}}, \quad \eta_{+1} = \frac{2}{\Delta x_{i+1}(\Delta x_i + \Delta x_{i+1})} \]

The cross derivative can be approximated by a simple combination of the central three point stencil applied to two dimensions. In the problems we encounter in this thesis, no on-sided approximations of the second derivative are needed.

2.4.3 ADI method

The finite difference approximations of a grid of values can be computed by using matrix vector multiplications. These matrix vector multiplications will be used in the time stepping scheme. For stable time stepping, matrix inversion is needed, but this gives rise to some numerical difficulties.

Note that for a difference scheme that uses three neighboring points, this matrix has three diagonals. In higher dimensional problems, these matrices can be very large and, because of the lexicographic ordering, the diagonals can be separated by a large number of zeros. In this case, there are more than three non-zero diagonals which makes the matrix harder to invert. For example if we have a matrix \( F \) which approximates finite differences in two dimensions, we can optimize the structure of this matrix by splitting. ADI stands for Alternating Direction Implicit, and it treats one direction implicitly per step. First the matrix \( F \) is split:

\[ F = F_0 + F_1 + F_2 + F_3, \]

where \( F_0 \) corresponds to the mixed derivatives, and \( F_1, F_2 \) and \( F_3 \) to the derivatives in the first, second and third direction. In the scheme one time step is split in substeps. In every sub step one direction is treated implicitly and the other directions explicitly, next, the other direction is treated implicitly and the rest explicitly and so on. In every sub step only one sub-matrix needs to be inverted, which has small bandwidth, therefore these implicit steps can be done fast. The mixed derivative matrix is not sparse and is therefore not treated in an implicit fashion. For the higher dimensional problems, we use the Hundsdorfer-Verwer scheme [61].
3

Efficient computation of exposure profiles for counterparty credit risk

In this chapter, three computational techniques for approximation of counterparty exposure for financial derivatives are presented. The exposure can be used to quantify so-called Credit Value Adjustment (CVA) and Potential Future Exposure (PFE), which are, as discussed in Chapter 2, of utmost importance for modern risk management in the financial industry. The three techniques all involve a Monte Carlo path discretization and simulation of the underlying entities. Along the generated paths, the corresponding values and distributions are computed during the entire lifetime of the option. Option values are computed by either the finite difference method for the corresponding partial differential equations, or the simulation based Stochastic Grid Bundling Method (SGBM), or by the COS method, based on Fourier-cosine expansions. In this chapter, numerical results are presented for early-exercise options. The underlying asset dynamics are given by either the Black-Scholes or the Heston stochastic volatility model.

---

The contents of this chapter are based on: C.S.L. de Graaf, Q. Feng, B.D. Kandhai and C.W. Oosterlee. Efficient computation of exposure profiles for counterparty credit risk. *International Journal of Theoretical and Applied Finance*, 17(4), 2014
3.1 Introduction

This chapter focuses on two important building blocks of CCR, Expected Exposure (EE) and Potential Future Exposure (PFE). The exposure is defined as the amount of money that may be lost if a counterparty defaults at a particular time, and cannot meet future payments that are agreed upon in the option contract. Because of the growing practical importance, there are recent articles from practitioners [7] and [8] that discuss the computation of the exposure. Since in particular the Heston model is widely used in practice, it is of interest to estimate the EE/PFE under the Heston model.

For European-style derivatives, without any opportunities to exercise contracts before the maturity date, it is obvious that CCR is important. When a counterparty defaults before the contract’s maturity, the investment in the OTC option will be lost and the payoff will not be paid out. In the case of early-exercise options, such as Bermudan options, CCR is also relevant. In this case, although the holder has the right to exercise at multiple moments during the life of the contract, if an option is exercised because of financial distress of a counterparty the return on investment is not as was originally expected, and therefore the option was most likely "mispriced". Recent studies by Klein and Yang [71] further elaborate on this issue.

Here, three numerical methods are presented to keep track of the option values and their distributions during the life of the option contracts. All methods presented contain essentially two elements, a forward sweep for generating future scenarios and a backward sweep to calculate exposures along the generated asset paths. The forward Monte Carlo method generates the asset paths from initial time up to maturity. Along the paths, option values are determined at each exercise time. Because of the complexity of this problem, efficient computation of the option prices is required. The COS Fourier option pricing method may seem a suitable candidate because of its speed and accuracy particularly for Lévy processes, see [39]. Also the finite difference method, approximating solutions to partial differential equations, may be suitable as it typically results in approximate option prices for a grid of underlying values. This feature may be exploited in the EE context, as all grid points can then be used to generate option densities. The recent development of the Stochastic Grid Bundling Method (SGBM), which is a Monte Carlo based method particularly suitable for high-dimensional early-exercise options, in [66] and [67], is another candidate because it also rapidly converges and is accurate.
3.2 Problem Formulation

3.2.1 CVA and Exposure of Bermudan options under Heston’s model

We will present methods for the computation of the exposure of Bermudan options under the Heston stochastic volatility asset dynamics [57], given by

\[
\begin{align*}
    dS_t &= rS_t dt + \sqrt{V_t} S_t dW^1_t, \\
    dV_t &= \kappa (\eta - V_t) dt + \sigma \sqrt{V_t} dW^2_t, \\
    dW^1_t dW^2_t &= \rho dt,
\end{align*}
\]

where \( W^1_t \) and \( W^2_t \) are Wiener processes, correlated by the parameter \( \rho \), \( \kappa \) is the speed of mean reversion parameter in the Cox-Ingersoll-Ross (CIR) process for the variance, \( \eta \) represents the level of mean reversion, and \( \sigma \) is the so-called volatility of the volatility parameter; \( r \) is the risk-free interest rate. The state of the process at time \( t_m \) is denoted by the pair \( (S_m, V_m) \) with \( S_m \) the price of the underlying and \( V_m \) the variance.

The exposure at a future time \( t < T \) is defined as in equation (2.17) in Chapter 2.

A Bermudan option is defined as an option where the buyer has the right to exercise at a set of (discretely spaced) time points. We denote the set of equally-spaced exercise times by

\[
T = \{t_1, t_2, \ldots, t_M\},
\]

where \( M \) denotes the number of exercise times, and the time difference is \( \Delta t \). At the start of the option \( t_0 \), exercise is not allowed.

At each exercise time, the exercise value, given by the payoff function, and the continuation value of the option are compared. The payoff function and the continuation value for the option at time \( t_m \) are, respectively, defined as:

\[
\begin{align*}
    \phi(S_m) &= \max(y(S_m - K), 0) \quad \text{with} \quad y = \begin{cases} 1, & \text{for a call} \\ -1, & \text{for a put} \end{cases}, \\
    c(S_m, V_m, t_m) &= e^{-r \Delta t} \mathbb{E}\left[U(S_{m+1}, V_{m+1}, t_{m+1}) \mid (S_m, V_m)\right],
\end{align*}
\]

where \( U(S_{m+1}, V_{m+1}, t_{m+1}) \) is the option value at time \( t_{m+1} \).

It is assumed that the holder of the option will exercise when the payoff value is higher than the continuation value, and then the contract terminates. At maturity \( t_M \), the option value is equal to the payoff value.
The following recursive scheme can be set-up to price a Bermudan option:

\[
U(S_m, V_m, t_m) = \begin{cases} 
\phi(S_M) & \text{for } m = M; \\
\max \left[ c(S_m, V_m, t_m), \phi(S_m) \right] & \text{for } m = 1, 2, \cdots, M - 1; \\
c(S_0, V_0, t_0) & \text{for } m = 0.
\end{cases}
\] (3.5)

By definition the exposure of an option equals zero once the option is exercised; otherwise, the exposure is equal to the continuation value of the option. The Bermudan option exposure at a future time \( t_m \) can thus be formulated as:

\[
E(t_m) = \begin{cases} 
0, & \text{if exercised, } m = 1, 2, \cdots, M - 1; \\
c(S_m, V_m, t_m), & \text{if not exercised, } m = 1, 2, \cdots, M - 1.
\end{cases}
\] (3.6)

In addition, \( E(t_0) = c(S_0, V_0, t_0) \) and \( E(t_M) = 0 \).

The key point of calculating the exposure at time \( t_m \) is to determine the continuation value.

### 3.3 Numerical Methods to Compute Expected Exposure

In this section, three methods are presented to compute the expected exposure for Bermudan options under the Heston dynamics. All three methods can also be used to simply calculate the value of a Bermudan option at time \( t_0 \). In combination with Monte Carlo forward path simulation, and based on the same common technique, they can be extended to value the exposure of Bermudan options.

#### 3.3.1 General pricing approach

The market state depends on two random variables, \((S_m, V_m)\), at time point \( t_m \) and therefore the exposure \( E(t_m) \) is also a stochastic variable. An option value distribution at future time points can be computed by generating scenarios, and therefore a Monte Carlo simulation is employed.

For the Monte Carlo simulation the highly accurate Quadratic Exponential (QE) scheme [5] is used here to generate the Heston stochastic volatility asset paths. Starting from simulated underlying values and variances, the exposures can be calculated by a backward valuation procedure. At each path, for each exercise time, the continuation value is calculated and compared to the exercise value on the path. When the exercise value is higher than the continuation value, the option is exercised at this path and the exposure for later time points is set to zero. At every time point the resulting exposure values for all paths generate a distribution, as is illustrated in Figure 3.1.
3.3. NUMERICAL METHODS TO COMPUTE EXPECTED EXPOSURE

The essential technique of modeling the exposure of Bermudan options can be presented as follows:

- Generating scenarios/paths by Monte Carlo simulation;
- Calculate continuation/option values and the exercise values to decide whether to exercise or not;
- Set the exposure at each path as the continuation value if the option is not exercised; otherwise the exposure equals 0;
- Compute the empirical distribution of the exposure at each exercise time;
- Calculate $EE$, $PFE_{2.5\%}$ and $PFE_{97.5\%}$ as defined in equation (2.17) and (2.22) in Chapter 2.

In the remainder of this section, we will describe three methods to calculate the required continuation/option values at the simulated paths.

3.3.2 The Finite-Difference-Monte-Carlo method

As outlined in the introduction, an often used option pricing technique is the finite difference method [97]. The method calculates option prices based on the option pricing partial differential equation, for an entire grid of underlying values and can therefore also easily be used to compute the sensitivities (for example, the derivatives of the option prices w.r.t. the asset prices). In the scope of this research, the resulting grid of option values facilitates to determine distributions of
option values at different time points. The method is called the *Finite-Difference-Monte-Carlo* (FDMC) method. Solving the Heston PDE to price European or American options is extensively studied, see, for example, [62],[55] and [56].

As described in Section 3.3.1, option values for all realized market states need to be computed. For European options it is well-known that option value \( U \) satisfies the PDE:

\[
\frac{\partial U}{\partial t} = \mathcal{A}U, \tag{3.7}
\]

where in the Heston case the spatial differential operator \( \mathcal{A} \) is given by

\[
\mathcal{A}U = \frac{1}{2} \sigma^2 v \frac{\partial^2 U}{\partial v^2} + \rho \sigma v \frac{\partial^2 U}{\partial s \partial v} + \frac{1}{2} v^2 \frac{\partial^2 U}{\partial s^2} + (\kappa [\theta - v]) \frac{\partial U}{\partial v} + rs \frac{\partial U}{\partial s} + \frac{\partial U}{\partial t} - rU. \tag{3.8}
\]

For American options, a linear complementarity problem is solved. Using the payoff function (3.3), the option value in this case satisfies:

\[
\frac{\partial U}{\partial t} \geq \mathcal{A}U, \tag{3.9a}
\]

\[
U(s) \geq \phi(s), \tag{3.9b}
\]

\[
(U - \phi(s))(\frac{\partial U}{\partial t} - \mathcal{A}U) = 0, \tag{3.9c}
\]

with one equality sign in either (3.9a) or (3.9b). Note that for each \( t \in [t_0, T] \) the option can be exercised. A discrete version results in the pricing of a Bermudan-style option.

In this research the Brennan-Schwartz [15] algorithm is used, which is a well-known technique from the literature. At each exercise time, this method first solves inequality (3.9a) as an equality, after which the option value is taken to be the maximum of this value and the exercise value.

The boundary conditions used are stated in Table 3.1. Note that at the \( V = 0 \) boundary a so-called degenerated boundary condition is imposed which is obtained by substituting \( V = 0 \) in (3.7).

The schemes used for discretizing (3.8) in asset and variance directions are second-order accurate central schemes or one-sided second-order schemes where needed at boundaries.

The option price is computed backwards in time, from maturity \( T \) back to time \( t_0 \). The equations that need to be solved as a result of the finite difference discretization are linear systems of equations. Such a system of equations can be represented as a matrix-vector problem where the operators are represented...
3.3. NUMERICAL METHODS TO COMPUTE EXPECTED EXPOSURE

<table>
<thead>
<tr>
<th>Boundary</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S \to \infty$</td>
<td>$U = 0$</td>
</tr>
<tr>
<td>$S = 0$</td>
<td>$U = K$</td>
</tr>
<tr>
<td>$V \to \infty$</td>
<td>$\frac{\partial^2 U}{\partial v^2} = 0$</td>
</tr>
<tr>
<td>$V = 0$</td>
<td>$\frac{\partial U}{\partial t} - r U + r S \frac{\partial U}{\partial S} + \kappa \eta \frac{\partial U}{\partial \eta} = 0$</td>
</tr>
</tbody>
</table>

Table 3.1: Heston model boundary conditions for a European put option.

by matrices and the (intermediate) solutions by vectors. For the time integration scheme a particular Alternating Direction Implicit (ADI) scheme, namely the Hundsdorfer-Verwer scheme is employed, which exhibits second-order convergence in time. Next to that, due to the splitting of matrices, it involves the inversion of tridiagonal matrices while in general for (fully) implicit schemes, the matrices are not tridiagonal and may have several non-zero diagonals. For more details we refer to [62].

To ensure that all paths are contained in the computational domain of the finite difference technique, the boundaries $S_{\text{max}}$ and $V_{\text{max}}$ are prescribed such that all Monte Carlo path values at all time points are contained.

The paths attain values that are, most likely, not grid points of the finite difference grid. From this grid, specific option values are determined by interpolation. Because this interpolation may introduce errors at every point, second-order accurate spline interpolation is used.

In general, only a small part of the discretized grid is a region of interest, therefore one can concentrate grid points in that region. This is done by stretching the grid so that a non-uniform grid results, applied in the variance as well as in the asset dimension [55]. As here we need option values at each exercise time for many combinations of spot and variance values, non-uniformity is even more important when we compute exposure.

Because tests show that the impact of the spot dimension on the error is highest, the non-uniform grid in [55] is slightly adjusted. The grid employed is a combination of a uniform and a non-uniform grid. An interval $[S_{\text{left}}, S_{\text{right}}]$ containing $K$ is introduced in which the mesh is uniform. We choose:

$$S_{\text{left}} = \lambda K \quad \text{and} \quad S_{\text{right}} = K,$$

(3.10)

where $\lambda \in [0.3, 0.7]$ can be chosen depending on the quantity that needs to be computed (PFE or EE). An accurate computation of EE requires accurate pricing around the mean, which implies a high value of $\lambda$, whereas for accurate computation of quantiles, also an accurate computation of extreme values is needed for which a smaller value of $\lambda$ should be chosen. So, when we compute the PFE, the dense region is shifted towards the outer regions of the domain.
Outside the interval \([S_{\text{left}}, S_{\text{right}}]\) the grid follows a hyperbolic sine function with:

\[
\begin{align*}
\xi_{\text{min}} &= \xi_0 = \sinh^{-1}\left(\frac{S_{\text{left}}}{d_1}\right), \\
\xi_{\text{int}} &= \frac{S_{\text{right}} - S_{\text{left}}}{d_1}, \\
\xi_{\text{max}} &= \xi_{m_1} = \xi_{\text{int}} + \sinh^{-1}\left(\frac{S_{\text{max}} - S_{\text{right}}}{d_1}\right),
\end{align*}
\]

where \(d_1\) is a scaling parameter, and \(\xi_{\text{min}} < 0 < \xi_{\text{int}} < \xi_{\text{max}}\). Now, to construct a grid of \(m_1 + 1\) points, the non-uniform adjustment \(0 = s_0 < s_1 < \ldots < s_{m_1} = S_{\text{max}}\) can be constructed via a uniform grid of \(m_1 + 1\) points between \(\xi_{\text{min}}\) and \(\xi_{\text{max}}\):

\[
\xi_{\text{min}} = \xi_0 < \xi_1 < \ldots < \xi_{m_1} = \xi_{\text{max}}
\]

and the function \(g\):

\[
g(\xi_i) = s_i, \quad i = 0, \ldots, m_1, \text{ where}
\]

\[
g(\xi_i) = \begin{cases} 
S_{\text{left}} + d_1 \sinh(\xi_i) & \text{if } \xi_{\text{min}} \leq \xi_i < 0, \\
S_{\text{left}} + d_1 \xi_i & \text{if } 0 \leq \xi_i \leq \xi_{\text{int}}, \\
S_{\text{right}} + d_1 \sinh(\xi_i - \xi_{\text{int}}) & \text{if } \xi_{\text{int}} < \xi_i \leq \xi_{\text{max}}.
\end{cases}
\]

Smaller values of \(d_1\) result in a lower density in \([S_{\text{left}}, S_{\text{right}}]\), whereas higher values of \(d_1\) will result in a higher density of grid points in this interval.

For the \(V\) direction more points at the boundary \(V = 0\) are desired and for larger values of \(V\) the mesh can be less dense. Let \(m_2\) be the number of points to be considered and \(d_2\) another scaling parameter. Define equidistant points \(v_0 < v_1 < \ldots < v_{m_2}\), given by \(v_j = j \cdot \Delta v\), with \(\Delta v = \frac{1}{m_2} \sinh^{-1}\left(\frac{V_{\text{max}}}{d_2}\right)\), for \(j = 0, \ldots, m_2\). Now the grid \(0 = v_0 < v_1 < \ldots < v_{m_2} = V_{\text{max}}\) is defined by:

\[
v_j = d_2 \sinh(v_j), \quad j = 0, \ldots, m_2.
\]

These grids are smooth in the sense that there are real-valued constants \(C_0, C_1\) and \(C_2\) such that:

\[
C_0 \Delta \xi \leq \Delta s_i \leq C_1 \Delta \xi \quad \text{and} \quad |\Delta s_{i+1} - \Delta s_i| \leq C_2 (\Delta \xi)^2.
\]

When the finite difference method is used to price a single option, only a single grid point at initial time is used. The FDMC method however uses a large portion of the grid points for option pricing at all exercise times which makes this method computationally attractive.

### 3.3.3 Stochastic-Grid-Bundling method

The Stochastic-Grid-Bundling Method (SGBM) is a Monte Carlo method that combines regression, bundling and simulation. It was proposed by Jain and Oosterlee in [66, 67] for pricing multi-dimensional Bermudan options under Black-Scholes dynamics. The SGBM method generates a direct estimator, a lower
bound for the option value, as well as an optimal early-exercise policy. Here, we extend the SGBM method (applied to the Black-Schole model in [66]) towards the Heston model, and exposure distributions along the time horizon are naturally obtained.

Suppose we deal with a Bermudan option with tenor $T$ and $M$ exercise dates. First a stochastic grid is generated, i.e. we generate $H$ paths of the underlying under the Heston model. It is easy to see that the option value at time $t_M = T$ is equal to the corresponding payoff value, which gives us the initial setting for the SGBM method at each path.

At time $t_k$, $k = M - 1, \ldots, 1$, these paths are clustered into $\beta$ bundles, based on their stock and variance values. The bundle set at time $t_m$ is denoted by $\{B^p, m\}_{p=1}^\beta$. Paths within the same bundle are assumed to share some common properties. We adapt the so-called multi-dimensional recursive bifurcation bundling method [67] by means of a rotation.

Within each bundle, basis functions $(g_k(S,V))_{k=0}^B$ are defined for the regression of option values\(^1\). The essential idea in SGBM is that, for paths in the $p$-th bundle $B^p, m$ at time $t_m$, a set of coefficients $(\alpha^p, m)^B_{k=0}$ exists, so that for the option values of these paths at time $t_{m+1}$, the following relationship holds

$$U(S_{m+1}, V_{m+1}, t_{m+1}) \approx \sum_{k=0}^B \alpha^p, m_k g_k(S_{m+1}, V_{m+1}).$$ \hspace{1cm} (3.11)$$

When the option values $U(S_{m+1}, V_{m+1}, t_{m+1})$ at the stochastic paths are determined, the coefficient set $(\alpha^p, m)^B_{k=0}$ can be obtained by regression. Equation (3.11) can be substituted into (3.4) which gives us:

$$e(S_m, V_m, t_m) = e^{-rT} E \left[ U(S_{m+1}, V_{m+1}, t_{m+1}) \bigg| (S_m, V_m) \right]$$

$$\approx e^{-rT} E \left[ \sum_{k=0}^B \alpha^p, m_k g_k(S_{m+1}, V_{m+1}) \bigg| (S_m, V_m) \right]$$

$$= e^{-rT} \sum_{k=0}^B \alpha^p, m_k E \left[ g_k(S_{m+1}, V_{m+1}) \bigg| (S_m, V_m) \right]$$

$$= e^{-rT} \sum_{k=0}^B \alpha^p, m_k f_k(S_m, V_m).$$ \hspace{1cm} (3.12)$$

When the functions $(f_k(\cdot, \cdot))_{k=0}^B$ are known, the continuation values at time $t_m$ can be computed, and, subsequently, the option values at time $t_m$ can be obtained with the scheme in (3.5). At time $t_0$, we deal with one bundle, as

\(^1\)For the definition and the choice of basis function see [73].
CHAPTER 3. COMPARATIVE STUDY: 2D HESTON MODEL

all paths originate from \((S_0, V_0)\), and the option value at time \(t_0\) is equal to the continuation value \(c(S_0, V_0, t_0)\). In this way, option values are calculated backward in time from \(t_M\) to \(t_0\). By (3.6), the exposure at each path along the time horizon is calculated as a by-product. We are thus able to determine the empirical exposure distribution at each time point for the calculation of EE and PFE.

We choose the basis functions \(\{g_k(S, V)\}\) such that analytic formulas for their expectations are available. Analytic formulas of these expectations bring exact information into the recursive procedure. Under the Heston dynamics, the basis functions are chosen as

\[
g_k(S, V) := (\log(S))^k,
\]

where \(k = 0, \ldots, B\).

When \(k = 0\), the basis function is the constant; when \(k > 1\), the expectation of \(g_k\) is the \(k\)-th moment of \(\log(S)\).

There is a well-known relationship between the moments of the model and its characteristic function. The joint characteristic function of Heston’s model is available (see, for example, [41]) and thus analytic formulas for the expectation \(f_k(\cdot, \cdot)\) can be derived. Note, however, that the stochastic grid is based on \((S, V)\)-values; only the regression is based on log-variables.

The optimal exercise strategy is determined by comparing the immediate exercise value and the continuation value, and the exercise decision is made when the immediate exercise value is highest. We store the strategy and the corresponding realized cash flows at each path during the backward procedure. The option value can be calculated as the mean of the discounted cash flow as in [73].

The SGBM method has some advantages compared to the Longstaff-Schwartz method, although both are based on regression. The Longstaff-Schwartz method only uses the ‘in-the-money’ paths to get the optimal stopping strategy and the corresponding cash flow. In SGBM, all paths are used and the optimal stopping strategy is merely a by-product. By applying bundling, the approximation of the linear coefficients can be optimized locally. Furthermore, information from the model dynamics is included by application of the analytic formulas for the expectation of the basis functions, whereas the Longstaff-Schwartz method only employs the dynamics in the path generation.

3.3.4 The COS-Monte-Carlo method

In the third computational method, we combine the generated stochastic MC grid with the COS method, introduced in [41]. Based on the same stochastic grid as, for example, in Section 3.3.3, the COS method is used for the calculation of the continuation values at each path along the time horizon. We call this combined method the COS-Monte-Carlo (CMC) method.
3.3. NUMERICAL METHODS TO COMPUTE EXPECTED EXPOSURE

As in [41], we work in the log-domain, denoted by \((x, u) := (\log(S), \log(V))\). Suppose that the path values \((x_m, u_m)\) at time \(t_m\) are known. We can write the joint density function at \(t_{m+1}\), conditioned on values at \(t_m\), as

\[
f_{x,u}(x_{m+1}, u_{m+1}|x_m, u_m) = f_{x|u}(x_{m+1}|x_m, u_{m+1}, u_m) \cdot f_u(u_{m+1}|u_m),
\]

where \(f_{x|u}(\cdot)\) is the conditional log-stock density, and \(f_u(\cdot)\) the conditional log-variance density. Notice that here we have \(f_u(u_{m+1}|x_m, u_m) = f_u(u_{m+1}|u_m)\).

The continuation value defined in (3.4) at time \(t_m\) is the expectation of the option value at time \(t_{m+1}\) w.r.t. the joint density function. One can choose a proper integration range \([a, b] \times [a_v, b_v]\) in the log-stock domain and the log-variance domain, so that the integral can accurately be approximated. We refer to [41] for details on the definition of this range based on the initial state \((x_0, u_0)\).

In this chapter, we are not only concerned with the option value at time \(t_0\), but also with continuation values along the time axis. To assure accuracy, we need a common integration range which is sufficiently large for all paths at each exercise time. We define it as

\[
[a, b] := \bigcup_{h, m=1}^{H, M} [a^h, b^h, m, b^h, m],
\]

\[
[a_v, b_v] := \bigcup_{h, m=1}^{H, M} [a_v^h, b_v^h, m, b_v^h, m],
\]

where \([a^h, m, b^h, m]\) and \([a_v^h, b_v^h, m]\) are the ranges for the log-stock and log-variance domains, respectively, for the \(h\)-th path at \(t_m\), by the suggestions in [39] and [41].

The integration can now be written as follows:

\[
e(x_m, u_m, t_m) \approx e^{-r t} \int_{a_v}^{b_v} f_u(u_{m+1}|u_m) \times \left[ \int_a^b U(x_{m+1}, u_{m+1}, t_{m+1}) f_{x|u}(x_{m+1}|x_m, u_{m+1}, u_m) dx_{m+1} \right] du_{m+1}.
\]

An analytic formula for the log-variance density \(f_u(\cdot)\) is available (see [41]), and the conditional log-stock density \(f_{x|u}(\cdot, \cdot)\) can be recovered from the corresponding characteristic function [41] by applying the COS expansion. The recovered density of the log-stock process from the characteristic function [41] is given by

\[
f_{x|u}(x_{m+1}|x_m, u_{m+1}, u_m) \approx \frac{2}{b-a} \times \sum_{n=0}^{N-1} \Re \left\{ \frac{n\pi}{b-a} \right\} e^{i n\pi \frac{x_m-u_m}{b-a}} \cos \left( \frac{n\pi}{b-a} (x_{m+1} - x_m) \right).
\]
CHAPTER 3. COMPARATIVE STUDY: 2D HESTON MODEL

where $\sum'$ indicates that the first term is multiplied by $\frac{1}{2}$; $\text{Re}(\cdot)$ returns the real part of the value; the function $\Phi$ is defined as $\Phi \left( \frac{n\pi}{b-a}; \nu_{m+1}, \nu_m \right) := \Phi \left( \frac{n\pi}{b-a}; 0, \nu_{m+1}, \nu_m \right)$, which is the characteristic function of the log-stock process.

We use the Gaussian-quadrature rule [41] for the approximation of the outer integral in (3.16). The log-variance integral range is discretized on a grid, denoted by $\{\xi_j, \hat{\xi}_0 = a, \hat{\xi}_j = b\}_{j=0}^{J}$. The characteristic function of the log-stock process at time $t_{m+1}$ conditioned on the log-variance and $(\nu_m, \nu_m)$ is denoted by $\Phi \left( \omega; \nu_m, \xi_j, \nu_m \right)$.

By interchanging the inner integral and the summation obtained by the COS expansion, the continuation value can be written as

$$c(x_m, \nu_m, t_m) \approx e^{-r\Delta t} \sum_{j=0}^{N-1} w_j f_u(\xi_j | \nu_m) \times \sum_{n=0}^{N-1} A_{n,j}(t_{m+1}) \text{Re} \left\{ \Phi \left( \frac{n\pi}{b-a}; \xi_j, \nu_m \right) e^{i \frac{n\pi x_m-a}{b-a}} \right\},$$

(3.18)

where the $w_j$ are the weights of the quadrature nodes $\xi_j$, $j = 0, 1, \ldots, J - 1$; $A_{n,j}(t_{m+1})$ is the Fourier-cosine coefficient defined as follows:

$$A_{n,j}(t_{m+1}) = \frac{2}{b-a} \int_a^b U(x_{m+1}, \xi_j, t_{m+1}) \cos \left( \frac{n\pi x_{m+1}-a}{b-a} \right) dx_{m+1} \right].$$

(3.19)

At time $t_M$, the values of the coefficients $A_{n,j}(t_M)$ can easily be obtained as the option value at time $t_M$ equals the payoff value. The expression for $A_{n,j}(t_M)$ becomes

$$A_{n,j}(t_M) = \begin{cases} G_n(a, 0) & \text{for a put,} \\ G_n(0, b) & \text{for a call,} \end{cases}$$

(3.20)

where the $G_n$-functions are the cosine coefficients of the payoff function, given as:

$$G_n(l, u) = \frac{2}{b-a} \int_l^u \phi(y) \cos \left( \frac{n\pi y-a}{b-a} \right) dy.$$

(3.21)

Fortunately, for specific payoffs, an analytic formula for $G_n(l, u)$ is available (see [41]).

From maturity $t_M$ a backward recursive calculation can be used to obtain all coefficients $A_{n,j}(t_m)$.

For the computation of the EE, however, interpolation is needed for the calculation of the continuation value at each time step for each path. We use interpolation to reduce the calculation costs, because the calculation of the function $\Phi \left( \frac{n\pi}{b-a}; \xi_j, \nu_m \right)$ takes significant CPU time. Instead of calculating the value of
function $u_m$ for each path and time step $t_m$, we compute an $N \times J \times J$ matrix $\Phi$ with elements $\left\{ \Phi \left( \frac{n \pi}{\beta}; \varsigma_j, \varsigma_p \right) \right\}$. For a fixed $p$, we extract the corresponding 2D slice from the 3D matrix to calculate the continuation value $c(x_m, \varsigma_p, t_m)$ at time $t_m$.

One can easily determine the vector $\{c(x_m, \varsigma_p, t_m)\}_{p=0}^{J}$, storing the continuation values of each path on the variance grid by matrix calculation. As in the FDMC method, by spline interpolation, an accurate continuation value $c(x_m, u_m, t_m)$ at each path can be obtained.

After the calculation of the continuation value, the exposure can easily be determined by applying the formulas in Section 3.2.1.

Compared to the COS method for pricing option values, the CMC method is significantly slower when the number of MC paths is high. One reason is that, at each exercise time, an additional calculation of the continuation value is performed, for which interpolation is required for each path. Next to that, as we need to choose a wider integration range to assure accuracy for EE and PFE, we also need a large number of Fourier-cosine terms and variance grid points to get converged results, which has a significant impact on the computational speed.

At the same time, the CMC method maintains the very high accuracy of the COS method. The errors due to the truncated integration ranges, the quadrature and the propagation error have been discussed in [41]. The error of the spline interpolation on the variance grid is small when $J$ is sufficiently large as the continuation value is a continuous function of the variance. Because of the high accuracy, we will use the results of the CMC method as reference values in the discussion of the numerical results.

3.4 Numerical Results

In this section we start with an assessment of the impact of stochastic volatility on the EE and PFE profiles. Next, we consider a detailed analysis of the convergence and accuracy of the methods by means of numerical experiments.

As there are no exact values available for the exposure of Bermudan put options under the Heston stochastic volatility model, we will use the converged results of the COS method as reference values. As mentioned in Section 3.3.4, the COS method is a highly accurate method for pricing Bermudan options. When valuing the exposure, the high accuracy is maintained as long as the integration range is properly chosen (see Section 3.3.4). We reduce the impact of Monte Carlo noise in the comparative analysis by using $10^5$ paths.

To investigate the proposed three methods, three different sets of parameters

---

2The convergence of the COS method has been discussed in [41], and we will set the number of Fourier terms to $N = 2^9$ and the number of the variance grid points to be $J = 2^9$. 

---
are tested, see Table 3.2. These test cases were used recently in [41], [56] and [100] and reference values are thus available for individual option prices. Moreover, in these test cases we stress the parameters of the stochastic volatility process by considering different levels for the initial variance, the mean-reversion parameters, vol-of-vol and correlation parameters. These parameters are chosen such that in Tests A and C the well-known Feller condition: $2\kappa\eta > \sigma^2$ is satisfied, while in Test B it is not. Apart from the different settings for the model parameters, we consider different maturities, interest rates and moneyness levels.

<table>
<thead>
<tr>
<th></th>
<th>Test A</th>
<th>Test B</th>
<th>Test C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spot ($S$)</td>
<td>10</td>
<td>100</td>
<td>9</td>
</tr>
<tr>
<td>Strike ($K$)</td>
<td>10</td>
<td>100</td>
<td>10</td>
</tr>
<tr>
<td>Interest ($\rho$)</td>
<td>0.04</td>
<td>0.04</td>
<td>0.10</td>
</tr>
<tr>
<td>Exercise Times</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Initial Vol ($\sqrt{v_0}$)</td>
<td>0.5745</td>
<td>0.1865</td>
<td>0.2500</td>
</tr>
<tr>
<td>Tenor ($T$)</td>
<td>0.25</td>
<td>0.25</td>
<td>1.00</td>
</tr>
<tr>
<td>Mean Reversion ($\kappa$)</td>
<td>0.80</td>
<td>1.15</td>
<td>5.00</td>
</tr>
<tr>
<td>Mean Var ($\eta$)</td>
<td>0.3300</td>
<td>0.0348</td>
<td>0.1600</td>
</tr>
<tr>
<td>Vol of Var ($\sigma$)</td>
<td>0.700</td>
<td>0.459</td>
<td>0.900</td>
</tr>
<tr>
<td>Correlation ($\rho$)</td>
<td>0.10</td>
<td>-0.64</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Table 3.2: Parameter sets for Test A, B and C.

### 3.4.1 Comparison of Black-Scholes to Heston to assess impact of stochastic volatility on exposure

Here it is shown that stochastic volatility clearly has an impact on exposure profiles. It is most significant for the PFE$_{97.5\%}$ quantile in the tests considered. We restrict the analysis to Tests A and B because in Test C the mean reversion level is not equal to the initial variance and thus it is not clear which level to use for the variance in the Black-Scholes model. In Figure 3.2, the results are plotted for the parameters from Tests A and B.

In general, independent of the underlying dynamics, the plots show that the EE starts at the initial option value, after that, the level drops because of the early exercise possibility. The PFEs also start at the initial option value, because at this stage there is no uncertainty, i.e. the minimum value for which the probability is higher than a specific benchmark equals this initial value. Starting from $t = 0$, PFE$_{2.5\%}$ drops to zero soon while PFE$_{97.5\%}$ is always higher than the EE. Due

\footnote{It is known that when the Feller condition is not satisfied, the variance process can become zero and numerical methods can suffer from this issue.}
3.4. NUMERICAL RESULTS

Figure 3.2: EE and PFE profiles under the Black Scholes and Heston model, differences are significant for 97.5% PFE.

Figure 3.3: Distribution of option values under Heston and Black-Scholes for Test B at exercise time 17 of 30. The PFE_{97.5\%} is shaded in red (BS) or blue (Heston) respectively, in the right plot the axes are changed to make the boundary more clear.

When the results for Black-Scholes are compared to Heston, one can conclude that the most significant difference is for PFE_{97.5\%}, in both cases. The difference in PFE_{97.5\%} is a factor 10 times larger than the difference for EE and PFE_{2.5\%}. Intuitively this makes sense, due to the fact that the mean reversion level is equal to the constant variance level in the Black-Scholes model, the EE is not heavily affected. However, since the volatility is stochastic, extreme cases may
occur more frequently (with the parameters chosen), resulting in fatter tails of
the distribution that have a significant impact on PFE.

The early exercise value also depends on the volatility so that for any path,
there is a different exercise value. From Figure 3.3 it can be seen that the dis-
tribution computed under the Black-Scholes dynamics is chopped off at a specific
maximum option value, whereas the distribution under the Heston dynamics has
a smoothly varying tail. The mass that is originally in the cut-off tail in the
Black-Scholes case is here located at the left-side boundary.

Although these results show that stochastic volatility has an impact on the
exposure profiles, a more rigorous analysis, based on market calibrated parame-
ters instead of model parameters, is presented in Chapter 6. Here instead we will
focus on the accuracy and numerical convergence of our proposed methods.

For the analysis of the accuracy and convergence of the three proposed meth-
ods, we concentrate on Tests B and C. The results for EE and PFE_{97.5\%}
obtained for the different methods are shown in Figure 3.4 and Table 3.3 (the results for
PFE_{2.5\%} are not shown in the table). The methods are tested on a set of
$H = 10^5$ generated Monte Carlo paths. In the following subsections, the convergence and

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Error & Measurement & Test B & Test C \\
\hline
EE & $||\text{FD} - \text{COS}||_2$ & 5.89e-03 & 1.90e-03 \\
\hline
 & $||\text{SGBM} - \text{COS}||_2$ & 8.06e-03 & 1.49e-03 \\
\hline
PFE_{97.5\%} & $||\text{FD} - \text{COS}||_2$ & 1.41e-02 & 2.71e-03 \\
\hline
 & $||\text{SGBM} - \text{COS}||_2$ & 4.59e-03 & 3.94e-03 \\
\hline
\end{tabular}
\caption{Relative $L_2$ difference between two methods for $10^5$ paths with the CMC method as benchmark.}
\end{table}
3.4. NUMERICAL RESULTS

error behavior of the FDMC method and SGBM is discussed.

3.4.2 Error FDMC

The error for pricing options with the finite difference method for the Heston PDE is extensively studied, see, for example, [62]. The error is mainly introduced near the boundaries, but can be controlled by a combination of a large number of grid points and the use of a non-uniform grid. In all finite difference computations the grids are non-uniform, as discussed in Section 3.3.2. The free parameter $\lambda$ in (3.10) to determine the region $[4K, K]$ in the spot direction is determined depending on the quantity that is being measured: For PFE a smaller value is desired, whereas for EE the value is larger, in any case $\lambda \in [0.3, 0.7]$. The variance grid is very dense around the $V = 0$ boundary, independent of the measured quantity. The number of grid points in spot ($S$) and variance ($V$) directions are denoted by $m_1$ and $m_2$, respectively. By experiment we know that the numerical error is dominated by the error in spot direction, and therefore the number of grid points in the $S$ direction is chosen as $m_1 = 2m_2$. With this fixed ratio, the decay in error is measured by decreasing a generic measure $\Delta s$ defined as $\Delta s := \frac{1}{m_2}$. If we decrease $\Delta s$ by increasing the number of grid points, the numerical convergence is second-order when we price a single Bermudan option in Test C, whereas it is almost second-order for Test B, see Figure 3.5. The grid is chosen to be very dense in the region of the initial market parameters $(S_0, V_0)$, and the price is extracted from the grid by accurate spline interpolation.

When EE and PFE$\theta$ are computed, multiple prices at each exercise time are needed. In this case, interpolation is needed for each path and exercise time which is expected to have an impact on the error. To investigate the scale of this

![Figure 3.5](image-url)
error the same convergence tests are done as in the single option case. In this case the finite difference solution is compared to the semi-analytic CMC method described in Section 3.3.4. The same random scenarios are used for computing the EE for the CMC and the FDMC methods. As shown in Figure 3.6, in both

Tests B and C the convergence of the error is similar for EE as it is for pricing a single option. The decrease of the error is of second-order in the number of grid points in Test C and almost second-order in Test B.

For PFE, in Test B the convergence from the start is only first-order. In this test the Feller condition not satisfied. In Test C, the convergence tends to get smaller than 1 after reaching $10^{-2}$. The mesh used for PFE has a dense region around the strike, whereas for a PFE computation the strike region is generally not of highest relevance, and therefore the convergence of this measure is slower. To enhance the accuracy of the PFE, the non-uniform grid in $S$ direction can be adjusted (which we leave for later study). Because PFE and EE are mostly computed independently, a conclusion is that measuring PFE or EE would imply using two different grids.

The convergence with respect to $\Delta t$ is not presented in this research because tests show that the error is dominated by the spatial error.

### 3.4.3 Error SGBM

Here we focus on the convergence of SGBM regarding the option value, the EE and PFE. We use five basis functions (including the constant) defined in Equation (3.13). In the recursive bifurcation bundling method, an essential property is that the number of bundles must be of the form $4^j, j = 0, 1, 2, \ldots$, for details.
3.4. NUMERICAL RESULTS

we refer to [67]. The bundling scheme is slightly adapted to deal with the two-dimensional Heston dynamics.

In the tests, a large number of paths $H = 10^5$, and bundles $\beta = 4^4$ are chosen.

In Section 3.3.3, there are two ways of calculating the option value at time $t_0$. One way is to estimate the coefficient set over all paths at time $t_1$ and to apply regression at time $t_0$ (the so-called direct estimator); the other is to store the optimal strategy and take the mean of the discounted cash flow (the so-called path estimator):

- The results calculated directly from the set of Monte Carlo paths is called direct estimator results;
- The results calculated by the second set of paths, but with the coefficients from the first set of paths is called the path estimator.

The numerical results for the path estimator should be similar to the results for the direct estimator. Table 3.4 presents the difference between the direct estimator and the path estimator for EE and PFE. Again the error is measured in the relative $L_2$ norm. We can see that the difference between the two results is only of order $10^{-3}$.

<table>
<thead>
<tr>
<th></th>
<th>EE</th>
<th>PFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test B</td>
<td>2.85e-03</td>
<td>6.30e-03</td>
</tr>
<tr>
<td>Test C</td>
<td>1.95e-03</td>
<td>5.93e-03</td>
</tr>
</tbody>
</table>

Table 3.4: The difference between the direct estimator and path estimator for EE and PFE. The number of bundles equals $4^4$ and the number of paths is $10^5$.

The option value is the maximum value obtained among all possible stopping rules, indicating that the option value calculated by the “optimal” strategy will be less than or equal to the real option value. This provides a criterion for convergence. The result calculated by the optimal strategy will be the lower bound of the Bermudan option value.

We examine the convergence of the Bermudan option value w.r.t. the number of bundles for SGBM. The tests are done for ten simulations, and the results are presented in Figure 3.7. We take the regression results of the direct estimator and the results of the optimal strategy of the path estimator for comparison. As we can see, in both Tests B and C, the results of the path and the direct estimator resemble each other better when the number of bundles increases. The two results are very close to the COS reference value for $\beta = 4^4$, see Table 3.5.

In addition to the convergence of the Bermudan option value, we examine the convergence of EE and PFE in Figure 3.8. The results of the CMC method is used as the reference value. With the same set of $10^5$ generated paths, we increases
CHAPTER 3. COMPARATIVE STUDY: 2D HESTON MODEL

Figure 3.7: Convergence with respect to the number of bundles $\beta = 4^j$, for the Bermudan option value; the total number of paths $H = 10^5$; the reference value in Test B equals 3.2066 and in Test C 1.4990. The red dashed line is the direct estimator, blue is the COS reference value, and the black dashed line the path estimator.

<table>
<thead>
<tr>
<th>Test</th>
<th>COS(reference)</th>
<th>Direct estimator(s.e.)</th>
<th>Path estimator(s.e.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test B</td>
<td>3.2066</td>
<td>3.2091(2.8613e-03)</td>
<td>3.1924(1.3768e-02)</td>
</tr>
<tr>
<td>Test C</td>
<td>1.4990</td>
<td>1.4964(7.2086e-04)</td>
<td>1.4926(3.0376e-03)</td>
</tr>
</tbody>
</table>

Table 3.5: The difference between the direct and the path estimator for a Bermudan option value when the number of bundles equals $4^j$. The results are computed via ten simulations (s.e. is standard error).

the number of bundles from 1 to $4^4$ for the calculation of the SGBM method. It shows that the error decreases when increasing the number of bundles. The EE results exhibit a higher accuracy than the PFE$_{97.5\%}$ results, when the number of bundles is equal to $4^4$.

The convergence of the EE and PFE, w.r.t the number of paths, is examined in Figure 3.9. We choose the number of bundles equal to $4^3$, and increase the average number of paths in each bundle. The differences of EE and PFE between the direct estimator and the path estimator are compared. The average number of paths in each bundle is increased from 25 to 2000. It shows that the difference between the path and the direct estimator decreases when the average number of paths in each bundle (i.e. the total number of paths) increases.

These results support the fact that SGBM converges (to the reference values) for Bermudan options, EE and PFE$_{97.5\%}$.
3.5. CONCLUSIONS

In this chapter, three different approaches for computing exposure profiles within the context of counterparty credit risk are presented. The underlying asset exposure is driven by the Heston stochastic volatility model and Bermudan put options are priced. In all three methods scenarios are generated by a Monte Carlo scheme and option values are priced at each path at each exercise time. The pricing procedure is done by either the FDMC method, SGBM or the CMC method.

The CMC method is a combination of the Monte-Carlo method and the COS.
method which can be used for computing exposures. We adapt the COS method to make it more general and thereby applicable to a wide range of possible states \((S_m, V_m)\), while maintaining its high accuracy. This comes at the cost of computational speed (under the Heston dynamics, particularly when the Feller condition is not satisfied). However, considering its high accuracy, this method is used as a benchmark value to analyze the accuracy and convergence of EE and PFE computed by FDMC and SGBM. By using this benchmark, it is shown that the FDMC method, SGBM and CMC method agree for multiple tests.

As a first result, it is shown that the impact of stochastic volatility on exposure profiles is most significant for PFE\(_{0.97.5\%}\). Whereas the distribution computed under the Black-Scholes dynamics suffers from the tail being chopped off at a certain maximum option value boundary, under the Heston dynamics this feature is not present.

Because any finite difference solution generates option values for an entire grid of underlying values, the FDMC method is promising. The computation time in this method is dominated by the computation of the solution on the grids at each exercise time. When these are stored, the EE computation boils down to an interpolation procedure for all paths at each exercise time. By using the COS method as a benchmark, it is shown that the error introduced by the interpolation is negligible. Compared to the CMC method, the error is within the range of \(10^{-3}\). A possible improvement of this method would be the implementation of a non-uniform adaptive grid, which adapts its dense region to the density of the generated paths at the specific exercise time, but this comes at a price in CPU time.

SGBM has been extended to the Heston model for computing exposures. We test the convergence of SGBM w.r.t the number of paths and the number of bundles in several ways. For the two considered tests, the computation of EE and PFE shows to be highly accurate with an error compared to CMC in the order of \(10^{-3}\). SGBM is an efficient Monte Carlo method when valuing exposure distributions along a time horizon.

For the FDMC method, it remains to be checked if we can compute accurate sensitivities and whether the method can be used for higher dimensional models. In higher dimensional models fewer grid points can be used and as such, the interpolation of states on the pricing grid becomes more important. In the next chapter we will investigate the transition from two to three-dimensional models and look into exposure sensitivities.
4

Efficient estimation of sensitivities for counterparty credit risk with the finite difference Monte-Carlo method

As shown in the previous chapter, the FDMC method is an interesting numerical technique to compute exposure profiles. Here, we further assess the applicability of this method to compute exposure profiles and its sensitivities for computationally challenging path-dependent options. We do this in the context of the Black-Scholes as well as the Heston Stochastic Volatility model with and without stochastic domestic interest rate for a wide range of parameters. In the case of fixed domestic interest rate, our results show that FDMC is accurate compared to the semi-analytic COS method and advantageously can compute multiple options on one grid. Even in the higher dimensional case, when stochastic domestic interest rate is included, we show that we can accurately compute exposures and sensitivities of discontinuous one-touch options by using a linear interpolation technique.

CHAPTER 4. SENSITIVITIES UNDER HESTON HULL-WHITE

4.1 Introduction

In the previous chapter we introduced the Finite Difference Monte Carlo (FDMC) method to calculate the exposure profiles of a derivative. This is done for Bermudan put options which have an early exercise feature at preset discrete time points. Similar to [78] and [79], the FDMC method uses the scenario generation from the Monte Carlo method. Option prices are computed on a grid with the finite difference method and option values per path are obtained by interpolation on this grid. The Expected Exposure (EPE) equals the mean of the resulting option price distribution whereas the Potential Future Exposure (PFE) is a quantile of this distribution. In practice, apart from EPE and PFE, the sensitivities to market factors (like spot value, interest rate and volatility) are required for hedging and control of the Counterparty Credit Risk (CCR) of derivatives portfolios.

In this chapter we extend our previous study. We again incorporate the highly relevant skew effect which is dominantly present in the Foreign Exchange (FX) market by choosing the Heston model to drive the underlying FX rate. In the case of constant interest rates, we consider the estimation of first and second-order sensitivities with respect to the spot FX rate. In contrast to the widely used bump-and-revalue method, we propose a path-dependent estimator that is leveraging from the already estimated local sensitivities on the finite difference grid. A rigorous analysis is performed in the case of barrier options which pose severe numerical challenges due to the knock-out feature that results in a discontinuous terminal condition. Similar discontinuities also arise in portfolios with instruments of different maturities, with the possibility of error propagation on the computation grid in time. Therefore we analyze such portfolios specifically in this work. We validate our results by comparing to the Monte Carlo COS method ([88]).

Next, we relax the assumption of a constant domestic interest rate by looking at the three-factor Heston Hull-White model. This implies the use of coarse grids for which the interpolation is vital. We therefore compare exposure quantities for discontinuous One-Touch (OT) options computed by a linear or a spline interpolation. Additionally we discuss the applicability of the path-dependent sensitivities with respect to initial variance and domestic interest rate. Again, the bump-and-revalue method acts as a benchmark.

4.2 Problem Formulation

4.2.1 CVA under the Heston Hull-White Model

In the Heston Hull-White model, the volatility and the domestic interest rate are modeled as a stochastic process such that the volatility smile and interest rate
4.2. PROBLEM FORMULATION

dynamics can be captured. The three-dimensional dynamics are given by:

\[\begin{align*}
    dS_t &= (R_d^1 - r^d)S_t dt + \sqrt{V_t} S_t dW^1_t, \\
    dV_t &= \kappa(\eta - V_t) dt + \sigma \sqrt{V_t} dW^2_t, \\
    dR_d^1 &= \lambda(\theta(t) - R_d^1) dt + \gamma dW^3_t, \\
    dW^i_t dW^j_t &= \rho_{i,j} dt, \quad \text{for } i \neq j \in [1,2,3].
\end{align*}\] (4.1)

where \(r^d\) is the foreign interest rate, \(\kappa\) the mean-reverting speed in the Cox-Ingersoll-Ross (CIR) process for the variance, \(\eta\) the level of the long term mean, \(\sigma\) the so-called volatility of volatility. The domestic interest rate \(R_d^1\) follows the Hull-White SDE, where \(\lambda\) is the mean reverting speed, \(\theta(t)\) the level of the long term mean deduced from the forward curve and \(\gamma\) the volatility of the short rate. The SDEs are coupled by the correlated Wiener processes. Note that the same dynamics will hold for an equity derivative with stochastic interest rate and constant dividend. The FX rate is modeled in more detail if it is also assumed that the foreign interest rate is stochastic, which we do in Chapter 5 and 6. The price \(\tilde{U}\) of an option with maturity \(T\), payoff function \(\phi(S_T, V_T, R_d^1)\), and with the initial value of the underlying volatility and domestic short rate equal to \(s\) and \(v\) and \(r_d^1\) respectively equals:

\[\tilde{U}(s, v, r_d^1, t_0) = \mathbb{E}\left[e^{-\int_{t_0}^{T} R_d^1 d\xi} \phi(S_T, V_T, R_d^1)|S_{t_0} = s, V_{t_0} = v, R_d^{1}_{t_0} = r^d\right].\] (4.2)

Because of the stochastic volatility and interest rate components, pricing formulas are three dimensional and an analytic option price is harder to obtain, or not available.

While computing CVA, we assume that the exposure and the counterparty’s default probability are independent. In the case of independence between discount factor, exposure and default probability, we can formulate the discrete expression (1.3) for CVA, in integral form as follows ([47]):

\[\text{CVA}(t_0, T) = (1 - \delta) \int_{t_0}^{T} \text{EPE}^*(t) dPD(t),\] (4.3)

where \(\delta\) is the recovery rate and \(PD(t)\) denotes the default probability of the counterparty at time \(t\).

In practice CVA is hedged and thus practitioners compute the sensitivity of the CVA with respect to its dependencies. We assume that the default probability is independent of exposure, such that the sensitivity with respect to \(\Theta\) (where \(\Theta\)
can be $S_0, V_0$ or $R^d_0$ can be rewritten as:
\[
\frac{\partial \text{CVA}(t_0, T)}{\partial \Theta} = \frac{\partial}{\partial \Theta} \left( (1 - \delta) \int_{t_0}^T \text{EPE}'(t) dPD(t) \right),
\]
\[
= (1 - \delta) \int_{t_0}^T \frac{\partial \text{EPE}'(t)}{\partial \Theta} dPD(t).
\]

(4.4)

Following the same arguments, the second derivative with respect to $\Theta$ can be computed as:
\[
\frac{\partial^2 \text{CVA}(t_0, T)}{\partial \Theta^2} = \frac{\partial}{\partial \Theta} \left( (1 - \delta) \int_{t_0}^T \frac{\partial \text{EPE}'(t)}{\partial \Theta} dPD(t) \right),
\]
\[
= (1 - \delta) \int_{t_0}^T \frac{\partial^2 \text{EPE}'(t)}{\partial \Theta^2} dPD(t).
\]

(4.5)

By computing these sensitivities in this way, we need an efficient computation of the derivatives $\frac{\partial \text{EPE}'(t)}{\partial \Theta}$ and $\frac{\partial^2 \text{EPE}'(t)}{\partial \Theta^2}$ for every $t \in [t_0, T]$.

To conclude, the CVA of a portfolio is determined by all the future Mark-to-Market (MtM) values of all the options in the portfolio ([9]). Further, if we want to compute the sensitivities, we also need the derivative at all future market scenarios. These requirements call for a valuation method that can compute option prices and derivatives for a wide range of market scenarios. In this chapter we will show that the FDMC method can compute these quantities fast and accurately.

### 4.3 Computation of Counterparty Exposure and Sensitivities

#### 4.3.1 The FDMC Method

As presented in the previous chapter, the FDMC method uses the scenario generation of the Monte Carlo method and the pricing approach of the finite difference method. The market states are simulated by the Quadratic Exponential (QE) scheme ([5]). Next, a grid in the $s, v$- and $r^d$-directions is created. This grid is chosen to be sufficiently large to capture all attained values of the scenario generation. On this grid, prices at any simulation date are calculated by the finite difference procedure. The specific state $(S_m, V_m, R^d_m, t)$ is interpolated on the grid to obtain option price $\hat{U}(S_m, V_m, R^d_m, t)$ at each path, for each time point. At every time point the resulting future option values for all paths generate a distribution...
4.3. COMPUTATION OF CCR AND SENSITIVITIES

and from this distribution the exposure profiles can be calculated. The EPE can be obtained by averaging over all the prices at all the time points. The higher (97.5%) and lower (2.5%) PFEs can be computed by taking quantiles.

In the case of a path-dependent barrier option, if the underlying state hits the barrier level $B$, the option is exercised at this path and the exposure for later time points is set to zero. The essential technique of modeling the exposure by the FDMC method can be presented as follows:

- generate scenarios/paths by Monte Carlo simulation;
- calculate option values and for barrier options, check which paths hit the barrier;
- set the exposure at each path equal to the option value if the option is not exercised; otherwise the exposure and all future exposures of this path are set equal to 0;
- compute the empirical distribution of the exposure at each exercise time;
- calculate EPE, PFE$_{2.5\%}$ and PFE$_{97.5\%}$.

One important difference between the FDMC method and other approaches such as the regression based SGBM method presented in [42], or the Monte Carlo COS method ([88]) is that the FDMC method is directly applicable to non-affine models (e.g. the SABR or the Heston Hull - White model with non zero correlation between $S_t$ and $R^d_t$). To compute exposures driven by non-affine models by the COS method or SGBM, an affine approximation is solved (see [54] and [42]). This is not necessary when the FDMC method is used.

4.3.2 The finite-difference method

For a European option with maturity $T$ and payoff function $\phi$ its risk-neutral value $U$ at $t_0 \leq T$ can be expressed using the conditional expectation under the risk-neutral measure $Q$ as follows [12]:

$$
\tilde{U}(S_{t_0}, V_{t_0}, R^d_{t_0}, t_0) = \mathbb{E}\left[e^{-\int_{t_0}^{T} R^d_u du} \phi(S_T) \right],
$$

(4.6)

where $\phi(\cdot)$ is the payoff function of the option. The finite difference procedure computes the price backward in time starting at maturity $t = T$ and continuing to $t = t_0$. Thus, the pricing function $u$ is defined as a function of $\tau = T - t$ such that $U(S_t, V_t, R^d_t, \tau) = \tilde{U}(S_{T-\tau}, V_{T-\tau}, R^d_{T-\tau}, T - \tau)$. The Feynman-Kac theorem links the
expectation (4.6) to the solution of a PDE by no arbitrage arguments, resulting in the following PDE:

$$\frac{\partial U}{\partial \tau} = \mathcal{A}U,$$

(4.7)

where in the case of an FX option driven by the Heston Hull-White dynamics, the spatial differential operator $\mathcal{A}$ is given by

$$\mathcal{A}U = \frac{1}{2} v \frac{\partial^2 U}{\partial s^2} + v \frac{\partial^2 U}{\partial v^2} + \frac{1}{2} \sigma v \frac{\partial^2 U}{\partial s \partial v} + \frac{1}{2} \gamma v \frac{\partial^2 U}{\partial r^2} + \left( r^d - r^f \right) \frac{\partial U}{\partial s} + \left( \kappa \eta - v \right) \frac{\partial U}{\partial v} + \left( \lambda \theta (T - t) - r^d \right) \frac{\partial U}{\partial r}$$

(4.8)

Note that by setting $\gamma$ and $\lambda$ equal to zero we end up with Heston PDE, where $r^d$ is fixed. For a given state $(S_\tau, V_\tau, R^d_\tau)$ at expiry, the payoff is known, and in the finite difference method, this is used as an initial condition.

**Barrier options** For a Down-and-Out barrier call or put option on an underlying $S_\tau$ with strike $K$ and barrier level $B$ the payoff function is equal to:

$$\phi(S_\tau) = \max(p(S_\tau - K), 0) \mathbb{1}_{(S_\tau > B)}$$

with $p = \begin{cases} 1, & \text{for a call,} \\ -1, & \text{for a put.} \end{cases}$

(4.9)

The payoff function for European options can be obtained from this by setting $B = 0$.

**One-touch options** In the case of one-touch options, the holder receives a predetermined payout $H$ whenever the underlying reaches $K$ anytime before maturity $T$. Thus the payoff at expiry is given as:

$$\phi(S_\tau) = H \mathbb{1}_{(S_\tau \geq K)}.$$

(4.10)

**Space discretization**

In the finite difference method this PDE is solved on a finite set of points, by discretizing in $s$, $v$- and $r^d$-direction. The domain to be discretized is chosen as $[0, S_{\max}] \times [0, V_{\max}] \times [-R_{\max}, R_{\max}]$, where $S_{\max}$, $V_{\max}$ and $R_{\max}$ are chosen to be sufficiently large to minimize the effect of the imposed boundary conditions and such that all simulated market scenarios can be interpolated on the grid.
4.3. COMPUTATION OF CCR AND SENSITIVITIES

Let $s_0 < s_1 < \ldots < s_m$, $v_0 < v_1 < \ldots < v_m$, and $r_0 < r_1 < \ldots < r_m$ be the discretization in $s$, $v$- and $r^d$-direction respectively, similar to [55]. In all dimensions the grid is chosen to be non-uniform. The $s$ dimension consists of a predefined interval $[S_{\text{left}}, S_{\text{right}}]$ in which points are uniformly spaced. $S_{\text{left}}$ and $S_{\text{right}}$ are chosen to contain the region of interest i.e., the region around the expected mean of the underlying. Following [55], for options without barriers we choose:

$$[S_{\text{left}}, S_{\text{right}}] = [0.5K, K].$$

Outside $[S_{\text{left}}, S_{\text{right}}]$, the points are distributed with the help of a hyperbolic sine function. In the barrier case, the non-uniform grid is chosen such that the dense region contains more than 95% of the non-exercised paths; generally, choosing

$$[S_{\text{left}}, S_{\text{right}}] = \begin{cases} [0.5K, B], & \text{for a up-and-out call or put,} \\ [B, 1.5K], & \text{for a down-and-out call or put,} \end{cases}$$

is found to be sufficient. For a portfolio of options however, we define $S_{\text{left}}$ and $S_{\text{right}}$ such that all possible strikes and barriers are included. In Figure 4.1, the two different non-uniform grids are shown.

![Figure 4.1: Non uniform grids in s-direction for a Down-and-Out Put (DOP) option (a) (where we choose $S_{\text{left}} = B = 120$ and $S_{\text{right}} = 140$) and for a portfolio of options (b) (with non-uniform points within $S_{\text{left}} = 100$ and $S_{\text{right}} = 150$).](image)

In $v$-direction the grid is chosen similar as in [62]. The grid is dense around $v = 0$. We do this because, for realistic test parameters, the expected mean
of the variance process is close to zero. In addition, because the Heston PDE in the \( v \)-direction is convection-dominated close to zero and the initial condition is non-smooth, numerical stability requires a high density of points in this region ([55]).

Also for the \( r_d \)-direction we follow [55]. The grid is dense around 0 and stretched symmetrically towards the boundaries \(-R_{\text{max}} \) and \( R_{\text{max}} \) by using a sinus hyperbolicus function.

The derivatives are approximated using central, forward and backward three point stencils. All stencils are second-order accurate. For more details we refer to [55].

### Boundary conditions

The options considered in this research are of the following type: European Call and Put options with strike \( K_C \) and \( K_P \) respectively, barrier options with strike \( K_B \) and barrier level \( B \), which can be down-and-out or up-and-out calls or puts and one-touch options with strike \( K_O \) and payout \( H \). The boundary conditions for the \( s \) dimension used in this research are stated in table 4.1. Note that in the case of non barrier options, the \( S_{\text{min}} \) and \( S_{\text{max}} \) converge to 0 or \( \infty \) respectively.

The boundary conditions in the volatility direction are imposed independently of the option type. In [38] it is shown that for a CIR process, like the variance process in the Heston model, the solution of the PDE at \( v = 0 \) satisfies the boundary condition that is obtained by inserting \( v = 0 \) into (4.8); this is also referred to as a degenerated boundary condition:

\[
\frac{\partial U}{\partial \tau} = \kappa \eta \frac{\partial U}{\partial v} + (r_d - r_f)s \frac{\partial U}{\partial s} - r_d U. \tag{4.11}
\]

In traditional literature (see eg. [97]), the maximum variance boundary for call options is imposed as \( U(s, V_{\text{max}}, r_d, \tau) = s \), but experiments show that this introduces a boundary layer. In combination with the PDE becoming convection-dominated
4.3. COMPUTATION OF CCR AND SENSITIVITIES

around $v \approx 0$ this can result in oscillations if no upwinding is applied. To prevent this problem but still use central schemes, the option value at the maximum variance boundary is assumed to satisfy:

$$\frac{\partial^2 U(s, v_{max}, r^d, \tau)}{\partial v^2} = 0.$$  \hspace{1cm} (4.12)

For the interest rate dimension, which is needed in the Heston Hull-White model, the boundary conditions are taken as:

$$\frac{\partial U(s, v, \pm R_{max}^d, \tau)}{\partial r^d} = 0.$$  \hspace{1cm} (4.13)

Using these discretizations, boundary and initial condition, the following initial value problem for stiff Ordinary Differential Equations (ODEs) is derived:

$$\begin{cases} u'(\tau) = A u(\tau) + g(\tau), \\ u(\tau_0) = \phi(s(T)), \end{cases}$$  \hspace{1cm} (4.14)

where $u(\tau)$ denotes the vector of discrete solutions $u_{i,k}(\tau) := U(s_i, v_j, r^d_k, \tau)$ ordered lexicographically, $A$ is the finite difference matrix, $g(\tau)$ is a vector determined by the boundary conditions and $s(T)$ denotes the grid in $s$-direction at maturity.

**Time discretization**

As the Heston Hull-White model is a three dimensional problem in space, the ODEs also have three space dimensions.

In this chapter we apply the Hundsdorfer-Verwer scheme to the two and three dimensional ODEs. For more details we refer to [61] for the derivation of the scheme and to [56] for more details on the three-dimensional ADI scheme in the context of the Heston Hull-White PDE.

4.3.3 Computing CVA and its Sensitivities

To estimate CVA, we need $EPE^*(t)$ at any time $t \in [t_0, T]$ during the life of the derivative. Next to that we need the probability of default at any time. Following [47], we define $q_l = q(t_{l-1}, t_l)$ as the probability that the counterparty will default in the interval $[t_l - dt, t_l]$. Using the so-called hazard rate $\lambda_{max}$ the survival probability $P_{\text{surv}}(t)$ is defined as:

$$P_{\text{surv}}(t) := e^{-\lambda_{max} t}.$$  \hspace{1cm} (4.15)
CHAPTER 4. SENSITIVITIES UNDER HESTON HULL-WHITE

Using this definition we can derive the probability to default in an interval \((t - dt, t)\) conditioned on no prior default as follows:

\[
q(t - dt, t) = P_{\text{surv}}(t) - P_{\text{surv}}(t - dt).
\]  

(4.16)

For any counterparty for which a Credit Default Swap (CDS) is available for protection, this entity can be calculated from the CDS spread. As shown in [102], the annual premium payment \(c\) of a CDS can be calculated as:

\[
c = \left(1 - \delta\right) \sum_{l=1}^{N} \frac{P(t_0, t_l)(q_{l-1} - q_l)}{\sum_{l=1}^{N} P(t_0, t_l)q_l dt + \sum_{l=1}^{N} P(t_0, t_l)(q_{l-1} - q_l) \frac{dt}{T}}.
\]  

(4.17)

where \(dt\) denotes the payment interval. In this research we assume annual premiums of 400 basis points, which correspond to a hazard rate of \(6.6 \cdot 10^{-2}\). Now, in a discrete setting, CVA can be calculated as:

\[
CVA = (1 - \delta) \sum_{l=1}^{N} q(t_{l-1}, t_l) \text{EPE}^* (t_l).
\]  

(4.18)

By using this expression, the first and second derivative of the CVA with respect to \(\Theta_0\) (where \(\Theta_0\) can be \(S_0, V_0\) and \(R_d^0\)) can be derived as follows:

\[
\frac{\partial CVA}{\partial \Theta_0} = (1 - \delta) \sum_{l=1}^{N} q(t_{l-1}, t_l) \frac{\partial \text{EPE}^* (t_l)}{\partial \Theta_0},
\]  

(4.19)

where in the second equality we assume independence between default and \(\Theta_0\) as well as between the recovery rate \(\delta\) and \(\Theta_0\). Note that the assumption of independence between default probability and \(\Theta_0\) can be relaxed by modeling the default probability as a stochastic process which depends on \(\Theta_0\), as is done in [60]. Similar, for the second derivative, we have

\[
\frac{\partial^2 CVA}{\partial \Theta_0^2} = (1 - \delta) \sum_{l=1}^{N} q(t_{l-1}, t_l) \frac{\partial^2 \text{EPE}^* (t_l)}{\partial \Theta_0^2}.
\]  

(4.20)

To compute \(\frac{\partial \text{EPE}^* (t)}{\partial \Theta_0}\) in (4.19), first, the derivative is rewritten as follows:

\[
\frac{\partial \text{EPE}^* (t)}{\partial \Theta_0} = \frac{\partial \text{EPE}^* (t)}{\partial S_t} \frac{\partial S_t}{\partial \Theta_0} + \frac{\partial \text{EPE}^* (t)}{\partial V_t} \frac{\partial V_t}{\partial \Theta_0} + \frac{\partial \text{EPE}^* (t)}{\partial R_d^t} \frac{\partial R_d^t}{\partial \Theta_0}.
\]  

(4.21)

\(^{1}\)In this specific case expression (4.19) will have an extra term, but the sensitivities can still be computed.
At every intermediate time point $t_l$, the finite difference method stores the prices for the entire grid in the vector $u_l' = EPE'(t_l)$. On this grid we can approximate $\frac{\partial EPE'(t_l)}{\partial \Theta_t}$ and $\frac{\partial^2 EPE'(t_l)}{\partial \Theta_t^2}$ by multiplying with the difference matrices $A_{\Theta_t}(t_l)$ and $A_{\Theta_t^2}(t_l)^2$ defined as follows:

$$\frac{\partial EPE'(t_l)}{\partial \Theta_t} \approx A_{\Theta_t}(t_l) u_l = \frac{\partial u(t_l)}{\partial \Theta_t} + O(\Delta \Theta_t^2), \quad (4.22)$$

$$\frac{\partial^2 EPE'(t_l)}{\partial \Theta_t^2} \approx A_{\Theta_t^2}(t_l) u_l = \frac{\partial^2 u(t_l)}{\partial \Theta_t^2} + O(\Delta \Theta_t^2). \quad (4.23)$$

So the partial derivatives of the exposure in equation (4.21) are obtained from the finite difference grid. The partial derivatives of the state variables with respect to the initial conditions are analyzed in the following subsections for all the possible choices of $\Theta$.

**Sensitivity with respect to initial FX rate**

To compute the sensitivity with respect to the initial underlying FX rate ($\Theta_0 = S_0$), we first note that future variance and future short rate are independent of $S_0$, such that $\frac{\partial V_t}{\partial S_0} = \frac{\partial R_d}{\partial S_0} = 0$. However $\frac{\partial S_t}{\partial S_0}$ is clearly non-zero and this can be computed by the pathwise Monte Carlo method. Because $S_0$ follows a Geometric Brownian Motion (GBM) in the Heston Hull-White model, we can assume:

$$S_t = S_0 e^{(R_d t - r_f - \frac{1}{2} V_t^2) t + \sqrt{V_t} \sqrt{t} Z}, \quad (4.24)$$

where $Z$ is a standard normal random variable. Consequently, following ([18]), for the first and second derivative, we have:

$$\frac{\partial S_t}{\partial S_0} = e^{(R_d t - r_f - \frac{1}{2} V_t^2) t + \sqrt{V_t} \sqrt{t} Z} = \frac{S_t}{S_0} \quad (4.25)$$

$$\frac{\partial^2 S_t}{\partial S_0^2} = 0. \quad (4.26)$$

Now, at any time point $t_l$ both partial derivatives from (4.21) can be computed for every path, such that $\frac{\partial EPE'(t_l)}{\partial S_0}$ is obtained by averaging.

---

2Note that for $\Theta_t = \theta^*_t$, $A_{\theta^*_t}(t)$ is a time dependent matrix; as in the case of stochastic interest rate, the drift can be time dependent because of the yield curve (see e.g., the Hull - White model).
CHAPTER 4. SENSITIVITIES UNDER HESTON HULL-WHITE

To compute \( \frac{\partial^2 CVA}{\partial S^2_0} \), we need \( \frac{\partial^2 EPE^*(t)}{\partial S^2_0} \), this yields:

\[
\frac{\partial^2 EPE^*(t)}{\partial S^2_0} = \frac{\partial}{\partial S_0} \left( \frac{\partial EPE^*(t)}{\partial S_t} \frac{\partial S_t}{\partial S_0} \right),
\]

\[
= \left( \frac{\partial^2 EPE^*(t)}{\partial S_t^2} \frac{\partial S_t}{\partial S_0} \right) \frac{\partial S_t}{\partial S_0} + \frac{\partial EPE^*(t)}{\partial S_t} \left( \frac{\partial \partial S_t}{\partial S_0} \right),
\]

\[
= \left( \frac{\partial^2 EPE^*(t)}{\partial S_t^2} \left( \frac{S_t}{S_0} \right)^2 \right). \tag{4.27}
\]

Where the second derivative of EPE with respect to \( S_t \) can be obtained from (4.23), and \( \left( \frac{S_t}{S_0} \right)^2 \) can be obtained from the scenario generation.

Sensitivity with respect to initial variance

In the case of sensitivity with respect to initial variance \( (\Theta_0 = V_0) \), the future short rate is independent of \( V_0 \) such that equation (4.21) can be simplified to:

\[
\frac{\partial EPE^*(t)}{\partial V_0} = \frac{\partial EPE^*}{\partial S_t} \frac{\partial S_t}{\partial V_0} + \frac{\partial EPE^*}{\partial V_t} \frac{\partial V_t}{\partial V_0}. \tag{4.28}
\]

The partial derivatives of the exposures can be extracted from the finite difference grid by (4.22) such that only \( \frac{\partial S_t}{\partial V_0} \) and \( \frac{\partial V_t}{\partial V_0} \) are unknown. In the Heston model, the variance is modeled by a square root process, which has its difficulties. In this model, the variance process can and will reach zero, such that a straightforward derivative of (4.24) with respect to \( V_0 \) is not defined for every path at every time. This also holds for \( \frac{\partial V_t}{\partial V_0} \), because when the discretized SDE of the variance process is differentiated with respect to \( V_0 \), the square root will appear in the denominator, which makes the derivative intractable. Furthermore, in [25] it is noted that the sensitivities of the variance process with respect to initial inputs can grow very quickly and potentially blow up. We therefore approximate these partial derivatives by a so called local bump and revalue approach as follows:

\[
\frac{\partial S_t}{\partial V_0} \approx \frac{\tilde{S}_m - S_m}{\epsilon_v}, \tag{4.29}
\]

\[
\frac{\partial V_t}{\partial V_0} \approx \frac{\tilde{V}_m - V_m}{\epsilon_v}, \tag{4.30}
\]

where \( \tilde{S} \) and \( \tilde{V} \) are modeled with \( V_0 + \epsilon_v \) as initial variance. Note that when an analytic expression for these partial derivatives is available the method will gain in efficiency.
4.3. COMPUTATION OF CCR AND SENSITIVITIES

Sensitivity with respect to initial domestic interest rate

To measure the sensitivity with respect to the initial domestic short rate $R^d_0$, (4.21) is simplified to:

$$\frac{\partial \text{EPE}^*(t)}{\partial R^d_0} = \frac{\partial \text{EPE}^*}{\partial S_t} \frac{\partial S_t}{\partial R^d_0} + \frac{\partial \text{EPE}^*}{\partial R^d_t} \frac{\partial R^d_t}{\partial R^d_0}$$

where $\frac{\partial V}{\partial R^d_0}$ is zero because the future variance is independent of the initial short rate. Similar as in the previous cases, the partial derivatives of EPE with respect to $S_t$ and $R^d_t$ can be derived from the finite difference grid by equation (4.22).

The partial derivatives of the state variables with respect to $R^d_0$ are computed along the path in the Monte Carlo simulation. For any discrete time point $t_l$ ($0 < l < N$), we can create a recursive formula for $\frac{\partial S_{t_l}}{\partial R^d_0}$:

$$\frac{\partial S_{t_l}}{\partial R^d_0} = \frac{\partial S_{t_l}}{\partial S_{t_{l-1}}} \frac{\partial S_{t_{l-1}}}{\partial R^d_0} + \frac{\partial S_{t_l}}{\partial R^d_{t_{l-1}}} \frac{\partial R^d_{t_{l-1}}}{\partial R^d_0}$$

where (when $S_t$ is driven by a geometric Brownian motion) we have:

$$\frac{\partial S_{t_l}}{\partial S_{t_{l-1}}} = S_{t_{l-1}}$$
$$\frac{\partial S_{t_l}}{\partial R^d_{t_{l-1}}} = S_{t_l} \Delta t$$

where $\Delta t = t_l - t_{l-1}$ is the uniform time increment in one time step. The interest rate is modeled by the Hull - White model ([59]) such that an Euler scheme as a discretization yields:

$$R^d_{t_{l+1}} = R^d_{t_l} + \lambda \left[ \theta(t) - R^d_{t_l} \right] \Delta t + \gamma \sqrt{\Delta t} Z_t,$$

where $Z \sim N(0,1)$. From this we can recursively derive:

$$\frac{\partial R^d_{t_{l+1}}}{\partial R^d_0} = (1 - \lambda \Delta t)^{(l-1)}.$$

The first time step gives us the initial condition:

$$\frac{\partial S_{t_1}}{\partial R^d_0} = S_0 \Delta t.$$

Using this recursive formula together with the finite difference approximations, we can estimate the sensitivity with respect to $R^d_0$ at any time point without the need of an extra Monte Carlo simulation.

Similar to the case of EPE, the computation of the first and second derivatives with respect to initial underlying states can be summarized as follows:
• generating scenarios/paths by Monte Carlo simulation;
• at each time point \( t_l \), for the entire grid, calculate option sensitivities \( \frac{\partial \text{PE}^*(t_l)}{\partial \theta} \) and \( \frac{\partial^2 \text{PE}^*(t_l)}{\partial \theta^2} \) and for barrier options, check if option is not exercised \((S_{t^*} < B \text{ for all } t^* \leq t_l)\);
• set the first and second derivatives at each path as the calculated sensitivities if the option is not exercised; otherwise set them equal to 0;
• compute the empirical distribution of the sensitivities at each exercise time;
• calculate \( \frac{\partial \text{PE}^*(t_l)}{\partial \theta} \) and \( \frac{\partial^2 \text{PE}^*(t_l)}{\partial \theta^2} \) by averaging.

4.3.4 Pricing a Portfolio

In this research, the finite difference grid is used to price multiple options with different strikes and maturities in one sweep on one grid. The portfolios considered here are constructed of European options and a first order exotic barrier option. The value \( \Pi \) of a portfolio of \( N \) options can be seen as the sum of the option prices:

\[
\Pi(t) = \sum_{i=1}^{N} U_i(S_t, K_i, T_i),
\]  

(4.38)

where \( K_i \) is the strike, \( T_i \) the maturity and \( U_i \) the price of the \( i \)-th option. In this Chapter option \( i \) can be a European call or put option, or a barrier option. We assume that all options in the portfolio can be netted.

Together with the Monte Carlo scenario generation, this gives us the exposure profile of the sum of the option values at any future time point. The duration of the portfolio is equal to the longest maturity in the portfolio:

\[
\tilde{T} = \max_{i \in [1,N]} T_i.
\]  

(4.39)

Again, at this maximum maturity, all the option prices on the grid are known, therefore the time is reversed such that the payoff formula (4.40) can be used as an initial condition which equals the sum of all individual payoff functions belonging to options with maturity equal to the maximum maturity \( \tilde{T} \):

\[
\phi_p(S_t) = \sum_{i=1}^{N} \phi_i(S_t, K_i, T_i) \mathbb{I}_{(T_i = \tilde{T})},
\]  

(4.40)

\(^3\)Here we slightly abuse notation, whereas \( U \) is now a function of \( S, K \) and \( T \) instead of a function of the risk factors earlier.
4.3. COMPUTATION OF CCR AND SENSITIVITIES

Important in the context of this research is that the option specific characteristics are only introduced by the initial condition and the boundary conditions. Because the portfolio consists of a sum of options, the boundary conditions for the portfolio will also be just a sum of these limiting conditions, such that our time stepping routine in the finite difference procedure can be updated as follows:

- From $u^l$, calculate $\tilde{u}^{l-1}$ by the ADI splitting scheme;
- update the portfolio value with possible other option values:

$$u^{l-1} = \tilde{u}^{l-1} + \sum_{i=1}^{N} \phi_i(s, K_i, T_i) \mathbb{I}_{(T_i = t_{l-1})},$$

where $s$ is a vector of the same size as $u$ consisting of all $s$ grid points.
- update all boundary conditions;
- if $l > 0$, repeat procedure.

By applying this time stepping procedure, $u^0$ will be the value of the portfolio at time $t = t_0$. For the computation of exposure of a portfolio over time, options that are not path-dependent can be included on one grid. By using only one grid, there will be no extra computational time for these extra options.

In the case of a portfolio of a call, put and a barrier considered in this research, the EPEs of the call and put option are computed using only one grid, whereas the EPE of the barrier option is computed using the algorithm from Section 4.3.1. In general, for the computation of path-dependent options, first a separate finite difference procedure needs to be done. Next, while computing this individual options exposure, for any scenario it needs to be checked if the simulated scenario should be exercised or not. At every time step, the portfolio exposure is then computed as the sum of the individual barrier exposure and the call and put option exposure. To summarize, the computational time of computing exposure for a portfolio, is determined by:

- one MC simulation to compute all the scenarios;
- one FD procedure to compute the price grid for all non path-dependent options;
- a separate FD procedure per path-dependent option.

Note that the same holds for path-dependent American options.
CHAPTER 4. SENSITIVITIES UNDER HESTON HULL-WHITE

Table 4.2: Model parameters for various test cases.

<table>
<thead>
<tr>
<th></th>
<th>Case A</th>
<th>Case B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spot ($S_0$)</td>
<td>1.364</td>
<td>138.1</td>
</tr>
<tr>
<td>Foreign short rate ($r^2$)</td>
<td>0.01</td>
<td>0.10</td>
</tr>
<tr>
<td>Initial variance ($V_0$)</td>
<td>0.029</td>
<td>0.029</td>
</tr>
<tr>
<td>Mean reversion speed of variance ($\kappa$)</td>
<td>4.42</td>
<td>1.50</td>
</tr>
<tr>
<td>Mean reversion variance level ($\eta$)</td>
<td>0.0240</td>
<td>0.0707</td>
</tr>
<tr>
<td>Vol of vol ($\sigma$)</td>
<td>0.46</td>
<td>0.63</td>
</tr>
<tr>
<td>Initial domestic short rate ($R_0^d$)</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>Mean reversion speed of interest rate ($\lambda$)</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Mean reversion interest rate ($\theta$)</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>Vol of short rate ($\gamma$)</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>$s,v$-correlation ($\rho_{1,2}$)</td>
<td>-0.45</td>
<td>-0.76</td>
</tr>
<tr>
<td>$s,r^d$-correlation ($\rho_{1,3}$)</td>
<td>0.501</td>
<td>-0.011</td>
</tr>
<tr>
<td>$v,r^d$-correlation ($\rho_{2,3}$)</td>
<td>-0.96</td>
<td>-0.96</td>
</tr>
<tr>
<td>Maturity ($T$)</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>Strike ($K$)</td>
<td>1.360</td>
<td>138.1</td>
</tr>
<tr>
<td>Barrier ($B$)</td>
<td>1.20</td>
<td>120</td>
</tr>
</tbody>
</table>

4.4 Numerical Results

The numerical results are divided into two parts. First we present our numerical study on the accuracy and convergence under the two-dimensional Heston model. In this case, all results are validated by the semi-analytical COS method.

Second, we further assess the accuracy of the interpolation schemes by looking into the numerically challenging Heston Hull-White model, where the number of grid points per dimension is smaller which can cause a larger interpolation error. We do this in combination with the EPE computation of OT options which are discontinuous during the entire lifetime of the option. Next to that we show the sensitivities with respect to $V_0$ and $R_0^d$.

The parameters are chosen according to Table 4.2. In test A the foreign interest rate is equal to the initial domestic rate, next to that the option is Out-of-The-Money (OTM) at inception. The level of the initial FX rate is set at 1.3639, which is a real market quoted EUR/USD FX rate from June 2014, whereas the other parameters satisfy characteristics observed in literature, such as negative correlation, low volatility and interest rates and small maturities (see eg., [87] and [2]). In this test, the well known Feller condition is satisfied. In test B the initial FX rate is set to 138.1, a real EUR/JPY FX rate from June 2014, the option is ATM at inception while the initial domestic interest rate is higher than the foreign interest rate. In this case the other model parameters are chosen...
such that the Feller condition is violated.\textsuperscript{4}

4.4.1 Heston Model

In the case of the Heston model, the domestic interest rate $R^d_t$ is assumed to be constant over time, such that

$$D(0,t) = e^{-\int_0^t R^d_s ds} = e^{-r^d t}.$$  

Furthermore, in the pricing PDE (4.8), $\lambda$ and $\gamma$ are assumed to be zero.

Single barrier options: Numerical Setup

Computing the exposure of barrier options is more challenging than computing the exposure of European options. Barrier options are path-dependent and have a discontinuous initial condition. It is this discontinuous nature of the payoff function in particular that may complicate accurate estimation of sensitivities, especially for higher-order ones. Because we have a benchmark solution for Down-and-Out Put (DOP) options, we do an extensive error analysis for this option type, but the method can also be applied to Down-and-Out Call (DOC), Up-and-Out Put (UOP) and Up and Out Call (UOC) options and all the other “In” (instead of “Out”) variants.

The computed EPE, PFE$_{2.5\%}$ and PFE$_{97.5\%}$ are shown in figures 4.2(a) and 4.2(b). The starting level of the EPE, equals the option price at $t = t_0$ and shows a small increase towards maturity. The PFE however, shows a more interesting behavior. Starting at the option price, the PFE is increasing over time and shows a steep growth close to maturity. Intuitively, the increase of the higher quantile makes sense, when moving $t^* \in [t_0,T]$ closer to maturity, the hitting probability conditioned on no prior barrier hit, will become smaller, such that for in-the-money paths, the price will resemble a straightforward European option value more and more. The mean (EPE) is not heavily affected because also the probability of the barrier being hit up to time $t^*$ is increased which will lower the option value.

Accuracy and Convergence

As a benchmark, the COS method can be applied to evaluate barrier options accurately and efficiently. For details on the pricing procedure using this Fourier cosine method we refer to \cite{41}. Here we use this efficient pricing technique by computing prices for an entire grid of possible market scenarios. Similar as the

---

\textsuperscript{4}When the Feller condition is not satisfied, the variance process can become zero and numerical methods can become unstable.
CHAPTER 4. SENSITIVITIES UNDER HESTON HULL-WHITE

Figure 4.2: Exposures ($EPE_{2.5\%}$ and $PFE_{97.5\%}$) and the first and second derivative profiles over time under the Heston dynamics for tests A and B. The dashed black line is computed using the FDMC method, whereas the dashed red line is computed using the COS method. In the case of the sensitivities, the results corresponding to the COS method are obtained using a Bump-and-Revalue (B&R) procedure whereas for the FDMC method the derivative is splitted as explained in Section 4.3.3.
4.4. NUMERICAL RESULTS

grid used for the finite difference procedure, this grid is chosen such that it is large enough to contain all future market scenarios generated by the Monte Carlo scenario generation. We choose 500 points in $x$- and 300 points in $u$-direction, both densely distributed around the expected means. Prices for all the scenarios are obtained by a spline interpolation on the COS grid.

For the sensitivities with respect to $S_0$, we run this procedure two (in the case of delta) or three (in the case of gamma) times with an initial condition bumped by $\epsilon$. From the resulting EPEs, the sensitivities are computed using the following finite difference formulas:

$$\frac{\partial \text{EPE}(t)}{\partial S_0} \approx \frac{\text{EPE}_{S_0+\epsilon}(t) - \text{EPE}_{S_0-\epsilon}(t)}{2\epsilon}, \quad (4.41)$$

$$\frac{\partial^2 \text{EPE}(t)}{\partial S_0^2} \approx \frac{\text{EPE}_{S_0+\epsilon}(t) - 2\text{EPE}_{S_0}(t) + \text{EPE}_{S_0-\epsilon}(t)}{\epsilon^2}. \quad (4.42)$$

Figures 4.2(a) to 4.2(f) show that the exposure profiles and sensitivities over time computed with the FDMC method resemble the results computed by the Monte Carlo COS method.

For an EPE computed over $N_T$ evaluation dates, the relative $L_2$ and $L_\infty$ errors are computed as:

$$\| \cdot \|_\infty := \max_{i=1,\ldots,N_T} \left| \text{EPE}^\text{COS}_i - \text{EPE}^\text{FDMC}_i \right| / \max_{i=1,\ldots,N_T} \left| \text{EPE}^\text{COS}_i \right|,$$  

$$\| \cdot \|_2 := \left( \sum_{i=1}^{N_T} \left( \text{EPE}^\text{COS}_i - \text{EPE}^\text{FDMC}_i \right)^2 \right)^{\frac{1}{2}} / \left( \sum_{i=1}^{N_T} \left( \text{EPE}^\text{COS}_i \right)^2 \right)^{\frac{1}{2}}. \quad (4.43)$$

Table 4.3: Relative $L_2$ and $L_\infty$ errors compared to the COS method using the linear or spline interpolation.

<table>
<thead>
<tr>
<th>Error</th>
<th>Quantity</th>
<th>Test A</th>
<th>Test B</th>
<th>Test A</th>
<th>Test B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Linear interpolation</td>
<td>Spline interpolation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$| \cdot |_\infty$</td>
<td>EPE</td>
<td>2.17 10^{-3}</td>
<td>3.71 10^{-3}</td>
<td>2.17 10^{-3}</td>
<td>3.70 10^{-3}</td>
</tr>
<tr>
<td></td>
<td>PFE_{0.75%}</td>
<td>4.91 10^{-3}</td>
<td>5.33 10^{-3}</td>
<td>4.92 10^{-3}</td>
<td>5.23 10^{-3}</td>
</tr>
<tr>
<td></td>
<td>$\partial p_{\text{EPE}}$</td>
<td>4.79 10^{-3}</td>
<td>6.62 10^{-3}</td>
<td>4.81 10^{-3}</td>
<td>6.62 10^{-3}</td>
</tr>
<tr>
<td></td>
<td>$\partial^2 p_{\text{EPE}}$</td>
<td>3.60 10^{-2}</td>
<td>3.71 10^{-2}</td>
<td>3.60 10^{-2}</td>
<td>3.71 10^{-2}</td>
</tr>
<tr>
<td>$| \cdot |_2$</td>
<td>EPE</td>
<td>1.82 10^{-4}</td>
<td>3.13 10^{-4}</td>
<td>1.82 10^{-4}</td>
<td>3.12 10^{-4}</td>
</tr>
<tr>
<td></td>
<td>PFE_{0.75%}</td>
<td>3.08 10^{-3}</td>
<td>5.22 10^{-3}</td>
<td>3.08 10^{-3}</td>
<td>5.23 10^{-3}</td>
</tr>
<tr>
<td></td>
<td>$\partial p_{\text{EPE}}$</td>
<td>3.85 10^{-3}</td>
<td>5.33 10^{-3}</td>
<td>3.85 10^{-3}</td>
<td>5.23 10^{-3}</td>
</tr>
<tr>
<td></td>
<td>$\partial^2 p_{\text{EPE}}$</td>
<td>2.60 10^{-2}</td>
<td>2.28 10^{-2}</td>
<td>2.59 10^{-2}</td>
<td>2.29 10^{-2}</td>
</tr>
</tbody>
</table>
In table 4.3 the errors between the FDMC method with 700 grid points in \( s \)- and 350 in \( v \)-direction are compared to the COS method. We can see that the relative error is below 1% in both EPE and the first derivative. The second derivative however, is accurate up to 5% in both \( L_\infty \) and \( L_2 \) norms. This is due to the fact that this absolute value of gamma is already in the range of \( 10^{-4} \) such that the errors from the finite difference discretization have a larger impact. Furthermore, we can see that the difference between a spline and linear interpolation is negligible.

![Figure 4.3: Error convergence of EPE and first- and second-order sensitivities for tests A and B by increasing the number of grid points (\( 2^m \times m \)) used in the finite difference computation. We use \( m \) in \( v \)- and \( 2m \) in \( s \)-direction. For every exposure computation, \( 10^5 \) paths are used simulated with a fixed seed to avoid noise. Here, a spline interpolation is used, but similar analysis performed using a linear interpolation is shown in table 4.3.](image)

The convergence with respect to the number of finite difference grid points is shown in figures 4.3(a) and 4.3(b), for the EPE and the first and second derivative with respect to \( S_0 \). In this case, the benchmark is the converged finite difference solution obtained with 700 and 350 points in \( s \)- and \( v \)-direction respectively. The convergence is shown to be first-order in the number of grid points.

In figures 4.4(a) and 4.4(b), we show the decline of the relative Standard Error (SE) in percentage of the mean, by increasing the number of paths for tests A and B. Here we computed this standard error using 10 Monte Carlo simulations with different seeds. Typically, the Monte Carlo convergence is expected to be \( 1/\sqrt{N_p} \) where \( N_p \) is the number of Monte Carlo paths. We see that for both tests, all quantities converge as expected.

Note that the computation time of the FDMC method heavily depends on the number of grid points. When we compare the computation time of the price grids, computed by COS or FDMC, we see that the FDMC method is significantly
4.4. NUMERICAL RESULTS

![Figure 4.4](image)

(a) Test A

(b) Test B

Figure 4.4: Convergence of the relative Standard Error (SE) of CVA, delta of CVA and gamma of CVA for tests A and B for increasing number of paths. Here, the number of finite difference grid points is set equal to 350 in \( v \)- and 700 in \( s \)-direction and the standard error is computed relative to the mean.

Portfolio of options

For the evaluation of CVA, we assume a recovery rate of 40%. The hazard rate is computed by assuming a 5 year CDS with a spread of 400 basis points paid quarterly. The Euro discount factors are taken from April 2014, the resulting survival probabilities up to one year are obtained as explained in Section 4.3.3. Because we assume absence of wrong- and right-way risk, we can compute the CVA for any CDS spread. In this case, CVA is a linear function of the CDS spread.

Different options in one portfolio can have different strikes and maturities. Due to these different maturities, the finite difference procedure is faced with a discontinuity in time. To assess the possible effect on the accuracy, we consider a portfolio of two European options with different strike and maturity. We again compare the resulting exposure profiles and CVA values with the Monte Carlo COS method. For the benchmark, we compute separate exposure profiles for every option with the Monte Carlo COS method and compute the EPE of the portfolios as the sum. This is similar for the sensitivities that are obtained by a bump-and-revalue procedure per option. In the FDMC method, the Call

---

5In our implementation the COS grid was obtained by subsequently pricing every grid point, which can probably be improved.
and Put options are computed simultaneously on one grid. The barrier option is computed on a separate grid because for every path termination needs to be checked. The resulting option prices per path are added to the portfolio values and from this, the mean and quantiles can be calculated.

Again, we assume the Heston dynamics to drive the underlying risk factors. The Heston parameters that drive the underlying are chosen as in case B of the previous subsection. All the options in the portfolio are written on this single FX rate. We consider two portfolios. Portfolio I consists of a Call, Put and a barrier option, while portfolio II consists only of the call and a put. Table 4.4 shows the option parameters for the two portfolios.

The results presented here also hold for portfolios consisting of an arbitrary larger number of options, but for illustrative reasons we present results for only three options.

In table 4.5 we show the CVA values. Here we computed the CVA as a percentage of the portfolio value. The sensitivities are quoted relative to the sensitivities of the initial portfolio. This way we can quantify the change between the CVA adjusted and the non CVA adjusted portfolio:

\[
\text{CVA}_\% := 100 \frac{\text{CVA}_\text{Π}}{\text{Π}},
\]

\[
\Delta_{S_0} := 100 \frac{\partial \text{CVA}}{\partial S_0} \bigg|_{\partial \Pi} 
\]

\[
\Gamma_{S_0} := 100 \frac{\partial^2 \text{CVA}}{\partial S_0^2} \bigg|_{\partial \Pi} 
\]

By looking at figures 4.5(a) and 4.5(b), we can see that the EPE drops at \( t = 0.4 \) when the put option expires. This discontinuity is captured nicely by the FDMC method, where the Put and Call option are computed on one finite difference grid. By looking at table 4.5 we can conclude that the resulting value adjustments are accurate compared to the Monte Carlo COS method. Next to that the difference between spline and linear interpolation is small.

In portfolio I, the call and put options have a bigger effect on the EPE than the barrier option. Also the higher PFE is heavily affected by the expiry of the
4.4. NUMERICAL RESULTS

Figure 4.5: Exposure, Delta and Gamma profiles for portfolios I and II over time for Case B computed with the FDMC method or with the COS method. Again the COS sensitivities are computed running 2 or three simulations from a bumped initial value. The results compared to the COS method are accurate up to an order of $10^{-3}$. 
CHAPTER 4. SENSITIVITIES UNDER HESTON HULL-WHITE

Table 4.5: CVA, delta and gamma for the portfolios I and II as a percentage of non-adjusted values. The percentages are computed by spline and linear interpolation both for the FDMC method as for the benchmark COS method. The sensitivities in the COS method are obtained by a bump and revalue technique.

<table>
<thead>
<tr>
<th></th>
<th>Linear interpolation</th>
<th>Spline interpolation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Portfolio I</td>
<td>Portfolio II</td>
</tr>
<tr>
<td></td>
<td>Portfolio I</td>
<td>Portfolio II</td>
</tr>
<tr>
<td>CVA FDMC</td>
<td>2.79%</td>
<td>2.77%</td>
</tr>
<tr>
<td>CVA COS</td>
<td>2.79%</td>
<td>2.77%</td>
</tr>
<tr>
<td>$\Delta S_0$ Splitting FDMC</td>
<td>23.49%</td>
<td>18.78%</td>
</tr>
<tr>
<td>$\Delta S_0$ B&amp;R COS</td>
<td>23.52%</td>
<td>18.85%</td>
</tr>
<tr>
<td>$\Gamma S_0$ Splitting FDMC</td>
<td>2.58%</td>
<td>2.42%</td>
</tr>
<tr>
<td>$\Gamma S_0$ B&amp;R COS</td>
<td>2.52%</td>
<td>2.38%</td>
</tr>
</tbody>
</table>

Next, if we look at the delta profiles in figures 4.5(c) and 4.5(d), we see that for portfolio I, the impact of the barrier option is reflected by a steep decrease at the expiry of the barrier option $T = 0.8$. As this barrier option is absent in portfolio II, this decrease is absent in the delta profile of portfolio II. This impact is also confirmed by looking at the sensitivity of CVA with respect to $S_0$ in table 4.5 where the difference between portfolio I and II is in the region of 25% for $\Delta S_0$.

In the case of gamma, shown in figures 4.5(e) and 4.5(f), the barrier option in portfolio I shows a steep increase at the expiry of the barrier option. The relative impact for gamma however is smaller than for delta. In table 4.5, we see that the difference in gamma between portfolio I and II is in the range of 3%.

We can furthermore see that the spline interpolation yield similar results as the linear interpolation. These results indicate that the effect of a barrier options in a portfolio can be more severe in the sense of sensitivities than in CVA itself. Clearly, a small change in the EPE profile can have a bigger impact on the first - and second - order sensitivity. Further, in Appendix A, we show that for barrier options the sensitivities are also more sensitive to changes in moneyness levels.
4.4. NUMERICAL RESULTS

4.4.2 Heston Hull-White model

As in this chapter the parameters are not calibrated, the mean reverting level in the short rate process $\theta(t)$ is assumed to be constant over time$^6$.

One-touch options

The one-touch option only delivers a fixed payoff at maturity and therefore has a discontinuous profile during the entire lifetime of the option. Next to that, the three dimensional dynamics of the Heston Hull-White model implies a coarser grid in every dimension. Because of this discontinuity and the coarser grid, the interpolation scheme is more important. In figure 4.6 we show the exposures for tests A and B. We can see that the 97.5% PFE in test B reaches the maximum payout level earlier than in test A. The lower mean reversion speed in combination with the higher vol of vol and long term mean of the variance process in test B cause fatter tails which imply a higher hitting probability.

The spline and linear exposures are very close. This is further shown in table 4.6, where we see that the $L_\infty$ and $L_2$ differences are in the range of 1%.

![Exposures and PFE for a one year one-touch option with barrier level 1.2S_0 and payout 100. Here we use 100 time points and the number of finite difference grid points is set equal to 100 in s- and 50 in v- and r^d direction.](image)

Figure 4.6: Exposures and PFE for a one year one-touch option with barrier level 1.2S_0 and payout 100. Here we use 100 time points and the number of finite difference grid points is set equal to 100 in s- and 50 in v- and r^d direction.

$^6$A study on real model impact on exposure in which parameters are calibrated to real market data is subject of the next chapter.
### Table 4.6: Relative $L_2$ and $L_\infty$ differences between exposure metrics computed using linear and spline interpolation for one year one-touch option with barrier level $1.25S_0$ and payout 100.

<table>
<thead>
<tr>
<th>Differences</th>
<th>Quantity</th>
<th>Test A</th>
<th>Test B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td></td>
<td>\cdot</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PFE_{97.5%}</td>
<td>$1.123 \times 10^{-2}$</td>
<td>$6.310 \times 10^{-3}$</td>
</tr>
<tr>
<td></td>
<td>PFE_{4.5%}</td>
<td>$4.327 \times 10^{-4}$</td>
<td>$1.066 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

| $|| \cdot ||_2$     | EPE         | $2.411 \times 10^{-4}$ | $3.002 \times 10^{-4}$ |
|                      | PFE_{97.5\%} | $6.454 \times 10^{-3}$ | $1.410 \times 10^{-3}$ |
|                      | PFE_{4.5\%}  | $7.700 \times 10^{-4}$ | $1.504 \times 10^{-3}$ |

**Other sensitivities**

Here we look at vega and rho for one-touch options over time. Note that, in this case, the bump and revalue method uses two Monte Carlo simulations and estimates the derivative by:

$$\frac{\partial \text{EPE}(t)}{\partial \Theta} \approx \frac{\text{EPE}_{t+\epsilon\Theta}(t) - \text{EPE}_t(t)}{\epsilon \Theta},$$ (4.48)

which is only first order accurate in $\epsilon \Theta$ which we choose as $0.01 \cdot V_0$ and $0.01 \cdot R_0^4$ respectively. Figures 4.7 show that the sensitivities computed by the splitting scheme and by the bump and revalue method agree over time. This is further confirmed by looking at the relative differences that are again in the range of 1% as presented in Table 4.7.

**Table 4.7:** Relative $L_2$ and $L_\infty$ differences between future exposure sensitivities computed using a splitting scheme and a bump and revalue technique for one year one-touch option with barrier level $1.25S_0$ and payout 100.

<table>
<thead>
<tr>
<th>Difference</th>
<th>Quantity</th>
<th>Test A</th>
<th>Test B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td></td>
<td>\cdot</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\frac{\partial \text{EPE}}{\partial \text{K}_0^4}$</td>
<td>$6.872 \times 10^{-3}$</td>
<td>$3.922 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

| $|| \cdot ||_2$     | $\frac{\partial \text{EPE}}{\partial \text{V}_0}$ | $6.582 \times 10^{-3}$ | $2.478 \times 10^{-3}$ |
|                      | $\frac{\partial \text{EPE}}{\partial \text{K}_0^4}$ | $1.946 \times 10^{-3}$ | $1.813 \times 10^{-3}$ |

### 4.5 Conclusion

In this Chapter we extend the FDMC method which is based on combining the Monte Carlo scenario generation with option valuation by solving a PDE on a
4.5. CONCLUSION

Figure 4.7: Derivatives of EPE over time for one-touch options. The derivatives are with respect to $S_0$, $V_0$, and $r_d^0$. The sensitivities are computed by the splitting method and the bump and revalue method. We use 100 time points and the number of finite difference grid points is set equal to 100 in $s$ and 50 in $v$ and $r^d$ direction.

grid. For the Heston model, we have shown that the FDMC is a computationally efficiently and accurately method compared to a benchmark Monte Carlo COS method and can therefore serve as an alternative to the widely used American Style Monte Carlo approach, which in application to exotic options can suffer from regression bias.

The sensitivities with respect to $S_0$ and $R_d^0$ are obtained efficiently by leveraging from the finite difference grid. Compared to a “brute force” bump - and - revalue technique the sensitivity results are accurate and no extra Monte Carlo simulations are needed which is a computational advantage. In the case of sensi-
tivity with respect to $V_0$, we show that when an analytic expression of the future variance with respect to initial variance is available, the same technique can be applied.

Under the Heston dynamics, we analyze the accuracy of the method by comparing with a benchmark solution and we assess the convergence of the solutions by first increasing the number of paths in the Monte Carlo simulation and secondly, increasing the number of grid points used in the finite difference procedure. As expected, the standard error converges by $1/\sqrt{N}$, where $N$ is the number of paths. By increasing the number of grid points, the relative error converges in first order.

Next we show that we can use the method to compute exposure profiles for a portfolio of options with different maturities. In this portfolio, the EPEs of all options that are not path-dependent (European options) can be efficiently computed on a single grid. The resulting discontinuity in time is captured and no significant error propagation is observed. The EPEs for path-dependent options have to be computed individually and are added to the portfolio before computing the means. The sensitivities can again be computed with small extra computational time. Results compared to the Monte Carlo COS method are accurate, as well as for linear as for spline interpolation.

To further assess the impact of the interpolation, one-touch options are considered which have a discontinuity over time. In combination with the computational challenging Heston Hull-White model, where less grid points can be used, the interpolation is essential. We found that even in this case, a linear interpolation is sufficient as differences are smaller than 1.2% both for exposures as for sensitivities.

We have shown that the FDMC method can be used to compute accurate exposure sensitivities and that it can be applied to three dimensional models. However, often portfolios are driven by more than four risk factors; therefore, it is interesting to investigate even higher dimensional problems. In order to overcome the curse of dimensionality, in the next chapter, we will decompose a high dimensional problem into multiple low-dimensional problems such that efficient forward Kolmogorov PDEs for the transition densities can be solved. This way we will have a full PDE framework for the computation of exposure.
Efficient exposure computation by risk factor decomposition

In derivative portfolios, all risk factors are correlated such that they cannot be modeled independently. Therefore efficient numerical techniques are needed to compute high dimensional exposure distributions in the future. In this setting, although its convergence is slow, the Monte Carlo method is considered industry standard because it scales linearly in the number of dimensions. In this chapter we take a different approach and use the finite difference method to solve the Kolmogorov backward and forward PDEs. We tackle the curse of dimensionality by decomposing the problem into solutions of low-dimensional PDEs and recombining these subsequently. We investigate our method by comparing exposure profiles for a realistic portfolio of Interest Rate, Cross Currency Swaps (CCYS) and FX call options. As a benchmark, we use the Monte Carlo method and show that our method is accurate. Next to that, our method can be used to obtain a highly effective control variate for variance reduction.

CHAPTER 5. CVA BY RISK FACTOR DECOMPOSITION

5.1 Introduction

Many traded portfolios consist of multiple underlying assets and derivatives on these assets; these are valued by models with multiple risk drivers such as stochastic interest rates, FX rates and stochastic volatility. Computing the future value distributions, for which closed-form solutions are typically not available, is a challenging high-dimensional computational problem. A typical benchmark for methods geared towards high-dimensional problems are basket options, where the payoff depends on a portfolio of assets, such as stocks, stock indices or currencies. These options serve the purpose of diversification and are therefore popular among investors [14]. For the valuation of these options, contributions have been made by, e.g., Jain and Oosterlee [66] for Monte Carlo estimation under early exercise and Reisinger and Wissmann [83] for PDE approximations.

In the context of CVA, netting needs to be taken into account and possible negative or positive parts of the exposure distribution are to be considered. This makes CVA estimation for simple plain vanilla instruments similar to the valuation of a series of basket options with different maturity, but is more complex if the instruments themselves are more complex, i.e., no closed-form or semi-closed-form pricing formulae are available. The computational problem in that case can be formulated as the estimation of nested conditional expectations, where the inner expectations are future derivative prices conditional on the underlying risk factors, and the outer expectations, taken over the risk factors, are the expected exposures for the derivative portfolio.

For computing these risk measures, typically, Monte Carlo methods are used to sample the future states of the underlying risk factors by discretization and simulation. Portfolio values for all these states can then be computed in multiple ways, including: fully nested Monte Carlo simulation [45, 19], where derivative values along those paths are estimated by an ‘inner’ Monte Carlo simulation; regression based approaches, introduced for American style options in Longstaff and Schwartz [73] and applied in the context of exposure in Karlsson et al. [69], Joshi and Kang [68]; or a grid or Fourier based method for the inner expectations and Monte Carlo sampling of the outer expectation [31].

All these different methods have their particular strengths and weaknesses. The naïve nested Monte Carlo method is generically applicable but computationally extremely expensive with a complexity of $O(\epsilon^{-4})$ for root-mean-square error (RMSE) $\epsilon$. This can be improved when the computational budget between inner and outer simulations is optimally balanced, e.g., $O(\epsilon^{-3})$ when trading off ‘inner’ bias and ‘outer’ variance [45], and $O(\epsilon^{-5/2-\delta})$ for arbitrary $\delta > 0$ for certain non-uniform estimators [19]. Regression based approaches have in a sense optimal complexity $O(\epsilon^{-2})$ as no extra ‘inner’ paths are required [20], but an additional bias creeps in when leaving out sets of regression variables in higher dimensions.
and in our experience regression methods still exhibit relatively slow convergence in applications. In practical portfolio CVA computations, only about $10^3$ to $10^4$ paths are feasible, giving a high variance of estimators for both of these approaches, in particular for the sensitivities with respect to market factors. PDE schemes, on the other hand, are very efficient for low dimensions but suffer from the ‘curse of dimensionality’ – a term used to describe the computational complexity which increases exponentially with the number of risk factors – which made Monte Carlo methods the industry standard technique for problems with dimensions larger than two or three.

Notwithstanding this, we propose a method based on PDEs not just for the inner, but also for the outer expectation. We address the curse of dimensionality by decomposition of the conditional expectations for individual and groups of risk factors, hence breaking the high-dimensional problem up into a sequence of lower-dimensional problems, whose solutions are then assembled into a truncated decomposition. These low-dimensional problems assume all other risk factors follow a deterministic term structure, such that numerical PDE methods are available which give sufficiently accurate approximations in acceptable time. We will find that for practically sufficient accuracy, the approximation requires solving only one one-dimensional base approximation and multiple two- and three-dimensional approximations which are used as corrections. Perhaps practically most relevant, the approximated portfolio dynamics underlying these expansions can be used to construct control variates for Monte Carlo estimators. As such, we do not see this methodology to compete with the above simulation approaches, but to be used in conjunction with them, as it provides a way to combine seamlessly the high accuracy of PDE schemes in low dimensions and the robustness of Monte Carlo schemes in high dimensions.

The method draws on a variety of related previous work, but requires substantial extensions for the present context:

- in terms of application and problem structure, the method extends Reisinger and Wittum [85] from basket options under Black-Scholes to portfolio risk management under a general class of models;

- specifically, for the exposure estimation problem formulated as nested conditional expectations, we solve a combination of forward and backward Kolmogorov equations and carry out numerical integration of the solutions;

- compared to the dimension-wise integration method of Griebel and Holtz [49], the future distribution here is unknown, such that forward PDEs have to be approximated;

- compared to Reisinger and Wittum [85], Reisinger and Wissmann [83], we use original variables as factors instead of principal components; this makes
the results more easily interpretable;

- compared to Reisinger and Wissmann [84] which provides error bounds for
  the constant coefficient setting, we deal with the case of non-linear SDEs
  and variable-coefficient PDEs, which requires a generalised definition of the
  expansion;

- instead of theoretical error bounds, we assess the method on a complex
  real-world example, as outlined in the following.

For the numerical illustration, we consider typical FX portfolios consisting
of multiple financial derivatives, namely Interest Rate Swaps (IRS), Cross Cur-
rency Basis Swaps (CCYS) and FX options. The risk factors are driven by a
multi-dimensional Black-Scholes-2-Hull-White (BS2HW) model which captures
stochasticity in the FX and the interest rates. The model is calibrated to market
data and for these different portfolios we compute EE and EPE and compare our
results with a full scale Monte Carlo benchmark, demonstrating vastly superior
computational complexity.

5.2 Problem formulation

We consider the general framework of a financial market described by a \( d \)-
dimensional stochastic process \( X = (X^d_t)_{0 \leq t \leq T} \). In our applications, \( X \) will be given
by a stochastic differential equation of the form

\[
\begin{align*}
    dX^i_t &= \mu^i(X_t, t) \, dt + \sigma^i(X_t, t) \, dW^i_t, \quad i = 1, \ldots, d, \quad t > 0, \\
    X^i_0 &= a_i, \quad i = 1, \ldots, d,
\end{align*}
\]

(5.1)\hspace{1cm} (5.2)

where \( W \) is a \( d \)-dimensional standard Brownian motion with a given correlation
matrix \( (\rho_{ij})_{i \leq j \leq d} \), and \( a \in \mathbb{R}^d \) a given initial state, \( \mu_i \) is the drift, \( \sigma_i \) the volatility.
We assume that the processes are written under a risk-neutral measure \( Q \), which
is relevant for derivative valuation as well as regulatory CVA computations.

We now consider derivatives on \( X \). For simplicity we focus on the European
case here. Let the payoff function \( \phi_k \) determine the amount \( \phi_k(X_T) \) to be received
by the holder at time \( T_k \) for \( 1 \leq k \leq N_T \). Then the arbitrage-free (undiscounted)
value \( V_k \) of this claim for \( 0 \leq t \leq T \) is then

\[
V(x, t) = \mathbb{E}[\phi(X_T)|X_t = x]
\]

(where we have dropped the subscripts \( k \) for simplicity) and satisfies the Kol-
5.2. PROBLEM FORMULATION

Mogorov backward PDE (see, e.g. [77])

\[
\frac{\partial V}{\partial t} + \sum_{i=1}^{d} \mu_i \frac{\partial V}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{d} \sigma_i \sigma_j \rho_{i,j} \frac{\partial^2 V}{\partial x_i \partial x_j} = 0, \tag{5.3}
\]

\[V(x, T) = \phi(x).\]

For Counterparty Credit Risk (CCR), not only the time 0 value is important, but also the evolution of this value after the deal has been entered. In this work we use several measures to quantify CCR. Let

\[\sum_{k=1}^{N_D} w_k(t) V_k(X_t, t)\]

be the value of a portfolio of \(N_D\) derivatives at time \(t\) that a financial institution holds, where \(w_k, 1 \leq k \leq N_D\), are derivative portfolio weights. Here, we set \(w_k(t) = 0\) for \(t > T_k\), the expiry of the \(k\)-th derivative.

In this case, following [80], Expected Exposure (EE) is defined as the expected value of the portfolio at time \(t > 0\):

\[\text{EE}(t) = \mathbb{E} \left[ \sum_{k=1}^{N_D} w_k(t) V_k(X_t, t) \mid \mathcal{F}_0 \right]. \tag{5.4}\]

where \(\mathcal{F}_0\) is the filtration at time \(t = 0\). This metric shows the dynamics of the mean portfolio value. Similarly, the Expected Positive Exposure (EPE) at a future time \(t < T\) is given by the expectation of the positive portfolio value. Taken into account the portfolio weights, this comes down to:

\[\text{EPE}(t) = \mathbb{E} \left[ \max \left(0, \sum_{k=1}^{N_D} w_k(t) V_k(X_t, t) \right) \mid \mathcal{F}_0 \right]. \tag{5.5}\]

Note that the Expected Negative Exposure (ENE) can be computed by taking the minimum instead of the maximum in equation (5.5), or from \(\text{ENE} = \text{EE} - \text{EPE}\).

In this chapter, we focus solely on the computation of EE and EPE. For exposure calculations at time 0, we then consider functionals of the form

\[V(x, 0; t) = \mathbb{E} \left[ \sum_{k=1}^{N_D} w_k(t) V_k(X_t, t) \right| X_0 = x \mid \mathcal{F}_0 \right]. \tag{5.6}\]

\[= \mathbb{E} \left[ \sum_{k=1}^{N_D} w_k(t) \mathbb{E} \left[ \phi_k(X_{T_k}) | X_t \right] \right| X_0 = x \right]. \tag{5.7}\]

\(^1\)Here we do not include collateralisation. This can be added by including a collateral model on top of the exposure dynamics.
CHAPTER 5. CVA BY RISK FACTOR DECOMPOSITION

where \( E \) is an exposure function, e.g., \( E(x) = x \) in the case of EE, or \( E(x) = \max(x,0) \) in the case of EPE. This can be adapted in the obvious way for American or path-dependent options.

Remark 1. To find \( V(x,0; t) \), the expected (positive) exposure at time \( t \) as seen from time 0, we could therefore solve backward PDEs in \((t, T)\) for \( V_k(x,t) \), and then another backward PDE over \((0, t)\) with \( E[\sum_{k=1}^{N_D} w_k(t)V_k(\cdot,t)] \) as terminal condition. This would however require the solution of a different backward PDE in the second stage for different \( t \). To avoid this, we utilise a forward PDE for the transition densities to compute exposures.

The reason we use the backward PDE to approximate \( V_k \) instead of using the density function, is that we need the value of derivatives in the whole state space reachable by \( X \).

Let \( p(x,t;a,0) \) denote the transition density function of \( X_t \) at \( x \) given state \( a \) at time \( t = 0 \), then

\[
V(x,0; t) = \int_{\mathbb{R}} p(y,t;x,0)E\left[ \sum_{k=1}^{N_D} w_k(t)V_k(y,t) \right] dy. \tag{5.8}
\]

Here, \( p \) is given through an adjoint relation, by the Kolmogorov forward equation:

\[
-\frac{\partial p}{\partial t} - \sum_{i=1}^{d} \frac{\partial}{\partial x_i} (\mu_i p) + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j} \left( \sigma_i \sigma_j \rho_{i,j} p \right) = 0, \tag{5.9}
\]

where \( \delta \) is the Dirac distribution centered at 0.

Hence, we can obtain expected (positive) exposures by the solution of one forward PDE for the density and one backward PDE for each derivative in the portfolio, plus one integration (for each \( t \)).

The principal difficulty in practice arises from the dimensionality \( d \) of \( X \). Although \( N_D \) is typically very large, the computational complexity is linear in \( N_D \), while it is exponential in \( d \). In the next section, we introduce an approximation technique which makes large \( d \) computationally manageable.

5.3 Approximation by risk factor decomposition

To introduce the concepts, we focus in Section 5.3.1 on the problem of approximating

\[
\begin{align*}
V(x, t) & = E[\phi(X_T)|X_t = x] \tag{5.10} \\
& = \int_{\mathbb{R}^d} \phi(y)p(y,T;x,t) dy, \tag{5.11}
\end{align*}
\]
where \( p(T; x, t) \) is the probability density function of \( X_T \). We first discuss an extension of the anchored-ANOVA concept to PDEs with variable coefficients and a “moving anchor”, in Section 5.3.2 we explain how the approximations can be used as effective control variates for unbiased Monte Carlo estimators. The application to the complete expected exposure problem (5.8) will be discussed in Section 5.3.3.

5.3.1 An anchored-ANOVA-type approximation

The proposed method extends anchored-ANOVA decompositions as considered by [49] in the context of integration problems. In the setting of (5.11), the methodology in [49] defines an anchor \( a \in \mathbb{R}^d \) and then, for a given index set \( u \subseteq \mathcal{D} = \{ i : 1 \leq i \leq d \} \), defines projections of the integrand \( f = \phi \) by \( f(a \setminus x_u) \), where \( a \setminus x_u \) denotes a \( d \)-vector such that

\[
(a \setminus x_u)_i = \begin{cases} x_i & i \in u, \\ a_i & i \notin u. \end{cases}
\]  

(5.12)

This leads to a decomposition of \( f \) into lower-dimensional functions which can be exploited for successive quadrature approximations (see [49] and the references therein).

In contrast to there, we cannot assume here that the joint probability density function for \( X_T \) is analytically known and therefore we have to consider the dynamics of \( X \). To improve the accuracy, e.g. to account for the term-structure of interest rates, here we do not concentrate the measure at a fixed anchor \( a \), but we allow the anchor to move with a certain conditional expectation of the underlying process, as detailed below.

We first define a deterministic approximation to \( X \) by

\[
\xi(t) = \mathbb{E}\left[ \frac{X_t}{X_t^u} \right],
\]

\[
\bar{X}_t^i = a_i + \int_0^t \mu_i(a_1, \ldots, a_{i-1}, \bar{X}_s^i, a_{i+1}, \ldots, a_d, s) \, ds + \int_0^t \sigma_i(a_1, \ldots, a_{i-1}, \bar{X}_s^i, a_{i+1}, \ldots, a_d, s) \, dW_i^s,
\]

(5.13)

In [83], the drift is first approximated by \( \mu(a, 0) \), and then eliminated by a coordinate transformation \( x \rightarrow x - \mu(a, 0) t \). The above construction takes into account more information about the term structure.

For given subset \( u \subseteq \mathcal{D} \), we then define a process

\[
X^u_t = \begin{cases} a_i + \int_0^t \mu_i(\xi(t), X^u_s, s) \, ds + \int_0^t \sigma_i(\xi(t), X^u_s, s) \, dW^i_s & i \in u, \\ \xi(t), & i \notin u, \end{cases}
\]

(5.13)
where \( \xi(t) \), \( X^u_t \) follows the notation from (5.12).

Here, we have replaced a subset of the processes by their conditional expectations, where all remaining coordinates are kept constant at their initial value. For instance, if \( X^1 \) is an exchange rate and \( X^2 \) and \( X^3 \) are the domestic and foreign short rate in a Hull-White model, then \( \xi^3 \) is the expectation of the exchange rate under a constant interest rate model, while \( \xi^2 \) is simply the expectation of the domestic short rate. The point is that the dynamics is effectively of dimension \(|\upsilon|\), and the expectation (5.10) can be approximated by lower-dimensional problems.

Accordingly, we define

\[
V_u(a; x, t) = \mathbb{E}[\phi(X^u_T)|X^u_t = x],
\]

which depends on \( a \) implicitly through the definition of \( X^u_t \) in (5.13).

Specifically, the backward PDE for (5.14) under (5.13) is

\[
\frac{\partial V_u}{\partial t} + \sum_{i \in u} \mu_i(\xi(t)|x_u) \frac{\partial V_u}{\partial x_i} + \frac{1}{2} \sum_{i,j \in u} \sigma_i(\xi(t)|x_u) \sigma_j(\xi(t)|x_u) \rho_{ij} \frac{\partial^2 V_u}{\partial x_i \partial x_j} = 0,
\]

\[
\text{subject to} \quad V_u(x, T) = \phi(x),
\]

The significance of the arguments \( \xi(t)|x_u \) in the coefficients of (5.15) is that the full information on the coordinates \( x_j, j \in u \) is used, while the solution \( V(a; a, 0) \) at the anchor point can be computed solving only a \(|a|\)-dimensional PDE in the variables \( x_j, j \in u \).

Remark 2. The generalisation from (5.14) to

\[
V_u(a; x, t) = \mathbb{E}[\exp(-\int_t^T \rho(X^u_s) \, ds) \phi(X^u_T)|X^u_t = x]
\]

is straightforward and the corresponding backward PDE instead of (5.15) is

\[
\frac{\partial V_u}{\partial t} + \mathcal{L}_u V_u + \rho(\xi(t)|x_u) V_u = 0.
\]

For future reference, we also have the forward PDE

\[
\frac{\partial p_u}{\partial t} + \sum_{i \in u} \frac{\partial}{\partial x_i} (\mu_i(\xi(t)|x_u) p_u) - \frac{1}{2} \sum_{i,j \in u} \frac{\partial^2}{\partial x_i \partial x_j} (\sigma_i(\xi(t)|x_u) \sigma_j(\xi(t)|x_u) \rho_{ij} p_u) = 0,
\]

\[
p_u(x, 0; a, 0) = \delta(x - a).
\]
5.3. APPROXIMATION BY RISK FACTOR DECOMPOSITION

From here onwards, we can follow the path of [49] and the references therein, to define a decomposition recursively through a difference operator $\Delta$, by $\Delta V_\emptyset = V_\emptyset$ and, for $u \neq \emptyset$,

$$\Delta V_u(a; x, t) = V_u - \sum_{w \subset u} \Delta V_w = \sum_{w \subset u} (-1)^{|w| - |u|} V_w,$$  \hspace{1cm} (5.19)

where the inclusion in the first summation is strict. This is indeed a decomposition because

$$V(a; x, t) = \sum_{u \subseteq \{1, \ldots, d\}} \Delta V_u = \sum_{k=0}^{d} \sum_{|u|=k} \Delta V_u.$$  \hspace{1cm} (5.20)

We note that for $u = \{i_1, \ldots, i_{|u|}\}$, $\Delta V_u(a; \xi(t)\setminus x_u, t)$ only depends non-trivially on the sub-set of coordinates $\{x_{i_1}, \ldots, x_{i_{|u|}}\}$, and satisfies (5.15). Therefore, $V_u$ and hence $\Delta V_u$ can be found by the solution of problems of dimension not higher than $|u|$.

As an example, consider $d = 3$ and $u = \{1, 2\}$, then

$$\Delta V_u(a; x, t) = V_{\{1, 2\}} - (\Delta V_{\{1\}} + \Delta V_{\{2\}} + \Delta V_\emptyset)$$
$$= V_{\{1, 2\}} - V_{\{1\}} - V_{\{2\}} + V_\emptyset.$$

We can now define an approximation by

$$V_{0,s}(a; x, t) = \sum_{k=0}^{s} \sum_{|u|=k} \Delta V_u = \sum_{k=0}^{s} c_k \sum_{|u|=k} V_u,$$  \hspace{1cm} (5.21)

where $c_k$ are integer constants, which also depend on $s$ and $d$ (we suppress this to keep the notation simple). The point is that the approximation $V_{0,s}(a; x, t)$ at the anchor can be found by solving PDEs of dimension at most $s$.

**Remark 3.** The approximations $V_{0,1}$ and $V_{0,2}$ have certain similarities with delta and delta-gamma approximations, respectively (see, e.g., [3]). In the latter, the derivative value is approximated by

$$V(X_t, t) \approx V(a, 0) + \frac{\partial V}{\partial t} t + (X_t - a)^T \nabla V + \frac{1}{2} (X_t - a)^T \nabla \nabla^T V(X_t - a),$$

where $a = X_0$, $\nabla V$ is the gradient or delta and $\nabla \nabla^T V$ the Hessian or gamma. One notable difference is that $V_{0,2}$ does not make any approximations if $V$ is two-dimensional or a sum of two-dimensional functions.
An alternative approximation, extending (5.21), is given by
\[
V_{r,s}(a;x,t) = \sum_{k=0}^{\min(r,s)} c_k \sum_{|a|=k} V_{d(1,...,r)}, \quad r + s \leq d. \tag{5.22}
\]
Here, we always retain the first \(r\) coordinates, and apply the splitting only to the remaining \(d-r\) coordinates. The approximation \(V_{e,s}(a;a,t)\) at the anchor can be found by solving PDEs of dimension at most \(r+s\). It is this approximation that we will use in the numerical computations later on in the chapter, with \(r = 1\) and \(s = 1\) or \(s = 2\). The coordinate \(x_1\) is chosen to capture most of the dynamics, either through prior knowledge or small pilot runs with reduced accuracy.

For instance, for \(s = 1\), \(r = 1\), we have
\[
V_{1,1}(a;x,t) = \sum_{1 < i \leq d} (V_{1,i} - V_{1}) + V_{1},
\]
i.e., \(c_0 = -(d - 2), c_1 = 1\) in (5.22), and, for \(s = 2, r = 1\),
\[
V_{1,2}(a;x,t) = \sum_{1 < i < j \leq d} (V_{1,i,j} - V_{1,i} - V_{1,j} + V_{1}) + \sum_{1 < i \leq d} (V_{1,i} - V_{1}) + V_{1},
\]
i.e., \(c_0 = \binom{d-1}{2}, c_1 = -(d - 2), c_2 = 1\).

### 5.3.2 A control variate

In some situations, the approximation (5.22) with small \(r\) and \(s\) will be sufficiently accurate. There is, however, no guarantee that increasing either \(r\) or \(s\) will improve the accuracy, or the computation may be prohibitively expensive due to the high dimensionality and number of the PDEs involved. In such cases, it will be valuable to simulate correction terms by a MC method – instead of computing them by PDEs – to obtain more accurate approximations.

Therefore, we now discuss a way to turn the approximation (5.22) into a control variate for a Monte Carlo scheme for (5.10). Denote by \(\omega_k : 1 \leq k \leq N_\omega\) a set of \(N_\omega\) independent samples of the \(d\)-dimensional Brownian motion \(W\), and \(X_T(\omega_k)\) the (strong) solution to (5.1) for a given sample path \(\omega_k\) of \(W\).

The standard estimator of \(V(x,0)\) in (5.10) is then
\[
\hat{V}^{N_\omega} = \frac{1}{N_\omega} \sum_{k=1}^{N_\omega} \phi(X_T(\omega_k)).
\]
5.3. APPROXIMATION BY RISK FACTOR DECOMPOSITION

Denote by Φ = ϕ(Xₜ) the random payoff, i.e., for event Nₜ we have Φ(ωₜ) = ϕ(Xₜ(ωₜ)), and

\[ \Phi_{r,s} = \sum_{k=0}^{s} c_k \sum_{|u|=k} \phi(X^{(i_1,\ldots,i_r)}) \]

(compare with (5.22)), then we define the estimator

\[ \hat{V}^{N_{r,s}} = \frac{1}{N_{r,s}} \sum_{k=1}^{N_{r,s}} (\Phi(\omega_k) - \alpha (\Phi_{r,s}(\omega_k) - V_{r,s})) \],

(5.25)

where V_{r,s} is given by (5.22) and α to be determined later to achieve the best variance reduction. By

\[ V = E[\phi(Xₜ)], \quad V^{(i_1,\ldots,i_r)} = E[\phi(X^{(i_1,\ldots,i_r)})], \]

we get

\[ E[\hat{V}^{N_{r,s}}] = V, \]

i.e., the standard estimator and the control variate estimator are both unbiased.

Typically, we will find that \( \alpha \approx 1 \) is optimal. Indeed, if \( \alpha = 1 \), then

\[ \hat{V}^{N_{r,s}} = V_{r,s} + \frac{1}{N_{r,s}} \sum_{k=1}^{N_{r,s}} (\Phi(\omega_k) - \Phi_{r,s}(\omega_k)) \],

(5.26)

i.e., we sample the approximation error of the truncated decomposition. The motivation for choosing \( \hat{V}^{N_{r,s}} \) over the standard estimator \( \hat{V}^{N_{r,s}} \) is that for given \( \omega_k \), \( \Phi_{r,s} \) and \( \Phi \) will be close and therefore the variance much reduced compared to the standard estimator.

The value of \( \alpha \) which minimises the variance is determined by the co-variances of the control variate and the standard estimator, which can be approximated by the estimators (see [44])

\[ \hat{\sigma}^2 = \frac{1}{N_{r,s}} \sum_{k=1}^{N_{r,s}} (\Phi_{r,s}(\omega_k) - V_{r,s})^2, \quad \hat{\rho} = \frac{1}{N_{r,s}} \sum_{k=1}^{N_{r,s}} (\Phi(\omega_k) - \Phi_{r,s}(\omega_k) - V_{r,s}) \]

and the optimal value is estimated as \( \alpha = \hat{\rho}/\hat{\sigma}^2 \). The variance reduction is then approximately

\[ \frac{\text{Var}[\Phi - \alpha (\Phi_{r,s} - V_{r,s})]}{\text{Var}[\Phi]} \approx 1 - \frac{\text{cov}(\Phi, \Phi_{r,s})}{\sqrt{\text{Var}(\Phi)\sqrt{\text{Var}(\Phi_{r,s})}}}. \]

(5.27)

Summarising, in situations where the approximation \( V_{r,s} \) from the previous section is not accurate enough, it can be corrected with a relatively small number of Monte Carlo samples of the correction terms (the right-hand sum) in (5.26).
5.3.3 Application to derivative portfolios

We now discuss approximations to expected exposures, as per (5.7). To benefit from the dimension reduction of risk-factor decomposition, we compute the solutions of all PDEs involved in the spirit of Section 5.3.1.

The general principle is to replace the process \( X \) in (5.7) by \( X^v \) as defined in (5.13). Thus, we approximate \( V \) in (5.8) by

\[
V_u(u; x, t) = \int_R p_u(y, t; x, 0) \sum_{j=1}^{N_u} V_{j,u}(y(t)|y_u, t) dy,
\]

where \( p_u \) is the transition density function of \( X^u \), which satisfies the forward PDE (5.17), (5.18) instead of (5.9). Similarly, \( V_{k,u} \) is given as the solution to (5.15), with payoff function \( \phi_k \) in (5.16). The key point is that we compute \( p_u \) and \( V_{k,u} \) by solving \( |u| \)-dimensional PDEs, and (5.28) is a \( |u| \)-dimensional integration problem, as \( p \) is a Dirac measure in dimensions \( D \setminus u \).

Then \( V_{r,s} \) is defined by (5.28) and (5.22). The complexity of the whole computation is linear in \( N \), and exponential in \( r + s \).

For products where model prices are available in closed form (such as most swaps), we will use those in the computations for speed and accuracy. For others (such as options), we will use numerical PDE approximations as outlined in the following section.

5.4 Numerical approximation of the Kolmogorov PDEs

Here, we outline the finite difference method used for the PDE solutions, as it is somewhat non-standard through the use of forward and backward equations. In particular, we extend the adjoint method from Itkin [65] from two to three dimensions and use it for the forward component. For a more general introduction to finite difference methods for pricing financial derivatives, see for example Tavella and Randall [97] and, specifically for interest rate derivatives Andersen and Piterbarg [4].

As usual, a grid is defined with one dimension per risk factor, and the partial derivatives in the Kolmogorov forward (5.9) or backward PDE (5.3) are approximated by finite differences on this grid.

5.4.1 Backward equations

We use a combination of second order central differences in the centre of the spatial domain, and upwinding for stabilisation of large drifts in outer parts of the
5.4. NUMERICAL APPROXIMATION OF THE KOLMOGOROV PDES

Let $U(\tau)$ be the solution to this semi-discrete Kolmogorov backward equation at time $\tau$, where $\tau = T - t$, then we have the following initial value problem

$$\frac{dU}{d\tau} = F(\tau)U, \quad \tau \geq 0, \quad U(0) = U_0,$$

(5.29)

with given matrix $F$, derived from the PDE, and initial vector $U_0$ given by the payoff at the mesh points. The boundary conditions are also derived from the payoff function.

We apply an Alternating Direction Implicit (ADI) splitting method. Let $F^n = F(\tau_n)$ be the discretisation matrix, as per (5.29), at time step $\tau_n = n\Delta \tau$, $n \in \mathbb{N}$, $\Delta \tau > 0$ a uniform step size, then this matrix is first decomposed into

$$F^n = F^n_0 + F^n_1 + F^n_2 + F^n_3,$$

where the individual $F^n_i$, $1 \leq i \leq d$ contains the contribution to $F$ stemming from the first and second order derivatives in the $i$th dimension. Following in ‘t Hout and Foulon [62], we define one matrix $F_0$ which accounts for the mixed derivative terms, and treat that fully explicitly. Here we show the scheme for a three-dimensional problem, and the two-dimensional case is obtained as special case by setting $F_3 = 0$. The Hundsdorfer-Verwer (HV) scheme [61],

$$Y_0 = U_{n-1} + \Delta \tau F^{n-1} U_{n-1},$$

$$(I - \theta \Delta \tau F^n_i) Y_j = Y_{j-1} - \theta \Delta \tau F^{n-1} U_{n-1}, \quad j = 1, 2, 3,$$

$$\bar{Y}_0 = Y_0 + \frac{1}{2} \Delta \tau \left[ F^n Y_3 - F^{n-1} U_{n-1} \right],$$

$$(I - \theta \Delta \tau F^n_i) \bar{Y}_j = Y_{j-1} - \theta \Delta \tau F^n Y_j, \quad j = 1, 2, 3,$$

$$U_n = \bar{Y}_3,$$

defines a second order consistent ADI splitting for all $\theta$, and can be shown to be von Neumann stable for $\theta \in \left[ \frac{1}{2} + \frac{1}{6} \sqrt{3}, 1 \right]$, see Haentjens and in ‘t Hout [55]. We use $\theta = 0.8 > \frac{1}{2} + \frac{1}{3} \sqrt{3} \approx 0.789$ in the computations. Accuracy and stability do not appear to be very sensitive to the choice of $\theta$ within the above range.

5.4.2 Forward equations

We use the adjoint relation between the Kolmogorov forward and the backward PDE. It can easily be seen that (for sufficiently smooth coefficients)

$$\frac{dP}{dt} = F^T(t)P, \quad t \geq 0, \quad P(0) = P_0,$$

(5.30)
with \( F \) from above, is a consistent scheme for the Kolmogorov forward equation. The initial datum for the discrete density function \( P \) is given by an approximation \( P_0 \) to the Dirac delta. We choose a mesh such that a mesh point coincides with the location of the Dirac delta. Then \( P_0 \) is set to zero for all other mesh points and to a large value at this particular point, chosen such that a numerical quadrature rule applied to \( P_0 \) gives 1. At the spatial boundaries, we apply zero Dirichlet conditions.

Hence, following Andreasen and Huge [6] and Capriotti et al. [24], we can first set up matrices that approximate the partial derivatives in the Kolmogorov Backward PDE, and use the transpose of these matrices for the Kolmogorov forward PDE.

Now we could apply HV splitting to (5.30) and obtain a second order (in time) consistent approximation to the forward PDE. Instead, we follow Itkin [65] to calculate the exact adjoint of the HV scheme for the backward equation, which results in a different scheme, i.e., the transposition and approximate factorisation do not commute. The forward scheme, adapted to the three-dimensional case from the two-dimensional case analysed in Itkin [65], then is

\[
(I - \theta \Delta t F_n^3)^T Y_0 = P_{n-1},
\]
\[
(I - \theta \Delta t F_n^2)^T Y_1 = Y_0,
\]
\[
(I - \theta \Delta t F_n^1)^T Y_2 = Y_1,
\]
\[
\tilde{Y}_0 = P_{n-1} + \Delta t \left( \frac{1}{2} (F_n^3)^T - \theta (F_n^3)^T Y_2 - \theta (F_n^2)^T Y_1 - \theta (F_n^1)^T Y_0 \right)
\]
\[
(I - \theta \Delta t F_n^2)^T \tilde{Y}_1 = \tilde{Y}_0,
\]
\[
(I - \theta \Delta t F_n^1)^T \tilde{Y}_2 = \tilde{Y}_1,
\]
\[
(I - \theta \Delta t F_n^0)^T \tilde{Y}_3 = \tilde{Y}_2,
\]
\[
P_n = \tilde{Y}_3 + \Delta t \left( ((F_n^{σ-1})^T - \theta (F_n^{σ-1})^T \tilde{Y}_3 - Y_2) - \theta (F_n^{σ-1})^T (\tilde{Y}_2 - Y_1) \right)
\]
\[
- \theta (F_n^{σ-1})^T (\tilde{Y}_1 - Y_0) + \frac{1}{2} (F_n^{σ-1})^T \tilde{Y}_3.
\]

By using this scheme, we have the adjoint relation

\[
P_n^T U_n = P_n^T U_0,
\]

i.e., using the backward equation to compute the conditional expectation function and then evaluating it at the location of the Dirac measure (left-hand side), or computing the density and then integrating over the “payoff” function, gives exactly the same result. Apart from the mathematical aestheticism of using the exact adjoint, this scheme has the following stability advantages. As discussed in In’t Hout and Wyns [64], splitting schemes do not give stable, and hence
convergent, solutions for Dirac initial data. A direct splitting of (5.30) is prone to wildly oscillating densities and may give meaningless expectations. By using the above scheme instead, we can be sure that even if we have no guarantee for the stability of the solution, the derived quantity of interest, e.g., EPE, is the same as for the backward equation with regular (for EPE, Lipschitz and piecewise smooth) data.\(^2\) The adjoint equation also ensures conservation of discrete probability in the interior of the mesh. The loss of mass at the boundaries is minimised by choosing the domain large enough.

With the grid of discrete transition probabilities, obtained by the numerical solution of the Kolmogorov forward PDE, and a grid of portfolio values from the backward PDE approximations, the exposures can be extracted by numerical quadrature applied to (5.8).

5.5 Case studies

In this section, we present the market set-up for the detailed numerical study in Section 5.6, where we compute exposures of portfolios with increasing complexity. The example consists of various exchange rate and interest rate products, and is chosen to be representative for the scenario where each product in a portfolio is exposed to one or more of a pool of risk factors. So for instance an exchange option of medium term maturity is exposed to interest rate risk in both the domestic and foreign markets, and conversely the domestic interest rate affects an exchange option as well as, say, a domestic zero coupon bond.

5.5.1 Driving risk factors

We consider a portfolio of FX and interest rate products. Each FX rate is assumed to be governed by a full Black-Scholes-2-Hull-White (BS2HW) model (see, e.g. Clark [27]). Every FX rate \(F^i_t\) thus has a stochastic domestic and foreign interest rate. For clarity of exposition, we consider the case where all exchange rates are relative to a single currency, in the case below the EUR, so that we can write the joint dynamics of FX and interest rates as

\[
\begin{align*}
    dF^i_t &= (R^d_t - R^f_t)F^i_t dt + \sigma^i(t)F^i_t dW^F^i_t, & 1 \leq i \leq m, \\
    dR^d_t &= \lambda_d(\Theta^d(t) - R^d_t) dt + \eta_d dW^d_t, \\
    dR^f_t &= \left[\lambda^f_i(\Theta^f_i(t) - R^f_t) - \eta^f_i(\rho_{i,i+1} + 1)\sigma(t)\right] dt + \eta^f_i dW^{L^f}_t, & 1 \leq i \leq m,
\end{align*}
\]  

\(^2\)But see Remark 1 why we are not using the the backward equation directly to compute exposures.
CHAPTER 5. CVA BY RISK FACTOR DECOMPOSITION

with given Brownian motions

\[(W^i)_{1 \leq i \leq 2m+1} = (W^{F,1}, W^{d,1}, W^{F,2}, W^{d,2}, W^{F,3}, W^{d,3}),\]  
with

\[d[W^i, W^j]_t = \rho_{i,j} \, dt, \quad \text{for } 1 \leq i, j \leq 2m + 1.\]

Here, \(R^d\) is the domestic short rate and \(R^{f,i}\) the foreign short rate for exchange rate \(F_i\), \(i = 1, \ldots, m\), with \(m + 1\) the total number of markets.

The interest rates are modeled by a mean-reverting Hull-White process with \(\Theta_d(t)\) and \(\Theta_f(t), 1 \leq i \leq m\), designed to fit the forward rate curve in the respective markets. More details on the calibration will be given in Section 5.5.3.

We do not consider stochastic volatility models such as Heston or SABR, as it is not straightforward to extend these models from a single currency pair to multiple pairs, while preserving important FX characteristics like symmetry and the triangle inequality (see Doust [35], De Col et al. [29]).

In the numerical examples, we choose EURUSD, EURGBP and EURJPY, i.e. \(m = 3\). Together with the interest rates in the EUR (domestic) and USD, GBP and JPY (all foreign) markets, this gives \(d = 2m + 1 = 7\) risk factors. This example is complex enough in the sense that a 7-dimensional PDE solution is not feasible and that we can demonstrate the effect of different decompositions.

5.5.2 Derivatives

We consider portfolios of cross-currency and interest rate swaps and FX options.

**Zero-coupon bonds** Although the portfolios below do not contain any bonds explicitly, the swaps studied here are all based on coupon legs which can be replicated by (and hence valued from the time 0 value of) a string of zero-coupon bonds that pay 1 at different times \(T\). The time \(t\) value a zero-coupon bond, \(P(t,T)\), under Hull-White for the short rate \(R\) is

\[P(t,T) = \frac{P(0,T)}{P(0,t)}e^{A(t,T) - B(t,T)R_t},\]  
where

\[A(t,T) = B(t,T)f(t) - \frac{B^2(t,T)}{4\lambda}(1 - e^{-2\lambda t}), \quad B(t,T) = \frac{1 - e^{-\lambda(T-t)}}{\lambda},\]
and

\[f(t) = \frac{-\log(P(0,t))}{t}, \quad t \geq 0,\]
see Filipovic [43]. For simplicity, we assume here a single curve framework and do not make a distinction between discounting and forwarding curves.
Cross-currency swap (CCYS) We consider a series of FX forward swaps settled in arrears, with floating payment and receiving leg. Note that both legs need to be valued in the domestic currency, for which the future exchange rate is used. Following Filipovic [43], the value $V_{ccy}(t, T_0, T)$ of the swap at time $t \leq T_0$ with payment dates $T_1, \ldots, T_N$ is

$$V_{ccy}(F_t, t; T_0, T) = C^d(t, T_0, T) - M F_t C^f(t, T_0, T),$$

where $F_t$ is the FX rate at time $t$, $C^d(t, T_0, T)$ and $C^f(t, T_0, T)$ are the domestic and foreign floating rate notes, and $M$ is the moneyness as a percentage of the (future) FX rate, e.g. 100% is referred to as At-The-Money (ATM), 105% as In-The-Money (ITM) and 95% as Out-of-The-Money (OTM). The values of the floating rate notes are

$$C^d(t, T_0, T) = P^d(t, T_0) \Lambda^d + \sum_{i=1}^{N_C} \left( P^d(t, T_{i-1}) - P^d(t, T_i) \right) \Lambda^d,$$

$$C^f(t, T_0, T) = P^f(t, T_0) \Lambda^f,$$

where $\Lambda^d = \Lambda^d/F_0$.

Interest rate swap Here, we consider a floating versus fixed interest rate swap, where the fixed leg is defined as a fixed coupon bond, such that the value with notional $\Lambda$ equals

$$V_{irs}(t) = \left( K \Delta T \sum_{i=1}^{N_C} P(t, T_i) + P(t, T) - P(t, T_0) \right) \Lambda,$$

where $K$ is the fixed rate and $\Delta T = T_i - T_{i-1}$ the interval between payment dates (see Filipovic [43]).

FX option Finally, we consider an FX call option on the EURUSD FX rate with strike $K$, notional $\Lambda_{opt}$ and maturity $T$ with payoff

$$\phi(F_T) = \max(0, F_T - K) \Lambda_{opt}.$$

The value of the option at time $t < T$, under the BS2HW model from Section 5.5.1, is equal to

$$V_{opt}(F_t, R^d_t, R^f_t, t) = E[\exp(-\int_t^T R^d_s \, ds) \phi(F_T)|X_t],$$

where $X_t = (F_t, R^d_t, R^f_t)$ follow (5.31). For more details on these options we refer to Filipovic [43] and Brigo and Mercurio [16]. We only note that the option value function is the solution of the corresponding three-dimensional backward PDE; see also Remark 2.
CHAPTER 5. CVA BY RISK FACTOR DECOMPOSITION

<table>
<thead>
<tr>
<th>Moneyness</th>
<th>EURUSD</th>
<th>EURGBP</th>
<th>EURJPY</th>
<th>IR swap</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>100%</td>
<td>95%</td>
<td>105%</td>
<td>100%</td>
</tr>
<tr>
<td>T</td>
<td>5Y</td>
<td>3Y</td>
<td>2Y</td>
<td>5Y</td>
</tr>
<tr>
<td>Λ</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>150</td>
</tr>
<tr>
<td>N</td>
<td>100</td>
<td>60</td>
<td>40</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 5.1: Parameters for the CCYS and IRS.

Contract parameters The CCYSs are driven by the FX rate $F$ and the foreign and domestic interest rate processes. The contract specific parameters are shown in Table 5.1. Note that the trades have different maturity, notional and moneyness. This moneyness is set as percentage of the (future) FX rate, e.g. using 105% is referred to as In-The-Money (ITM).

The interest rate swap is traded on the EUR interest rate, at-the-money, where the swap rate is such that the initial value of the trade at inception ($t = 0$) is equal to zero. The specific parameters can be found in Table 5.1.

The option strike is set at $K = 0.95F_0^1$ and the maturity to $T = 4$. The notional of this option is set to $Λ_{opt} = 100$.

5.5.3 Market and calibration

As the behaviour and accuracy of the approximation method from Section 5.3 may depend significantly on the model parameters, we perform a careful calibration to market data before undertaking numerical tests.

We use a data set from 2 December 2014. At this time, the interest rates were low, which resulted in low yields for all markets. Shown in Figure 5.1(b) are the yields up to 5 years, computed by polynomial interpolation of the market forward curve $f(t)$ as per (5.35). They are used for bond pricing via (5.34) and to calibrate $Θ ∈ \{Θ_{d}(t), Θ^{i}_{f}(t), 1 ≤ i ≤ m\}$ in (5.31) by

$$Θ(t) = \frac{1}{λ} \frac{df(t)}{dt} + f(t) + \frac{σ^2}{2λ^2} (1 - e^{-2λt})$$

with $f(t)$ from (5.35), which gives an exact fit to bond prices (see Filipovic [43]).

For the FX volatilities in (5.31a), we assume a piecewise constant volatility function which follows the ATM volatilities over time,

$$σ(t) = \sum_{i=1}^{N_M} σ_{i} 1_{(T_{i-1} < t ≤ T_i)},$$

where $0 = T_0 < \ldots < T_{N_M} = T_{max}$ is a partition of the interval $[0,T_{max}]$ for some maximum maturity $T_{max}$ such that $σ_t$ is the volatility corresponding to the time
5.5. CASE STUDIES

Figure 5.1: Volatilities and yields over time on 2 December 2014.

The resulting different volatility levels are shown in Figure 5.1(a).

The Hull-White parameters are calibrated to fit the prevailing yield curve and swaption data using analytic expressions, i.e. θ is calibrated by (5.36) and η and λ by a least-squares fit to co-terminal swaptions that terminate in 10 years (1×9, 2×8, 3×7, 4×6, 5×5, 6×4, 7×3, 8×2 and 9×1), as these swaptions can also be used to replicate the CVA of a swap that matures in 10 years, as is shown in Sorensen and Bollier [94].

The correlations between the observable factors (exchange and interest rates) are estimated using weekly historical time-series data from the preceding 3 year period. The estimated correlation matrix is regularised to be positive semi-definite by setting any negative eigenvalues in a singular value decomposition to zero [82]. The resulting full correlation matrix is shown in Appendix B.4 together with the other parameters.

Such a step-by-step calibration is industry practice in CCR modeling.
5.6 Results

In this section, we analyze the accuracy of the numerical approximations from Section 5.3, applied to the market described in Section 5.5. We compare our results to an accurate approximation computed by a brute-force Monte Carlo method. In particular, we study the relative differences expressed in percentages,

\[
\epsilon_{L_2} = 100 \frac{\sqrt{\sum_{i=1}^{N_T} (V_{rs}(t_i) - V_{MC}(t_i))^2}}{\sqrt{\sum_{i=1}^{N_T} V_{MC}(t_i)^2}},
\]

\[
\epsilon_{\infty} = 100 \frac{\max_{i=1}^{N_T} |V_{rs}(t_i) - V_{MC}(t_i)|}{\max_{i=1}^{N_T} |V_{MC}(t_i)|},
\]

where \( r = 1 \) and \( s = 1 \) or \( s = 2 \) to denote two- or three-dimensional corrections (see (5.22)). For the computations, \( t_i \) are chosen to coincide with the swap payment dates, so that \( N_T = N_C \) is the total number of coupon payments. In addition, we study the mean difference (MD) over time relative to the sum of all notionals (\( N_{total} \)) in the portfolio expressed in basis points (0.01%),

\[
MD = \frac{10000}{N_T N_{total}} \sum_{i=1}^{N_T} |V_{rs}(t_i) - V_{MC}(t_i)|.
\]

Similarly, the normalized standard error for EE is defined as

\[
SE = 100 \frac{\sqrt{\sum_{i=1}^{N_T} SE(t_i)^2}}{\sqrt{\sum_{i=1}^{N_T} EE(t_i)^2}}, \quad SE(t_i) = \frac{Std(EE(t_i))}{\sqrt{N}}
\]

where \( Std(EE(t_i)) \) is the standard deviation of \( EE(t_i) \), and analogous definition for EPE.

The settings used for the numerical methods (domain and mesh sizes, time steps etc) are reported in Appendix B.5.

In the following, we use up to seven risk factors, \( d = 7 \), where (see (5.1) and (5.31))

\[
(X^d_1, \ldots, X^d_7) = (F^d_1, R^d, F^d_2, R^d_3, F^d_3, R^d_4)
\]

are the EURUSD, EURGBP, EURJPY exchange rates \( F^d \), and EUR, USD, GBP, JPY short rates \( R^d \) and \( R^{d-i} \), respectively.

To simplify the notation (5.28), we will use, e.g., the shorthand

\[
V(F^2, R^d, R^{d-3}) \equiv V_{(2,4,7)}(F^2_0, R^d_0, R^{d-3}_0, F^d_0, R^{d-3}_0, t) = V_{(4,2,7)}(X^4_0, X^2_0, X^7_0, X^4_0, X^2_0, X^7_0, t),
\]
where we suppress the dependence on the anchor $a = X_0$ and time $t$, and it is understood implicitly that the arguments identify the function being used. Also note that the function is evaluated at $X_0$, since we are interested in exposures conditional on the current state $\mathcal{F}_0$.

### 5.6.1 Case A: Single EURUSD CCYS

As a first test case, we focus on the BS2HW model (Section 5.5.1) for a single CCY swap trade on EURUSD. In this case, only the three relevant risk factors $F^1, R^d$ and $R^f$ (the EURUSD exchange rate, the EUR and USD short rates) are used, so effectively $d = 3$.

In that case

$$V_{1,1} = V(F^1) + \left( V(F^1, R^d) - V(F^1) \right) + \left( V(F^1, R^f) - V(F^1) \right)$$

$$V_{1,2} = V_{1,1} + \left( V(F^1, R^d, R^f) - V(F^1, R^d) - V(F^1, R^f) + V(F^1) \right)$$

The first term in the first line is a one-dimensional approximation with only the EURUSD rate stochastic, and then corrects separately for stochastic domestic (EUR) and stochastic foreign (USD) interest rate, respectively. Note how in this particular case equation (5.40b) with the three-dimensional corrections simplifies to the exact solution (5.40c).

In Figures 5.2(a) and 5.2(b), the one-, two- and three-dimensional decomposed approximations $V_{1,0}, V_{1,1}$ and $V_{1,2}$ are plotted together with the full scale Monte Carlo approximation $V_{MC}$. The exposure decreases over time, which means that the FX rate drives the expected value of this swap down. The EPE is positive by definition and increases over time.

Both the decomposed two- and three-dimensional approximations are reasonably close to the full Monte Carlo estimator, but while the results for Case A in Table 5.2 show that the maximum error is over 5% in the two-dimensional case, it is around 0.5% for the three-dimensional approximation. Note that the latter result should be exact, so that any discrepancy are finite difference errors. These may seem large, but we note that the standard error of Monte Carlo with the industry standard of $10^4$ paths would be approximately $0.1\% \sqrt{4 \cdot 10^8/10^4} = 2\%$ (see Table 5.2). While we could certainly improve the finite difference accuracy by finer meshes, we do not investigate this further at this point.\(^4\)

\(^4\)See Appendix B.7 for an analysis of the accuracy of the finite difference schemes used.
Figure 5.2: EE (left) and EPE (right) for a single ATM CCYS (5.2(a) and 5.2(b)), by three CCYS (5.2(c) and 5.2(d)), and three CCYS and an IR swap (5.2(e) and 5.2(f)). The risk factors are driven by BS2HW and the EURUSD FX rate is taken as a base.
5.6. RESULTS

<table>
<thead>
<tr>
<th></th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EE</td>
<td>EPE</td>
<td>EE</td>
</tr>
<tr>
<td>$e_L^2$ (%)</td>
<td>1D</td>
<td>2.07</td>
<td>6.11</td>
</tr>
<tr>
<td></td>
<td>2D</td>
<td>1.20</td>
<td>2.83</td>
</tr>
<tr>
<td></td>
<td>3D</td>
<td>0.038</td>
<td>0.14</td>
</tr>
<tr>
<td>$e_{Lw}$ (%)</td>
<td>1D</td>
<td>2.16</td>
<td>9.27</td>
</tr>
<tr>
<td></td>
<td>2D</td>
<td>1.50</td>
<td>5.66</td>
</tr>
<tr>
<td></td>
<td>3D</td>
<td>0.049</td>
<td>0.55</td>
</tr>
<tr>
<td>MD (bp)</td>
<td>1D</td>
<td>8.58</td>
<td>18.29</td>
</tr>
<tr>
<td></td>
<td>2D</td>
<td>4.43</td>
<td>7.35</td>
</tr>
<tr>
<td></td>
<td>3D</td>
<td>0.17</td>
<td>0.47</td>
</tr>
<tr>
<td>SE (%)</td>
<td>0.13</td>
<td>0.085</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Table 5.2: Errors of exposure approximations for Cases A, B and C. The finite difference approximations are computed with $m_1 = 60$ grid points and 500 time steps. The errors are expressed in percentages together with the standard error of the Monte Carlo benchmark with $4 \times 10^6$ paths and 1000 time steps. The SE is defined in 5.38 as the root sum squared of the standard errors over time relative to the root sum squared of the sampled EE or EPE.

5.6.2 Case B: Three CCYS (EURUSD, EURGBP and EURJPY)

In the case of three CCYS, there are $d = 7$ risk factors, such that adding three-dimensional corrections will no longer yield an exact solution. In this case, we make an a priori choice of the base risk factor, namely the EURUSD FX rate. The two-dimensional decomposed approximation for the portfolio of three FX swaps now has two extra terms to correct for the other FX rates and two extra terms to correct for the foreign interest rates:

$$V_{1,1} = V(F^1) + \left(V(F^1, F^2) - V(F^1)\right) + \left(V(F^1, F^3) - V(F^1)\right) + \left(V(F^1, R^1) - V(F^1)\right)$$

$$+ \left(V(F^1, R^1) - V(F^1)\right) + \left(V(F^1, R^2) - V(F^1)\right) + \left(V(F^1, R^3) - V(F^1)\right)$$

$$= V(F^1) + \sum_{j=2,3} \Delta V(F^1, F^j) + \Delta V(F^1, R^3) + \sum_{j=1,2,3} \Delta V(F^1, R^j),$$

where the short-hand $\Delta V(X^i, X^j) \equiv V(X^i, X^j) - V(X^i)$ was used in the last line.

For the three-dimensional corrections, there are in principle $\binom{7}{3} = 15$ extra terms, because in addition to $X_1$, we choose 2 out of the $d - 1 = 6$ remaining

---

5See Section 5.6.6 for an assessment of alternative choices.
96  

CHAPTER 5. CVA BY RISK FACTOR DECOMPOSITION

factors. However, only 8 of them are non-zero, as given here:

\[
V_{1,2} = V_{1,1} + \Delta V(F_1, F_2, F_3) + \Delta V(F_1, F_2, R_1^d) + \Delta V(F_1, F_3, R_1^d) + \sum_{j=1,2} \Delta V(F_1, F_2, R_{j}^d) + \sum_{j=1,3} \Delta V(F_1, F_3, R_{j}^d) + \Delta V(F_1, R_1^d, R_{1,1}^d),
\]

(5.41)

where the short-hand \( \Delta V(X^i, X^j, X') \equiv V(X^i, X^j, X') - V(X^i, X') - V(X^i, X') + V(X^i) \) was used.

The corrections can be interpreted in the following way. For instance in the first term, \( V(F_1, F_2, F_3) \) treats all exchange rates stochastic but with deterministic interest rates (simple Black-Scholes models); and in the last term, \( V(F_1, R_1^d, R_{1,1}^d) \) comes from the full BS2HW model for the EURUSD rate where all other processes are approximated by their expectations (see Section 5.3.1). Similar interpretations can be found for the other correction terms.

In contrast, we have for instance

\[
\Delta V(F_1, R_1^d, R_{1,1}^d) = V(F_1, R_1^d, R_{1,1}^d) - V(F_1, R_1^d) - V(F_1, R_{1,1}^d) + V(F_1) = 0,
\]

because in this approximation where (the EURJPY exchange rate) \( F_3 \) is deterministic, the (JPY) short rate \( R_{1,1}^d \) has no impact on the exposure of the swaps. Similar arguments hold for the other 6 corrections that vanish.

Figures 5.2(c) and 5.2(d) show again the one- to three-dimensional decomposed approximations of the exposures together with the full scale Monte Carlo approximation \( V_{MC} \). The different expiries for the different swaps are reflected clearly in the profiles. At \( T = 2 \), the EURJPY swap expires, which results in an upwards jump in total EE and EPE whereas at \( T = 3 \) the EURGBP swap expires and the exposures drop.

Both the two- and especially the three-dimensional approximation are close to the full Monte Carlo estimator; see especially the rows relating to Case B in Table 5.2. The mean differences are now relative to the sum of all notionals in the portfolio. As none of the correction terms accounts simultaneously for, say, the EURJPY rate, EUR and JPY short rates, even the decomposition with three-dimensional approximations is not exact in this case, such that all errors arise from a combination of truncation error for the expansion and discretization error for the PDE solution (as well as a slightly smaller Monte Carlo error for the benchmark).

We give a table of all the correction terms separately in Appendix B.7. Table B.8 there shows that while some terms are dominant, like the one involving the EURUSD and EURGBP rates among the two-dimensional corrections and the one with EURUSD and the associated short rates among the three-dimensional ones, it is certainly not the case that the other terms are negligible. This demonstrates that the problem is genuinely high-dimensional and does not have a lower
superposition dimension (see [101]), by which we mean, loosely speaking, that the solution cannot be expressed as a linear combination of solutions to low-dimensional problems. Indeed, while the approximations $V_{0,3}$ and hence $V_{1,3}$ would be exact for EE here because of its linearity, $V_{1,2}$ is not. For EPE, generally only decompositions involving all risk factors, such as $V_{0,7}$ or any $V_{r,s}$ with $r + s = 7$ will be exact. What the results show however is that approximations with much smaller $r$ and $s$ and therefore much lower computational cost can be sufficiently accurate.

In Section 5.6.5 we will show the variance reduction achieved in Case B by using the above approximation as control variate.

5.6.3 Case C: Case B with an additional IRS in EUR

By adding an IR swap to the portfolio, the risk factors do not change, as the EUR interest rate was already modeled as a factor that drives the FX rate in Section 5.6.2. In Figures 5.2(e) and 5.2(f), one observes the exposure levels are increased because of the interest rate swap. Again, the different profiles match closely over time. In Table 5.2 one sees more clearly that, for all error measures, both the EE and EPE errors improve significantly when we include three-dimensional corrections, as expected.

One thing to note is that the errors for EE are almost exactly the same as in Case B, because the IR swap exposure is valued exactly by all models that have $R$ in them, e.g., $V(F^1, R)$.  

5.6.4 Case D: Case C with an additional EURUSD FX call option

We now add an FX option on the EURUSD rate with payoff:

$$\phi(F_T^1) = \max(F_T^1 - K, 0) \Lambda_{opt}$$

at maturity $T = 4$, so that the exposure level increases first, but drops at $t = 4$, as seen in Figures 5.3(a) and 5.3(b). This option is in the money with a strike value of $K = 0.95F_{0}^1$, but the exposure impact of a single option is typically small due to its relatively lower value compared to the swaps. To make the impact more noticeable for our numerical studies, we set the notional of the option to $\Lambda_{opt} = 10$.

In the previous examples, Sections 5.6.1–5.6.3, analytical pricing formulas were available for the swaps in the portfolio, which made Monte Carlo estimation of exposures straightforward. With the FX option added, we use the regression-based MC method, similar to the popular Longstaff-Schwartz algorithm [73] for American options. An outline of the algorithm used is given in Appendix B.6.
CHAPTER 5. CVA BY RISK FACTOR DECOMPOSITION

Figure 5.3: EE (left) and EPE (right) for a portfolio of CCYS, a call option on the EUR USD FX rate and a floating vs fixed interest rate swap. The factors are driven by BS2HW and the EURUSD FX rate is taken as the base factor.

Because of this regression, all paths need to be stored and less paths can be used, which results in a higher standard error. We also note that the benchmark in this case is not perfect due to the limited span of the basis functions. The resulting bias is not taken into account in the SE given in Table 5.3. Nonetheless, the discrepancy between PDE results and Monte Carlo behaves as expected – see Table 5.3 – with a sharp improvement as more corrections are added.

<table>
<thead>
<tr>
<th></th>
<th>$e_{L_2}$ (%)</th>
<th>$e_{L_{\infty}}$ (%)</th>
<th>MD (bp)</th>
<th>SE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EE</td>
<td>23.2 1.43 0.32</td>
<td>22.6 1.56 0.36</td>
<td>33.7 2.08 0.52</td>
<td>0.86</td>
</tr>
<tr>
<td>EPE</td>
<td>26.4 4.83 1.29</td>
<td>37.9 5.44 1.87</td>
<td>319 12.5 2.62</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 5.3: Errors of exposures of a portfolio with three CCYS, a IRS and a FX call option. The errors are expressed in percentages together with the standard error of the Monte Carlo benchmark with $2 \cdot 10^5$ paths. The SE is defined in 5.38 as the root sum squared of the standard errors over time relative to the root sum squared of the sampled EE or EPE.

5.6.5 Variance reduction

We now correct the bias in the results in the preceding sections by applying the methodology of Section 5.3.2, at much reduced variance compared to the standard estimator. For simplicity, we restrict ourselves to the case where derivatives are priced analytically, so that no regression is required. The obtained variance reduction, calculated as the ratio of the two variances with and without control
5.6. RESULTS

variate, is shown in Figure 5.4 for the different decomposed approximations and test cases from Sections 5.6.1 to 5.6.3. To be clear, a value of, e.g., 0.5% means that the variance is reduced by a factor of 200, while a value of 100% implies that no variance reduction at all was achieved.

In Figures 5.4(a) and 5.4(b), the variance reduction factor for EE and EPE in the case of a single CCYS is in the range of 100-1000. For the two-dimensional control variate, the reduction is greater for shorter maturities and increases with time. For the three-dimensional control variate, the variance reduction is constant, which is due to the exactness of the control variate. The only discrepancy between the control variate and the Monte Carlo estimator is the (time) discretization error, which also explains why the variance is not equal to zero.

Figures 5.4(c) and 5.4(d) show the variance reduction in the case of three CCYS. Here, the three-dimensional decomposed approximation is no longer exact and thus, also in the three-dimensional case, the variance increases over time. The sharp drop at $T = 2$ and $T = 3$ is due to the expiry of the EURJPY and EURGBP CCYSs. The variance reduction is largest between time 3 and 5 because during that time period only the EURUSD CCYS is not terminated. When we use two-dimensional corrections, the reduction is around 200 for EE and 50 for EPE. When three-dimensional corrections are taken into account, we obtain a reduction by a factor of $10^5$ for EE and 200 for EPE.\footnote{The extreme variance reduction for EE and 3D is more clearly seen quantitatively from Figure 5.5 later on, where the data are plotted on a log scale.}

Figures 5.4(e) and 5.4(f) show the variance reduction in the case of three CCYS and an IRS. The EE results resemble those for three CCYS alone, because the stochasticity of the IRS due to $R^d$ is modelled exactly by the corrections, as explained at the end of Section 5.6.3. For EPE, the variance is reduced by a factor around 30 for two-dimensional corrections and a factor of 50 for three-dimensional corrections.

The set-up is chosen so that the finite difference accuracy is comparable to the Monte Carlo accuracy. As the finite difference approximation converges at a higher order (for low dimensions, typically up to about 3 or 4) than the Monte Carlo sampling, the computational effort of the Monte Carlo component will become dominant if higher accuracy is required.

As a side note, we remark here that the different PDEs in the decomposition can be solved fully in parallel, and parallel to the Monte Carlo runs. We report here sequential run times.

From these run times for Case B, presented in Table 5.4, we can see that, using the 2D approximation as a control variate, we gain a computational speed up by a factor 50, as the computation time for the 2D corrections by finite differences is negligible. Using 3D corrections gives us a speed up factor 12; in spite of the greater variance reduction, the increased time for the PDE solutions eats up any
benefit over the 2D corrections.

<table>
<thead>
<tr>
<th>Solver</th>
<th>time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D FD</td>
<td>1.53 (3.65)</td>
</tr>
<tr>
<td>2D FD</td>
<td>2.35 (10.67)</td>
</tr>
<tr>
<td>3D FD</td>
<td>344 (601.34)</td>
</tr>
<tr>
<td>Full MC</td>
<td>4251 (7738)</td>
</tr>
<tr>
<td>2D CV MC</td>
<td>81.5 (258.4)</td>
</tr>
<tr>
<td>3D CV MC</td>
<td>30.1 (107.6)</td>
</tr>
</tbody>
</table>

Table 5.4: Computational times of the individual solvers in seconds. For the FD solvers, we use \( m_1 = 60 \) grid points and 100 or 500 (in brackets) time steps (see also Appendix B.5). The full Monte Carlo result is obtained with \( 4 \cdot 10^6 \) paths and 100 or 1000 (in brackets) time steps. For the control variate results, we reduce the number of paths commensurate with the variance reduction. From Figure 5.4, Case B, the variance reduction factor is seen to be around 50 in the case of 2D and 200 in the case of 3D corrections. Therefore, we choose \( 8 \cdot 10^4 \) paths in the case of 2D corrections and \( 2 \cdot 10^4 \) in the case of 3D corrections.

5.6.6 Other base risk factors

In the previous sections, the EURUSD FX rate was chosen as the base risk factor \textit{a priori}, based on the fact that the derivatives driven by this FX rate have the highest maturity. However, there is no guarantee that this choice as a base is optimal in any sense. Often, one may have prior knowledge what the main driving factors are. Failing that, it would be practically feasible to estimate the variance reduction achieved by different factors by a relatively small number of samples in a trial run, and then do the actual large scale estimation with the best performing base factor. In Figures 5.5(a) and 5.5(b), the different EE and EPE profiles for Case B are shown when we choose the EURGBP or EURJPY FX rate as a base compared to choosing the EURUSD rate. Table 5.5 shows that the error is smallest when we choose the EURUSD rate.

In Figures 5.5(c) and 5.5(d), we show the resulting variance reductions of the control variates. Clearly, the EURJPY FX rate as a base does not reduce the variance after \( T = 2 \), which makes sense as that is when the base risk factor does not affect any non-terminated derivative in the portfolio. The EURUSD FX rate performs best especially after time \( T > 3 \) when the EURUSD rate affects the only non-terminated derivative in the portfolio. Note that we use a log scale on the y-axis for visualization.
Figure 5.4: Variance reduction for different test cases with 2D and 3D corrections. The variance with and without control variate is compared. For this computation we simulate $10^6$ paths.
CHAPTER 5. CVA BY RISK FACTOR DECOMPOSITION

Figure 5.5: Case B. EE and EPE results for different risk factors as a base for the decomposed approximation. The profiles are shown in Figures 5.5(a) and 5.5(a). The variance reductions by using two-dimensional approximations are shown in Figures 5.5(c) and 5.5(d), and for three-dimensional corrections in Figures 5.5(e) and 5.5(f).
5.7 Conclusion

This chapter is motivated by the computation of exposure profiles for portfolios depending on a moderate to large number of risk factors. For this problem, the industry standard technique is to employ forward Monte Carlo sampling to compute future scenarios. This scales linearly in both the number of dimensions (risk factors) and products. However, valuing the whole book across all scenarios is still a big computational challenge and relatively large standard errors have to be tolerated given the relatively low feasible number of sample paths.

We therefore propose another approach which exploits the accuracy of PDE approximation schemes for low-dimensional estimation problems, through an anchored-ANOVA-style splitting of the high-dimensional problem into a sequence of lower-dimensional ones. As the problem is stated in the form of nested conditional expectations, we use a combination of forward and backward PDEs to generate exposure profiles for all future times with a single backward and forward sweep.

This chapter provides a proof of concept rather than a fully worked out black box algorithm. The detailed analysis of a moderately sized, realistic and fully calibrated test case showed the scale of the benefits achievable by risk factor decomposition coupled with numerical PDE solutions, used as standalone approximation or – if necessary – employed as a control variate for a Monte Carlo estimator. Some of the computational savings were dramatic, with a speed-up factor of 50 compared to standard Monte Carlo estimation, by using a sum of two-dimensional estimators as control variate.

Further tests, not reported here, indicate that for longer maturities (10 years) some three-dimensional terms in the decomposition become significant for an accurate standalone approximation as well as for effective variance reduction.

Table 5.5: Errors of exposures for Case B for different choices of the base risk factor. The errors are expressed in percentages together with the standard error of the Monte Carlo benchmark with $10^6$ paths. The SE is defined in 5.38 as the root sum squared of the standard errors over time relative to the root sum squared of the sampled EE or EPE.
CHAPTER 5. CVA BY RISK FACTOR DECOMPOSITION

It remains to investigate in practice the scalability with respect to the number of risk factors and number of products in the derivative portfolio. Both Monte Carlo sampling and the PDE based method scale linearly in the latter. As for the number of risk factors, the number of 2D correction terms is linear, and the number of 3D ones quadratic in the dimension.

For larger portfolios, one can consider using approximations of the form $V_{0,2}$ or $V_{0,3}$ instead of $V_{1,2}$ or indeed a combination of the two in the sense that only a subset of $V_u$ with $|u| = 2$ and $|u| = 3$ are computed and included in the approximation, chosen adaptively by estimation of the individual terms. We find evidence for this in Table B.8, which shows that only a subset of the correction terms are significant. This also suggests that not all terms have to be computed with the same relative accuracy, and one can follow the principle of Griebel and Holtz [49] to divide the total computational budget optimally between all correction terms. In this process, it is not necessary to use the same numerical method for all terms. Indeed, we have used a combination of closed-form and numerical solutions in the tests, and the framework is rich enough to use the best available method (e.g., Fourier-based methods for affine models, PDEs for early exercise options, or conceivably even a Monte Carlo method for strongly path-dependent derivatives) for a given sub-problem. Moreover, for the computation of a specific correction terms only a subset of derivatives has to be considered, namely those affected by the risk factors considered in that correction. So it is conceivable that even for a large derivative portfolio each correction term only requires consideration of a small fraction of the derivatives. For instance, if the portfolio contains derivatives on equities, commodities and FX, the risk factor decomposition may provide a way to decompose the exposure computations into smaller sub-computations.

The practical challenge will be to develop a framework which allows this to happen in a generic way, and which is easily adaptable as different derivatives are entered or the modelling framework changes.

In this, and the previous chapters, the focus has been on the methodology of computing exposure with the help of PDE methods. The results confirm that using PDEs in the context of exposure calculations is accurate and therefore promising. This is why we will now apply the FDMC method to address a complex and urgent question in derivative pricing: what is the impact of the volatility smile and stochastic rates on default risk? We will investigate this by looking at exposures (and more specifically CVA) computed with and without stochastic volatility and rates in real market scenarios in the next chapter.
The methodologies developed in earlier chapters prove that we can use PDEs in the context of realistic exposure calculations. We will therefore now apply these methods to answer an urgent open question in derivative pricing. For derivatives that depend on foreign exchange and interest rates, it is important to incorporate stochastic volatility and interest rates in the models. Although, it is known that these risk factors are important for pricing, in default risk the impact of including these factors as SDEs is unclear. Therefore, in this chapter, we investigate the impact of stochastic volatility and interest rates on Counterparty Credit Risk (CCR) for FX derivatives. For this, we will employ the FDMC method developed in earlier chapters. We analyze two real life cases in which the market conditions are different, namely during the 2008 credit crisis where risks are high and a period after the crisis in 2014, where volatility levels are low. The Heston model is extended by adding two Hull-White components which are calibrated to fit the EURUSD volatility surfaces. We then present future exposure profiles.
and CVAs (Credit Value Adjustments) for plain vanilla Cross-Currency Swaps (CCYS), Barrier and American options and compare the different results when Heston-Hull-White-Hull-White or Black-Scholes dynamics are assumed. It is observed that the stochastic volatility has a significant impact on all the derivatives. For CCYS, some of the impact can be reduced by allowing for time dependent variance. We further confirmed that Barrier options exposure and CVA is highly sensitive to volatility dynamics and that American options’ risk dynamics are significantly affected by the uncertainty in the interest rates.

6.1 Introduction

In contrast to the valuation of single derivatives, exposure calculations are typically done on the level of counterparties and therefore require the simultaneous handling of multiple risk factors with complex interdependencies. Although the Black-Scholes model gives very valuable insight into CCR, it assumes the volatility and the interest rates of the underlying risk factors to be constant. This is unrealistic and its impact in particular for long dated and path-dependent derivatives is unknown. It is more realistic to assume that the underlying driving risk factor (e.g. FX rate for cross-currency derivatives), its volatility and the interest rates are driven by stochastic processes. Here we study rigorously and for the first time, the impact of these extra degrees of freedom on exposures for various popular FX derivatives using the computationally challenging Heston-2-Hull-White (H2HW) model fully calibrated to real market data for a wide range of model and contract parameters.

In academia an important contribution in pricing FX derivatives assuming high dimensional models is done in [52] and [51], where the authors show how these high dimensional models can be calibrated. In [87], [34] and [26], it is observed that stochastic volatility models that are perfectly calibrated to European options can produce different prices for path-dependent barrier options. In [1], the importance of the interest rate dynamics for the valuation of FX options is shown. However, the choice of interest rate parameters in their work was made for illustrative purposes only. Moreover, in all these works the focus is on the possible impact on prices, whereas here we extend the analysis by looking at the impact on future exposures.

These future exposures are used to determine the valuation adjustments like CVA, DVA [47] and more recently KVA [46]. It requires simulating the future states of the risk factors and then valuing the derivative in all these future states. Recent contributions by academics and practitioners with guidelines on how to achieve this include [78], [79] and [81]. To the best of our knowledge, these
6.2. PROBLEM FORMULATION

6.2.1 Models

The main model used in this chapter is the H2HW model, where Heston is extended with two Hull-White (HW) processes to incorporate the stochastic dynamics of the interest rates. To assess the impacts, this model serves as a benchmark. Assuming these dynamics, one can price cross-currency and fixed-income products. As in the Heston case, the SDE for the variance process enables us to capture the volatility smile. The four-dimensional dynamics of the H2HW model are given by:

\[
\begin{align*}
\text{d}X_t &= (r^d_t - r^f_t)X_t \text{d}t + \sqrt{V_t}X_t \text{d}W^1_t, \\
\text{d}V_t &= \kappa (\eta - V_t) \text{d}t + \gamma \sqrt{V_t} \text{d}W^2_t, \\
\text{d}R^d_t &= \lambda_d (\theta_d (t) - R^d_t) \text{d}t + \eta_d \text{d}W^3_t, \\
\text{d}R^f_t &= \left[ \lambda_f (\theta_f (t) - R^f_t) - \eta_f \rho_{1,4} \sqrt{V_t} \right] \text{d}t + \eta_f \text{d}W^4_t, \\
\text{d}W^i_t \text{d}W^j_t &= \rho_{i,j} \text{d}t, \ \text{for } i \neq j \in [1, 2, 3, 4].
\end{align*}
\]
where the variance process (6.1b) is a CIR process with \( \kappa \) as the mean-reverting speed, \( \eta_v \) as the level of the long term mean, \( \gamma \) the so-called volatility-of-volatility.

The short rate (6.1c) ( or (6.1d) ) are both modeled by a HW process with \( \lambda_f \) (\( \lambda_f(t) \)) as the mean-reverting speed, \( \theta_d(t) \) (\( \theta_f(t) \)) as the yield level deduced from the forward curve and \( \eta_d \) (\( \eta_f \)) the short rate volatility of the domestic (foreign) interest rate process. Note that the drift adjustment for \( dR^d_t \) comes from the change of measure from the foreign to the domestic risk-neutral measure ([50])b.

This model allows to price hybrid products whose payoff \( g((X_t,V_t,R^d_t,R^f_T)_{0 \leq T \leq T}) \) depends on different stochastic components of the model. Pricing a European option means finding \( u(\cdot) \) such that

\[
    u(x,v,r^d,r^f,\tau) = \mathbb{E} \left[ e^{-\int_0^\tau r^d_s \, ds} g(X_T,V_T,r^d_T,r^f_T) \mid X_t = x, V_t = v, R^d_t = r^d, R^f_t = r^f \right].
\]

Via Feynman-Kac, one can show that the time-reversed pricing mechanism \( u \) of a contingent claim \( g(X_T,V_T,R^d_T,R^f_T) \) has to satisfy the following PDE

\[
\begin{align*}
    \frac{\partial u}{\partial \tau} &= \frac{1}{2} \left( \kappa^2 + \frac{\eta^2}{\theta^2} \right) \frac{\partial^2 u}{\partial x^2} + \frac{1}{2} \left( \gamma^2 + \frac{\eta^2 }{\theta^2} \right) \frac{\partial^2 u}{\partial v^2} + \frac{1}{2} \frac{\eta^2}{\theta^2} \frac{\partial^2 u}{\partial (r^d)^2} + \frac{1}{2} \frac{\eta^2}{\theta^2} \frac{\partial^2 u}{\partial (r^f)^2} \\
    &+ (r^d - r^f) \frac{\partial u}{\partial x} + \kappa (\bar{u} - v) \frac{\partial u}{\partial v} + \lambda_d (\theta_d (T - \tau) - r^d) \frac{\partial u}{\partial r^d} \\
    &+ \lambda_f (\theta_f (T - \tau) - r^f - \rho_{1,4} \eta_f \sqrt{\bar{u}}) \frac{\partial u}{\partial r^f} \\
    &+ \rho_{1,2} \gamma \sqrt{\bar{u}} \frac{\partial^2 u}{\partial (x \partial v)} + \rho_{1,3} \eta_d x \sqrt{\bar{u}} \frac{\partial^2 u}{\partial x \partial r^d} + \rho_{1,4} \eta_f x \sqrt{\bar{u}} \frac{\partial^2 u}{\partial x \partial r^f} \\
    &+ \rho_{2,3} \gamma \sqrt{\bar{u}} \frac{\partial^2 u}{\partial (v \partial r^d)} + \rho_{2,4} \gamma \sqrt{\bar{u}} \frac{\partial^2 u}{\partial v \partial r^d} + \rho_{3,4} \eta_f \eta_d \frac{\partial^2 u}{\partial r^d \partial r^f} \\
    &- \rho_{3,4} \eta^2 \eta_d \frac{\partial^2 u}{\partial r^d \partial r^f} - \rho^d u, \\
    u(x,v,r^d,r^f,\tau = 0) &= g(x,v,r^d,r^f),
\end{align*}
\]

(6.2)

(6.3)

where \( \tau := T - t \).

In order to assess the impact of stochastic volatility we compare against the Black-Scholes-2-Hull-White (BS2HW) model which assumes a deterministic volatility. In this case, we consider the following two possibilities:

1. Using the ATM implied volatility at maturity of the derivative;
2. Using a piecewise constant volatility function which is aligned with the ATM implied volatility curve over all maturities.

In Appendix B.1 we show the effective volatilities for the different models.

On the other hand, the impact of stochastic interest rates is assessed by looking into the Heston model with constant interest rates \( r^d \) and \( r^f \).
6.2. PROBLEM FORMULATION

6.2.2 Derivatives

We consider Cross-currency Swaps (CCYS), Up-and-out Barrier Call (UOC) and American Put/Call options. The former is the simplest derivative connecting FX and fixed-income markets and is highly relevant from the perspective of exposure and outstanding trading volumes. In addition, UOC and American options are interesting because of their path-dependence and early-exercise possibilities. More technical details about the pricing of these derivatives can be found in [92] and [43].

Cross-currency swap

A cross-currency swap is a derivative where the buyer pays interest and a notional in one currency and receives in another. Through this work we assume a swap where the buyer receives interest in USD and pays it in EUR. We present results for fixed notional cross-currency basis swaps where both legs are floating and notional are exchanged at inception and maturity. At the money (ATM) refers to the case where all cashflows (notional exchanges and future payments) are based on a domestic and foreign principal amount that is implied by the FX \( X_t \) spot rate. For the in the money (ITM) and out of the money (OTM) cases, the FX rates 1.25 \( X_t \) and 0.75 \( X_t \) are used.

CCYS differ from the other derivatives considered here in the sense that it is possible to compute its prices analytically without assuming an SDE to drive the FX rate. For exposure however, future scenarios are needed and for that the FX rate needs to be driven by a model.

Up-and-out call option

Next, we look into an up-and-out barrier call option. This first-order exotic option terminates (value becomes 0) when the underlying FX rate exceeds a predefined barrier level \( B \), but otherwise delivers a payoff equal to that of a European call. The price of an up-and-out barrier call option with barrier \( B \), strike \( K \) and maturity \( T \) equals:

\[
u(x, v, r^d, r^f, t) = \mathbb{E} \left[ e^{-\int_t^T r^d dt} (X_T - K)^{+} \mathbb{1}_{\{r(B, X) > T\}} \right | X_t = x, V_t = v, R^d_t = r^d, R^f_t = r^f \],
\]

where \( r(B, X) := \inf\{ t \geq 0 : X_t \geq B \} \). Note that future exposure as well as prices are now path-dependent because of the knockout probability.

American options

The holder of an American option is able to receive the payoff upon request any time before maturity. In practice, these options are approximated as Bermudan
options with many time points at which early exercise is allowed. Clearly by increasing the number of time points the option will converge to the American option.

For pricing, time is discretized by taking $N_T$ number of uniformly distributed time points between 0 and $T$ with $\Delta t := \frac{T}{N_T}$. At each exercise time, the exercise value (given by the payoff function) and the continuation value of the option are compared in a backward induction scheme. The payoff function and the continuation value for the option at discrete time $t_m$ are defined as:

$$\phi(X_t) = \max(y(X_t - K), 0) \quad \text{with} \quad y = \begin{cases} 1, & \text{for a call} \\ -1, & \text{for a put} \end{cases}$$

$$c(X_m, v_m, r^d_{m+1}, r^f_{m+1}, t_m) = E \left[ e^{-r^d_{m+1} \Delta t} u(X_{m+1}, v_{m+1}, r^d_{m+1}, r^f_{m+1}, t_{m+1}) | (X_m, v_m, r^d_m, r^f_m) \right],$$

where $u(X_{m+1}, v_{m+1}, r^d_{m+1}, r^f_{m+1}, t_{m+1})$ is the option value at time $t_{m+1}$.

It is assumed that the holder of the option exercises the option when the payoff is higher than or equal to the continuation value, after which the contract terminates. At maturity $T = t_{N_T}$, the option value is equal to the payoff value.

The recursive scheme to price a Bermudan option in equation (3.5) is slightly adjusted in this higher dimensional case:

$$u(X_m, v_m, r^d_m, r^f_m, t_m) = \begin{cases} \phi(X_{N_T}) & \text{for } m = N_T; \\ \max \left\{ c(X_m, v_m, r^d_m, r^f_m, t_m), \phi(X_m) \right\} & \text{for } m = 1, \ldots, N_T - 1; \\ c(X_0, v_0, r^d_0, r^f_0, t_0) & \text{for } m = 0. \end{cases}$$

In comparison to Barrier options the exercise of an American option is even more dependent on the chosen model. For Barrier options, the exercise level is equal to the barrier which is static, for Americans options however, this exercise level is determined by the continuation value which depends on the chosen model and the future state variables. For example, in the simple Black-Scholes world with no dividends/foreign exchange rate and positive domestic risk-free rate it is not beneficial to exercise an American Call option. However, this is not the case when both rates are time-dependent. As we describe later in more detail, American options seem to be very sensitive to the uncertainty in fixed-income markets.

### 6.2.3 Algorithm

For the computation of various value adjustment, one can create a new valuation framework as is done in for example [21] and [17], but here we choose another approach. We compute the exposures from the distribution of the future value of
6.2. PROBLEM FORMULATION

a derivative. The valuation of the CCYS is performed by combining Monte-Carlo (MC) simulations with analytical pricing. As there is no general analytic price for American or Barrier options under the H2HW dynamics, we apply the the FDMC method as described in Chapter 3 and 4. The slightly adjusted pseudo-code for the algorithm is as follows:

1. Select the number of time points $N_T$;
2. Construct a joint meshgrid for every modeled underlying risk factor;
3. On this grid, solve the PDE backwards in time using the FD method, and store these price grids for every time point;
4. Select the number of paths $N_p$ and set $t = 0$;
5. Use the MC method to simulate $N_p$ realizations of the underlying risk factors at $t + \Delta t$, where $\Delta t = \frac{T}{N_T}$;
6. For each of the simulated states, compute the respective derivative price by interpolation on the price grid;
7. In case early exercise/termination is allowed, set the derivative price to the exercise/termination value for paths which have been exercised/terminated at time $t$;
8. Calculate the exposure metrics using the resulting distribution of the derivative prices;
9. If needed discount the exposure metrics;
10. If $t < T$ take $t = t + \Delta t$ and go to step 5, else quit.

By definition the exposure of an early exercise option equals zero once the option is exercised; otherwise, the exposure is equal to the value of the option.

For the scenario generation we use the Euler discretization for the FX and interest rates. As the Euler scheme is proved to be inefficient for the simulation of the volatility dimension, the QE scheme as described in [5] is used. For the scenario generation we used $10^7$ paths and 100 time points such that in all test cases, the relative standard error of the mean was smaller than 1%.

To solve the pricing PDE (6.2.1) on a grid in every dimension, we expanded the FD implementation from Chapters 3 and 4. Note that pricing claims on underlyings driven by the H2HW model requires solving a 4 dimensional PDE which, in the case of exotic options, will result in large computational requirements\(^1\). The number of grid points used for the pricing grids equals 80, 40, 20

\(^1\)In case of European options the PDE can be reduced to a two dimensional PDE by switching to the forward measure.
and 20 in \( s \)-, \( v \)-, \( r^d \)- and \( r^t \)-dimension respectively and 100 points in time. Using this number of points we have compared European call option prices obtained by the finite difference method with a Monte Carlo simulation with \( 10^7 \) paths and 100 time points. For both periods we found that the relative difference was smaller than 0.5%. Furthermore, the error analysis in [92] shows that this number of points is sufficient to accurately (relative error smaller than 0.5%) price up-and-out Barrier call and American options. Because of the large computational demands, the discretized equation is again solved via the Alternating Direction Implicit (ADI) method which is numerically stable and computationally attractive as it only inverts sparse matrices. Because of its second order convergence we choose the Hundsdorfer-Verwer scheme (see e.g. [28], [61] and [63]).

6.3 Market data and model calibration

As noted earlier, the assessment of the additional degrees of freedom and their impact is performed by simulating exposures for derivatives on the EURUSD FX rate. We start by calibrating the models to the prices observed in the market during two different periods. The calibrated parameters are then used in computing the implied densities and exposure metrics. In the following section the differences between the models are analysed by comparing these metrics.

6.3.1 Data

Our data set contains market data from two intrinsically different periods. We analyse the impact of the different models in these two distinct market scenarios in order to get a robust view of the model implications. The first dataset presented in Figure 6.1(a) contains implied EURUSD volatilities as observed in 2008 on the 29th of September which represents the market conditions during the 2008 credit crisis. The second dataset is extracted from the market as seen on December 2nd 2014 and is presented in Figure 6.1(b). Comparison the two datasets reveal that:

1. In 2008, during the crisis, the volatility was relatively high for short maturities, but steadily decreasing with time-to-maturity for higher maturities;
2. The reverse is true in the case of 2014 December 2nd - volatilities are at historically low levels, but are increasing for larger maturities;
3. In both cases the volatility smile ‘flattens’ as maturity increases;

Note that in the EURUSD case USD is the domestic currency and EUR is the foreign.
4. Qualitatively speaking, the skew is more profound in 2014. In addition the skew is also more visible in the implied smiles of the shorter maturity vanilla options.

The Hull-White components used in the H2HW model allow to fit the term structure of the market interest rates. The fitted yield curves of the periods are presented in Figure 6.2. The comparison reveals the following:

1. Interest rate levels are moderate in 2008 and quite low in 2014;
2. The rate differences are large for shorter tenors, but decrease with time in 2008 and vice versa in 2014;
3. In 2008 domestic (USD) rate is lower than foreign (EUR) and the other way around in 2014.

6.3.2 Calibration

The calibration procedure used in this paper is described in more detail in [92]. For each strike \( K \) and maturity \( T \) implied volatilities, combined with yield curve data are converted into call option prices \( C_{mkt}(K,T) \). The goal of the calibration procedure is then to find parameter vector \( \vec{\theta} \) such that the following is minimized:

\[
\sum_{K,T} \left( C_{mkt}(K,T) - C_{model}(\vec{\theta})(K,T) \right)^2.
\]  

(6.5)

Note that for the H2HW model, 14 parameters have to be calibrated which is a demanding task for the commodity hardware. Therefore, the parameter estimation is performed in stages:

1. The Hull-White parameters are calibrated to fit the prevailing yield curve and swaption data using analytic expressions. We used co-terminal swaptions that terminate in 10 years (1x9, 2x8, 3x7, 4x6, 5x5, 6x4, 7x3, 8x2 and 9x1), as these swaptions can also be used to replicate the CVA of a swap that matures in 10 years, as is shown in [94].

2. The Heston model is fitted to the FX implied volatility surface using the COS method as presented in [40] using constraint optimization. In the fitting procedure market volatilities are converted to prices which are then used to calibrate the parameters.

3. The correlations between the observable factors (exchange and interest rates) are estimated using historical time-series data⁴;

⁴These correlations are assumed to be equal in 2008 and 2014.
CHAPTER 6. SMILE AND DEFAULT

Figure 6.1: EURUSD implied volatilities as a function of moneyness, which is defined as $K/X_0$, for different expiries and different periods.

Figure 6.2: Domestic (USD) and foreign (EUR) yield curves.
6.3. MARKET DATA AND MODEL CALIBRATION

4. H2HW parameters are filled in with values from the previous stages. The remaining correlations $\rho_{c,v}, \rho_{c,r}$ then are fitted using grid search algorithm\(^5\) over the feasible values\(^6\).

The resulting implied volatilities are shown in table 6.1. The discrepancies are highest close to the strike where the convexity tends to be high. This is especially visible for the cases with short maturities where the smile is most profound. An increase in error level can be observed as maturities increase which could be explained by the non-uniformity of options in time. The data contains more options of shorter maturities hence the goal function penalizes errors in shorter maturities more than those in the long term. The mean relative errors expressed as percentages of the market implied volatility averaged at 4.4% in 2008 and 3.0% in 2014 and are mostly dominated by far out of money strikes. We have also done this calibration routine by minimizing implied volatilities instead of prices and the results were similar. An improvement of this calibration could be achieved by weighting the errors by the corresponding option liquidity if such data were available.

Table 6.1: Absolute calibration errors measured as $|\sigma_{\text{market}} - \sigma_{\text{Heston}}|$.

(a) 2008

<table>
<thead>
<tr>
<th></th>
<th>10% Δ Call</th>
<th>25% Δ Call</th>
<th>ATM</th>
<th>25% Δ Put</th>
<th>10% Δ Put</th>
</tr>
</thead>
<tbody>
<tr>
<td>6M</td>
<td>0.35%</td>
<td>0.50%</td>
<td>0.78%</td>
<td>0.33%</td>
<td>0.68%</td>
</tr>
<tr>
<td>1Y</td>
<td>0.26%</td>
<td>0.57%</td>
<td>0.79%</td>
<td>0.32%</td>
<td>0.72%</td>
</tr>
<tr>
<td>3Y</td>
<td>0.46%</td>
<td>0.14%</td>
<td>0.12%</td>
<td>0.28%</td>
<td>0.67%</td>
</tr>
<tr>
<td>10Y</td>
<td>0.42%</td>
<td>0.23%</td>
<td>0.15%</td>
<td>0.13%</td>
<td>0.45%</td>
</tr>
</tbody>
</table>

(b) 2014

<table>
<thead>
<tr>
<th></th>
<th>10% Δ Call</th>
<th>25% Δ Call</th>
<th>ATM</th>
<th>25% Δ Put</th>
<th>10% Δ Put</th>
</tr>
</thead>
<tbody>
<tr>
<td>6M</td>
<td>1.20%</td>
<td>0.70%</td>
<td>0.11%</td>
<td>0.11%</td>
<td>0.14%</td>
</tr>
<tr>
<td>1Y</td>
<td>1.00%</td>
<td>0.53%</td>
<td>0.03%</td>
<td>0.18%</td>
<td>0.19%</td>
</tr>
<tr>
<td>3Y</td>
<td>0.01%</td>
<td>0.07%</td>
<td>0.17%</td>
<td>0.26%</td>
<td>0.31%</td>
</tr>
<tr>
<td>10Y</td>
<td>0.26%</td>
<td>0.04%</td>
<td>0.05%</td>
<td>0.13%</td>
<td>0.04%</td>
</tr>
</tbody>
</table>

The calibrated parameters can be found in Appendix B.2.

\(^5\)In our experiments option values were rather insensitive to the changes in these correlations, so the grid was much coarser than in previous stages.

\(^6\)Note that the correlations have to be picked such that the covariance matrix remains positive semi-definite.
CHAPTER 6. SMILE AND DEFAULT

6.4 Case studies

The assessment of the stochastic interest rate is made by comparing the results of H2HW to Heston and the stochastic volatility by contrasting H2HW to BS2HW. For the CCYS, we first look at the CVA as a percentage of the notional \( P \), which is computed as a discrete sum over time (with \( T_K \) the number of equidistant time points):

\[
CVA^M = \frac{100}{P} \sum_{i=1}^{T_K} \text{EPE}^M(t_i)PD(t_{i-1},t_i),
\]

(6.6)

where \( M \) can be the full H2HW, Heston or BS2HW model with constant or piecewise constant volatility and \( PD(t_{i-1},t_i) \) is computed with the help of survival probabilities as explained in Section 4.3.3.

To further quantify the differences between the models we look at the Mean Difference (MD) of exposures over time relative to the notional \( P \) expressed in basis points (6.7). This way we can assess model differences as the swap spread difference which is independent of the counterparties creditworthiness. As a reference we take the exposure computed by the full H2HW model and compare the exposures modeled under Heston or BS2HW with constant and piecewise constant volatility:

\[
\text{MD}_{\text{CCYS}} = \frac{10000}{P \cdot T_K} \sum_{i=1}^{T_K} \left| F^{\text{H2HW}}(t_i) - F^M(t_i) \right|,
\]

(6.7)

where \( F^{\text{H2HW}} \) is the exposure metric (EE, EPE, ENE or quantile) modeled under H2HW and \( F^M \) is the exposure metric modeled under \( M \) where again \( M \) can be the Heston or BS2HW model with constant or piecewise constant volatility. For the barrier and American options we look at the relative differences of CVA where we take the H2HW value as a reference:

\[
D_{\text{Option}} = \frac{100}{T_K} \sum_{i=1}^{T_K} \left| \frac{\text{CVA}^{\text{H2HW}}(t_i) - \text{CVA}^M(t_i)}{\text{CVA}^{\text{H2HW}}(t_i)} \right|.
\]

(6.8)

6.4.1 Cross-currency swaps

The first derivative to be analysed is the domestic receiver CCYS. Since exposure of swaps heavily depends on interest rate risk, we disregard the Heston model with constant interest rates.
6.4. CASE STUDIES

Densities

In figure 6.3 we show the densities for the future FX rates for two maturities and the two time periods. We see that the Heston model implies fatter left tails. Another observation is that adding stochastic volatility results in a shifted peak. Furthermore, the uncertainty in the interest rates makes the peak milder and produces more likely tail events, in particular for longer maturities.

Exposures and CVA

Figure 6.4 presents the three exposure metrics for the two periods and for different moneyness levels. A first observation is that the difference between the one-sided exposures (EPE and ENE) modeled by H2HW or BS2HW with piecewise constant volatility is small for an ATM traded swap. The differences are larger for other moneyness levels. When we compare 2008 with 2014, we see that the exposures modeled in 2008 are more positive than in 2014. This is in line with the future yields as presented in figures 6.2.

The CVA of the CCYS is presented in table 6.2, for different CDS spreads, ranging from 0 to 400 bps. We can see that both in 2008 and in 2014, the Black Scholes model with ATM volatility is significantly different than the H2HW model. The Black Scholes model with piecewise volatility is closer to the H2HW model, but in the ITM and OTM case, the difference is more profound which can be explained by the larger skew effect.

The mean differences for exposures (EE, EPE and ENE) are shown in table 6.3. For EE the differences for both Black-Scholes models are in the range of 10 bps upfront. However for the one-sided exposures (EPE and ENE), the differences are in the range of 50 bps for the piecewise constant volatility model and even in the range of 150 bps for the constant volatility model. These results suggest that the skew impact is significant, but is generally lower when the term structure of volatility is fully taken into account.

Quantiles

Figures 6.5(a) and 6.5(b) present the evolution of the quantiles for the 10-year ATM CCYS under H2HW and BS2HW with constant and piecewise constant volatility. Clearly, the differences in quantiles are more pronounced than the differences in exposures. The quantiles modeled by a piecewise constant volatility Black-Scholes model are close to the H2HW quantiles at the start, but diverge towards maturity. This is similar for the ITM and OTM cases as shown in figures 6.5(c) to 6.5(f).

In table 6.4 we show the MDs of the quantiles in basis points. The differences are in the range of 500 bps for the piecewise constant volatility model and even
Figure 6.3: Densities of the FX rate generated by calibrating H2HW, Heston, and BS2HW with constant and piecewise constant volatility to 2008 (left) or 2014 (right) data. Figures 6.3(a) and 6.3(b) show the unconditional distribution after $T = 3$ years, whereas figures 6.3(c) and 6.3(d) show the distributions after $T = 10$ years.
6.4. CASE STUDIES

Figure 6.4: Exposures metrics under H2HW and BS2HW for a 10-year floating vs floating CCYS for three different moneyness levels and two different periods. At the money (ATM) refers to the case where all cashflows (notional exchanges and future payments) are based on a domestic and foreign principal amount that is implied by the FX ($X_t$) spot rate. For the in the money (ITM) and out of the money (OTM) cases, the FX rates $1.25X_t$ and $0.75X_t$ are used.
Table 6.2: CVA for a CCYS for the different models and moneyness levels. The CDS spreads range from 100 to 400 bps. The recovery rate is assumed to be constant at 0% and the CVA is quoted as a percentage of the notional.

(a) 2008

<table>
<thead>
<tr>
<th>Bps</th>
<th>ITM</th>
<th>ATM</th>
<th>OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H2HW</td>
<td>BS2HW ATM vol</td>
<td>BS2HW piecewise vol</td>
</tr>
<tr>
<td>100</td>
<td>1.88</td>
<td>1.81</td>
<td>1.87</td>
</tr>
<tr>
<td>200</td>
<td>3.58</td>
<td>3.46</td>
<td>3.57</td>
</tr>
<tr>
<td>300</td>
<td>5.13</td>
<td>4.96</td>
<td>5.12</td>
</tr>
<tr>
<td>400</td>
<td>6.54</td>
<td>6.32</td>
<td>6.53</td>
</tr>
</tbody>
</table>

(b) 2014

<table>
<thead>
<tr>
<th>Bps</th>
<th>ITM</th>
<th>ATM</th>
<th>OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H2HW</td>
<td>BS2HW ATM vol</td>
<td>BS2HW piecewise vol</td>
</tr>
<tr>
<td>100</td>
<td>1.58</td>
<td>1.69</td>
<td>1.60</td>
</tr>
<tr>
<td>200</td>
<td>3.02</td>
<td>3.22</td>
<td>3.06</td>
</tr>
<tr>
<td>300</td>
<td>4.34</td>
<td>4.62</td>
<td>4.37</td>
</tr>
<tr>
<td>400</td>
<td>5.53</td>
<td>5.89</td>
<td>5.60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bps</th>
<th>ITM</th>
<th>ATM</th>
<th>OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H2HW</td>
<td>BS2HW ATM vol</td>
<td>BS2HW piecewise vol</td>
</tr>
<tr>
<td>100</td>
<td>0.57</td>
<td>0.69</td>
<td>0.56</td>
</tr>
<tr>
<td>200</td>
<td>1.08</td>
<td>1.31</td>
<td>1.05</td>
</tr>
<tr>
<td>300</td>
<td>1.53</td>
<td>1.86</td>
<td>1.50</td>
</tr>
<tr>
<td>400</td>
<td>1.93</td>
<td>2.35</td>
<td>1.89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bps</th>
<th>ITM</th>
<th>ATM</th>
<th>OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H2HW</td>
<td>BS2HW ATM vol</td>
<td>BS2HW piecewise vol</td>
</tr>
<tr>
<td>100</td>
<td>0.15</td>
<td>0.15</td>
<td>0.09</td>
</tr>
<tr>
<td>200</td>
<td>0.28</td>
<td>0.28</td>
<td>0.18</td>
</tr>
<tr>
<td>300</td>
<td>0.39</td>
<td>0.39</td>
<td>0.25</td>
</tr>
<tr>
<td>400</td>
<td>0.49</td>
<td>0.49</td>
<td>0.30</td>
</tr>
</tbody>
</table>
6.4. CASE STUDIES

Figure 6.5: 2.5% and 97.5% quantiles under H2HW and BS2HW for a 10-year floating vs floating CCYS for three different moneyness levels and two different periods. At the money (ATM) refers to the case where all cashflows (notional exchanges and future payments) are based on a domestic and foreign principal amount that is implied by the FX ($X_t$) spot rate. For the in the money (ITM) and out of the money (OTM) cases, the FX rates $1.25X_t$ and $0.75X_t$ are used.
in the range of 1500 bps for the constant volatility model. These observations lead to the conclusion that the effects of volatility smile on quantiles cannot be captured neither by a deterministic nor even piecewise constant volatility version of the Black-Scholes model.

6.4.2 Barrier Options

In this section we look into path-dependent UOC barrier options on the EURUSD exchange rate. In addition to the comparisons between H2HW and BS2HW with the different volatilities, we also compute the exposures with the Heston model, such that we can investigate the impact of stochastic interest rates. For the barrier options, we took a maturity of three years. For these type of maturities, barrier levels up to 130% are common choices. We have investigated the barrier levels from 110% to 130% and found that the results were similar. Here we show results for three year UOC options with barrier level 110%.

Densities

The densities shown in Figures 6.6(a) and 6.6(b) are computed by conditioning on the FX rate being lower than barrier level $B$ during the lifetime of the option. We observe the effect of stochastic volatility by noting the fatter left tails of the distributions generated by the Heston models. In addition, the Black-Scholes densities peak further away from the up-and-out barrier level $B$. Again the stochasticity of the interest rates results in a lower peak for the H2HW model compared to Heston.

Exposures and CVA

The probability of hitting the barrier impacts the dynamics of the EE as shown in figures 6.7(a) and 6.7(b). The initial barrier prices ($EE(t_0)$) for the different models are considerably different. These differences between the models persist over time. Comparing 2008 to 2014, we can see that the EEs in 2008 are lower than in 2014 due to lower volatility levels. Next to that the difference between the stochastic volatility and the deterministic volatility models is bigger in 2014 than in 2008 because of the relatively higher skew in 2014.

In table 6.5 we can see that the mean relative differences in CVA is always more than 7% and can be up to 42%. Note that the Heston model is closer to the H2HW model compared to the BS2HW models. Therefore we can conclude that for barrier options the smile has a bigger impact on CVA than the stochastic interest rates.
Table 6.3: Mean Differences (MD\textsubscript{CCYS}) between the exposures of a 10-year floating vs floating CCYS modeled by the Heston 2 Hull-White model or the Black-Scholes model with constant or piecewise constant volatility. The differences are in bps relative to the notional for three different moneyness levels and two different periods.

<table>
<thead>
<tr>
<th></th>
<th>ITM</th>
<th>ATM</th>
<th>OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>EE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant volatility</td>
<td>2.1</td>
<td>6.6</td>
<td>2.3</td>
</tr>
<tr>
<td>piecewise constant volatility</td>
<td>2.0</td>
<td>2.9</td>
<td>2.8</td>
</tr>
<tr>
<td>EPE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant volatility</td>
<td>14.6</td>
<td>37.8</td>
<td>16.6</td>
</tr>
<tr>
<td>piecewise constant volatility</td>
<td>66.4</td>
<td>138.5</td>
<td>167.2</td>
</tr>
<tr>
<td>ENE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant volatility</td>
<td>16.6</td>
<td>40.6</td>
<td>18.1</td>
</tr>
</tbody>
</table>

Table 6.4: Mean Differences (MD\textsubscript{CCYS}) between the quantiles of a 10-year floating vs floating CCYS modeled by the Heston 2 Hull-White model or the Black-Scholes model with constant or piecewise constant volatility. The differences are in bps relative to the notional for three different moneyness levels and two different periods.

<table>
<thead>
<tr>
<th></th>
<th>ITM</th>
<th>ATM</th>
<th>OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q_{0.975}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant volatility</td>
<td>1010</td>
<td>206</td>
<td>1258</td>
</tr>
<tr>
<td>piecewise constant volatility</td>
<td>531</td>
<td>510</td>
<td>659</td>
</tr>
<tr>
<td>Q_{0.925}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant volatility</td>
<td>818</td>
<td>1351</td>
<td>1034</td>
</tr>
<tr>
<td>piecewise constant volatility</td>
<td>167</td>
<td>434</td>
<td>205</td>
</tr>
</tbody>
</table>

Figure 6.6: Densities of the FX rate after three years generated by calibrating H2HW, Heston, and BS2HW with constant and piecewise constant volatility to 2008 (left) or 2014 (right) data. The densities are conditioned on $X_t < B$, where the barrier is set at 110% of $X_0$. 
Figure 6.7: EE and 97.5% quantiles computed under H2HW and BS2HW for a 3 year up-and-out barrier call option with barrier level at $1.1X_0$ for the two periods.

Table 6.5: Differences from and relative to H2HW CVA for up-and-out barrier call options for the 2008 and 2014 periods. The recovery rate is fixed at 0% and the counterparty’s credit spread is assumed to be 50 or 200 bps.

<table>
<thead>
<tr>
<th></th>
<th>2008</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 bps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heston</td>
<td>13.6%</td>
<td>7.9%</td>
</tr>
<tr>
<td>BS2HW ATM vol</td>
<td>38.6%</td>
<td>40.4%</td>
</tr>
<tr>
<td>BS2HW piecewise vol</td>
<td>42.5%</td>
<td>40.5%</td>
</tr>
<tr>
<td>200 bps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heston</td>
<td>13.6%</td>
<td>7.8%</td>
</tr>
<tr>
<td>BS2HW ATM vol</td>
<td>38.7%</td>
<td>40.4%</td>
</tr>
<tr>
<td>BS2HW piecewise vol</td>
<td>42.5%</td>
<td>40.5%</td>
</tr>
</tbody>
</table>
6.4. CASE STUDIES

Quantiles

Again the differences in quantiles computed by the different models are bigger than those in EEs as figures 6.7(c) and 6.7(d) show. We found that the mean relative differences in 97.5% quantiles between a deterministic volatility model and the full H2HW model can be more than 70% in 2008.

One important observation about the quantile is that the model risk for up-and-out barrier options is smaller if the quantile increases. In other words: the difference between the models becomes smaller for higher quantiles. This is due to the fact that the payoff function of a barrier option has a maximum value, namely $B - K$, therefore the maximum exposure is also limited by this value. Whereas in the case of European options the difference can even increase for higher quantiles.

6.4.3 American options

Most traded options are American style and thus allow for early exercise. The early exercise decision depends on the continuation and the intrinsic value of the claim at future time points. The implied exercise boundary can be highly model dependent and therefore it is interesting to explore its effect on exposure dynamics. In this section we consider American put and call options for the two different time periods and analyse the effect of stochastic volatility and rates on the exposure metrics.

Exposures and CVA

The EE profiles for American calls, shown in figures 6.8(a) and 6.8(b), reveal that in 2008 the EE is dropping, whereas in 2014 it does not drop and even slightly increases. This can be explained by looking at the yields shown in figure 6.2. In FX options the foreign yield plays a similar role as dividends in equity options. It is well known that for zero dividend paying stocks the value of American call options are equal to their European counterparts. When dividends are non-zero, it can be profitable to exercise call options early. We will see a similar effect induced by the interplay between the foreign and domestic interest rate which will be controlled by the stochasticity of the rates. In 2008, the foreign Euro yield is higher than the domestic USD yield which causes a negative drift and implies early exercise with a drop in the exposure as a result. In 2014 the domestic USD yield is higher than the foreign Euro yield such that the drift is positive and because of this positive drift, there is no early exercise such that the EE does not drop. In Appendix B.3, the governing drifts over time for the different periods are shown.
CHAPTER 6. SMILE AND DEFAULT

Figure 6.8: EE computed under H2HW and BS2HW for 3 year American call and put options two periods.
6.5. DISCUSSION

Table 6.6: Differences between models expressed as CVA differences from and relative to the H2HW model for American call and put options for the 2008 and 2014 periods. The recovery rate is fixed at 0% and the counterparty’s credit spread is assumed to be 50 or 200 bps.

<table>
<thead>
<tr>
<th></th>
<th>call</th>
<th>put</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2008</td>
<td>2014</td>
</tr>
<tr>
<td>50 bps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heston</td>
<td>27.6%</td>
<td>5.2%</td>
</tr>
<tr>
<td>BS2HW ATM vol</td>
<td>5.0%</td>
<td>2.8%</td>
</tr>
<tr>
<td>BS2HW piecewise vol</td>
<td>7.7%</td>
<td>2.9%</td>
</tr>
<tr>
<td>200 bps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heston</td>
<td>27.3%</td>
<td>5.2%</td>
</tr>
<tr>
<td>BS2HW ATM vol</td>
<td>5.0%</td>
<td>2.8%</td>
</tr>
<tr>
<td>BS2HW piecewise vol</td>
<td>7.7%</td>
<td>2.9%</td>
</tr>
</tbody>
</table>

For the exposures of put options, shown in figures 6.8(c) and 6.8(d), we see that it is always profitable to exercise early as the profiles are decreasing. Looking at the different models we see that the EE modeled under the Heston dynamics is significantly different. Independent of the option type, the Heston model is always underestimating the EE. We can quantify this by looking into the relative differences of the CVAs. Again, as a reference we take the H2HW model and compare the CVAs modeled under Heston and BS2HW with constant and piecewise constant volatility. The resulting differences are shown in table 6.6.

Quantiles

Figure 6.9 shows that the effect seen in exposures is again larger for quantiles. Also for quantiles, the 2014 American call mimics a European quantile profile because of the lack of exercise opportunity. In all other cases we see that the quantile computed with the Heston model is decreasing close to maturity relatively more than the other models. We have computed the mean relative difference between quantiles computed by H2HW and Heston and found that the difference can be more than 30%. For the quantiles and for EE we see that the BS2HW models are closer to H2HW than Heston such that we can conclude that for American options, the stochastic interest rates have a bigger impact than stochastic volatility.

6.5 Discussion

In this chapter we assess the impact of uncertainty in the interest rates and volatility on Counterparty Credit Risk metrics of EURUSD derivatives and Cross-Currency Swaps (CCYS). The analysis is performed by looking at two distinct
Figure 6.9: 97.5% quantiles computed under H2HW and BS2HW for 3 year American call and put options for the two periods.
6.5. DISCUSSION

periods with different market conditions. These periods include the 2008 credit crisis, where high volatilities and moderate interest rates were present, and the aftermath of the crisis in 2014, where volatility is low and fixed income market yields are close to zero. We investigate the impact of using a Heston-2-Hull-White, Heston or Black-Scholes-2-Hull-White with constant or piecewise constant volatility model on exposure metrics of plain vanilla CCYS, Up-and-Out Barrier Call (UOC) and American put and call options.

In the case of CCYS we see that the Black-Scholes model with piecewise constant volatility approximates the H2HW model better than a constant volatility model. However the accuracy of this approximation depends on moneyness of the trade. We found that the EE of ITM or OTM trade modelled by a piecewise constant volatility model is significantly different from the EE modeled by a full H2HW model. Furthermore, the assessment of extreme scenarios (in our study represented by a 2.5% or 97.5% quantile) is much more sensitive to the stochasticity of the volatility levels. This also depends on the market conditions as we observed that the difference between the models with and without skew is bigger in 2014 than in 2008.

When analysing the UOC options we confirm their sensitivity to volatility skew. We see that the models that are calibrated to the same European option market can lead to more than 40 percent differences in the UOC CVAs. This confirms and extends the results that are described in earlier work. We see that for barrier options the stochastic volatility has a bigger impact on exposure and quantiles than the stochastic interest rates.

Similarly as in the previous cases, American options are significantly affected by the risk in volatility levels. Moreover, we observed a substantial difference in models with and without uncertainty in yield curves. The Heston model with constant interest rates shows a different early exercise strategy than the models with stochastic interest rates. We reason this to be due to the possible sign change of the drift of the FX process when short rates are modeled as SDEs. Because of this different early exercise strategy, the stochasticity of interest rates had the biggest impact on the exposure of American options.
7 Discussion

“Market crashes are not randomly occurring lightning bolts; they are the consequence of the madness of crowds who are busy avoiding the last mania as they participate in what will turn out to be the current one”

(Derman [33])

In this work, we studied the applicability of finite difference methods to compute future exposure of financial derivatives. This method is already used in option pricing and, in this context, has advantages compared to Monte Carlo methods. First, the method is able to compute path dependent option values. Second, it is able to compute sensitivities almost without any additional computation. However, because of the curse of dimensionality, this method is typically only applied to low dimensional problems. Here, we investigated if these advantages of finite difference methods also hold for computing exposure. Next, we looked into dimension reduction techniques to circumvent the curse of dimensionality that is normally present in grid based numerical schemes. In this chapter, we will first summarize the contributions of the individual chapters which answer the research questions that are posed in the introduction of this thesis. Secondly,
we will provide some ideas for future research. Finally, we present some thoughts on the future challenges in computational finance.

7.1 Contributions of individual Chapters

Chapter 3. In Chapter 3, three different approaches for computing exposure profiles within the context of counterparty credit risk are presented. The underlying asset exposure is driven by the two-dimensional Heston stochastic volatility model, and this is used to price path-dependent Bermudan put options. In all the three methods scenarios are generated by using a Monte Carlo scheme and option values are priced at each path at each exercise time. Pricing the scenarios is done using either a regression based technique, the COS method or the finite difference method. The combination of finite differences and Monte Carlo sampling, known as the Finite Difference Monte Carlo method (FDMC), is extensively studied in the first part of the thesis.

The Stochastic Grid Bundling Method (SGBM) is based on Monte Carlo simulations. It uses regressed value functions, together with bundling of the Monte Carlo paths to approximate continuation values at different time steps that are needed for Bermudan options. On the other hand, the COS-Monte Carlo (CMC) method combines the Monte-Carlo method with the COS method. This Fourier based technique generates fast and accurate option prices, and this can therefore be efficiently applied to the generated Monte Carlo states. The FDMC method, uses option values on a grid that are obtained by solving the partial differential equation associated with the pricing problem. This PDE on a grid is solved with the finite difference method.

In this chapter, we show that exposure and Potential Future Exposure (PFE) generated with the FDMC method, SGBM or CMC method agree for multiple test cases. We conclude that the FDMC method is able to use all the option values on the pricing grid by interpolating simulated Monte Carlo states on that grid. We get efficient and accurate expected exposures and PFEs using linear interpolation of the Monte Carlo states on the grid.

Chapter 4. In this chapter we looked into the accuracy of the FDMC method in more detail. We did this by investigating exposure profiles and its sensitivities that are driven by the Heston model, and the three-dimensional Heston Hull-White model which is computationally more challenging.

The sensitivities with respect to initial values of the driving risk factors are obtained efficiently by leveraging from the finite difference grid, such that no additional Monte Carlo simulations are needed. For a discontinuous barrier as well as one-touch options, we compare obtained the sensitivities to a “brute force” bump-and-revalue technique. The sensitivities are accurate in the case of
7.1. CONTRIBUTIONS OF INDIVIDUAL CHAPTERS

the Heston Hull-White model, where less grid points per dimensions can be used, because of addition of the third dimension. We found that even in this case, linear interpolation is sufficiently accurate.

Finally, we showed that we can use the FDMC method to compute exposure profiles for a portfolio of multiple options written on one of the underlying risk factors. These options can have different maturities, resulting in discontinuities in time, and we show that this does not cause error propagation.

Chapter 5. The forward Kolmogorov PDE for the transition densities can be solved on a grid, instead of simulating the future distribution of risk factors. Together with option values on the same grid, obtained analytically or via solving the backward Kolmogorov PDE, we can arrive at exposure distribution without the need of the Monte Carlo simulation.

However, solving partial differential equations in dimensions higher than 4 is computationally demanding. Typically the portfolios in finance are driven by more than 4 risk factors, which makes the FDMC method inapplicable in this setting. In this chapter we circumvent this problem by looking into dimension reduction techniques.

In realistic test cases we confirmed that by decomposing the problem, we can approximate exposure profiles for portfolios consisting of multiple derivatives driven by even 7 different risk factors by solving only one, two and three dimensional PDEs. These PDEs can be solved in a fast and accurate way by using finite difference schemes together with splitting schemes for the time discretization. Adding three dimensional corrections in all cases improves the accuracy of the results, but this requires more computational time. The results show that we can use PDE methods to approximate realistic high dimensional exposures.

Chapter 6 Finally in Chapter 6, we looked into the role of the volatility smile, and stochastic interest rates in default risk. By analyzing real market data for two different time periods, we confirmed that the inclusion/exclusion of stochastic volatility has a significant impact on exposure dynamics. Especially those of more complex path-dependent derivatives. The impact of stochastic interest rate varies depending on the situation in the fixed-income markets. Moreover, this impact is most significant in the case of American options. All the results are magnified when one looks at the quantiles of the future option price distribution. We conclude that these factors have to be taken into account, and an industry standard piecewise constant volatility model is not sufficient whenever considering expected or worst case scenario outcomes in Counterparty Credit Risk (CCR).
7.2 Future research on value adjustments

In the context of efficient numerical techniques for value adjustments, important open questions remain. In the past years, developments in computing hardware, mathematical finance as well as smart numerical schemes have been employed to tackle this problem.

**GPUs.** Graphics Processing Units (GPUs) are now extensively used by financial institutions. This hardware was originally used in computer gaming, and can process multiple instructions in parallel generating a big speed-up (up to 10 to 50) compared to traditional CPUs [89]. GPUs are efficient for xVA calculations, as they can compute Monte Carlo simulations in parallel. As a consequence, results can be obtained within an hour, which otherwise would have taken days. Applying GPUs for PDE methods is a more challenging task. At every time step a multi-diagonal system is constructed and solved. This can be done for instance by Gaussian elimination efficiently on a serial computer. However, solving the system is not a parallel routine. In [36], [37] and [98], it is shown that by breaking dependencies, GPUs can be efficiently applied, thereby speeding up the computations with a factor of approximately 20.

Applying GPUs in the context of risk factor decomposition, as explained in Chapter 5, is a promising research direction. The individual corrections are independent, and therefore they can be computed in parallel. This will speedup the computation, and in addition to that it might be possible to employ even higher order corrections.

**AAD.** An important bottleneck in xVA calculations is computing the sensitivities. One way of doing this is by a so-called “bump and revalue” routine. For this routine, the number of scenarios needed are doubled for obtaining a single sensitivity. Even when GPUs are employed, this is not efficient as a large number of sensitivities are needed.

A more promising approach to this problem is Adjoint Algorithmic Differentiation (AAD). AAD is a mathematical method for computing sensitivities that was able to achieve a computational speedup in the order of 50,000 when it was first introduced in a bank [89]. Although it was presented in finance with the tempting title “Adjoint Greeks made easy” [22], the method is typically not straightforward.

Implementing AAD requires a high level of programming skills, whereas employing adjoint techniques for PDEs is more straightforward. As we have shown in Chapter 5, the forward Kolmogorov PDE can be solved by transposing the systems related to the backward Kolmogorov technique. A promising direction for future research would be to use this adjoint relation between the forward and
7.3. **FUTURE CHALLENGES**

backward Kolmogorov PDE in the context of complex path dependent derivatives. Together with risk factor decomposition, exposures of high dimensional portfolios consisting of path-dependent options, can then be efficiently computed.

### 7.3 Future challenges

By charging value adjustments, financial institutions are now able to hedge counterparty risks that were previously considered “unhedgeable”. This has increased the complexity of option pricing. In what follows, we discuss additional challenges that still remain in this framework.

**Liquidity risk.** In many publications on xVA, and also in this thesis, the main focus is on the computational challenge of computing future exposures. It is assumed that the default probability and the recovery rate can be implied from the market. As explained in the introduction, the default probabilities can be estimated from Credit Default Swap (CDS) quotes. However, for up to 80% of the counterparties, these quotes are not available in the market, which is a source of liquidity risk. For approximating a CDS spread, non-conventional methodologies are needed, which rely on the available information of the illiquid counterparties. Techniques from data science and machine learning are now increasingly being used for this purpose.

**Wrong-Way Risk (WWR).** Another source of risk that is not considered in this work is the correlation between exposure and probability default. If the default probability of a counterparty increases together with an increasing exposure to the counterparty, the phenomenon called Wrong-Way Risk (WWR) is encountered. A typical example is when a bank sells a CDS on a reference entity, which has a similar profile as the bank itself, to a counterparty. Now, when the reference entity’s credit spread is deteriorating, the value of the CDS increases for the counterparty. However, as the bank has a similar profile as the reference entity, its default probability also increases. Incorporating WWR should thus heavily increase the CVA charge. After the recent Brexit, the credit spreads jumped and interest rates fell at the same time, which resulted in big single-day losses for banks [90].

The difficulty of modeling WWR is twofold. First, similar to the liquidity risk in CDS quotes, there is not enough data to extract default probabilities, making it very hard to calibrate a possible model that incorporates the correlation between the default probability and the exposure. Secondly, there is no liquid product that enables institutions to hedge away this risk.
**Systemic risk.** The inclusion of value adjustments drives financial institutions to trade with counterparties with better creditworthiness. This will have an influence on the network of trading institutions, which plays an important role in the context of systemic risk. This risk refers to the possible collapse of an entire financial system. An example of such a collapse is the 2008 credit crisis. After this crisis, there has been an increase in the research on network theory in finance. Different techniques have been used to study the formation and evolution of financial networks [10]. It is shown that in the event of a crisis, the structure of an interbank network changes significantly [95]. However, this change is still not fully understood. In addition, the information that is needed to infer such an interbank network is difficult to obtain, because of the limited availability of data. As a result analyzing these financial complex systems is not straightforward.

The annual review [76] of the German financial sector, published in June 2016 by the International Monetary Fund (IMF), stated that the Deutsche bank is a possible source of systemic risk. At the same time, the bank failed the annual stress test of the Fed [91], which increases the worry for a possible “spillover” effect due to the connectedness and global importance of the bank.

How to incorporate this systemic risk into the pricing framework for financial derivatives is an open question. The credit crisis made institutions consider the creditworthiness of their counterparties. Systemic risk, in addition, will make institutions look at the creditworthiness of all counterparties of the counterparty they are trading with. It might be possible that in the future, to trade an individual derivative with one counterparty, the full interbank network needs to be taken into account to arrive at a fair price. This outlook forms a major future challenge for scientists, market participants and regulators.
A

CVA and sensitivities for barrier options

For the evaluation of CVA, we assume a recovery rate of 40%. The hazard rate is computed by assuming a 5 year CDS with a spread of 400 basis points paid quarterly. The Euro discount factors are taken from April 2014, the resulting survival probabilities up to one year are obtained as explained in section 4.3.3. Because we assume absence of wrong - and right - way risk, we can compute the CVA for any CDS spread. In this case, CVA is a linear function of the CDS spread. To investigate the barrier effect on CVA, we compute CVA and its sensitivities as a function of the barrier level, in figure A.1. The strike is set at -the - money (\(S_0 = K\)) and for barriers lower than \(K\) we look at Down - and - Out Put (DOP) options, while for barriers higher than strike we compute Up - and - Out Call (UOC) options. As expected for barrier options with a barrier close to strike, the CVA as well as the sensitivities decrease to zero.

To further investigate the impact of the CVA and its sensitivities we look at
the difference between the measures with and without CVA adjustment:

\[ U^* = U + \text{CVA}, \]

\[ \frac{\partial U^*}{\partial S_0} = \frac{\partial U}{\partial S_0} + \frac{\partial \text{CVA}}{\partial S_0}, \]

\[ \frac{\partial^2 U^*}{\partial S_0^2} = \frac{\partial^2 U}{\partial S_0^2} + \frac{\partial^2 \text{CVA}}{\partial S_0^2}. \]

Table A.1: The change of price, delta and gamma by adding CVA to barrier options for various moneyness levels. The significant numbers are computed up to 1% for CVA and delta and 5% for gamma due to the standard deviation.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>OTM $K = 0.9S_0$</th>
<th>ATM $K = S_0$</th>
<th>ITM $K = 1.1S_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVA</td>
<td>1.99%</td>
<td>2.00%</td>
<td>1.99%</td>
</tr>
<tr>
<td>Delta</td>
<td>4.70%</td>
<td>3.47%</td>
<td>3.66%</td>
</tr>
<tr>
<td>Gamma</td>
<td>-15%</td>
<td>-0.39%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Test B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVA</td>
<td>3.96%</td>
<td>3.98%</td>
<td>3.98%</td>
</tr>
<tr>
<td>Delta</td>
<td>16.75%</td>
<td>9.96%</td>
<td>13.3%</td>
</tr>
<tr>
<td>Gamma</td>
<td>-13%</td>
<td>-5.1%</td>
<td>2.0%</td>
</tr>
</tbody>
</table>

In Table A.1 we show the impact of CVA on the option price and its sensitivities. The CVA delta and gamma are quoted as percentages of the risk-free valuation ($U$) where no CVA is charged. We can conclude that, for different moneyness levels, the CVA adjustment is stable, but the delta and gamma differ significantly. Furthermore, figures A.1(a) to A.1(f) show that the impact of skew on CVA and its sensitivities cannot be ignored. We see that when the Heston or the Black-Scholes dynamics are assumed, gamma can even differ in sign.
Figure A.1: CVA as a function of the barrier level. The CVA is calculated assuming a LGD of 40% and a fixed CDS spread of 400 basis points. The number of paths is equal to 100,000 and for the FD computation we take 250 grid points in $V$- and 500 grid points in $S$-direction. The standard error is less than 3% for all barrier levels.
Market data

B.1 Implied volatility bootstrapping

The resulting different volatility levels that are used for the BS2HW models are shown in Figure B.1 together with the mean of the volatility over time when it is modeled as an SDE in the H2HW model.

Figure B.1: Effective volatilities over time for the three different models and two different periods.
Table B.1: Calibrated parameters for the H2HW model for the 2008 and 2014 periods.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>2008</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>FX rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_0$</td>
<td>1.41</td>
<td>1.25</td>
</tr>
<tr>
<td>Volatility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$v_0$</td>
<td>0.0204</td>
<td>0.0071</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.2174</td>
<td>0.2475</td>
</tr>
<tr>
<td>$\bar{v}$</td>
<td>0.0039</td>
<td>0.0160</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.1000</td>
<td>0.1037</td>
</tr>
<tr>
<td>Domestic IR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^d_0$</td>
<td>0.0297</td>
<td>-0.0036</td>
</tr>
<tr>
<td>$\lambda^d_0$</td>
<td>0.0100</td>
<td>0.0100</td>
</tr>
<tr>
<td>$\eta^d_0$</td>
<td>0.0105</td>
<td>0.0087</td>
</tr>
<tr>
<td>Foreign IR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^f_0$</td>
<td>0.0489</td>
<td>0.0002</td>
</tr>
<tr>
<td>$\lambda^f_0$</td>
<td>0.0100</td>
<td>0.0100</td>
</tr>
<tr>
<td>$\eta^f_0$</td>
<td>0.0067</td>
<td>0.0070</td>
</tr>
<tr>
<td>Correlations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{S,v}$</td>
<td>-0.0796</td>
<td>-0.2079</td>
</tr>
<tr>
<td>$\rho_{S,r^d}$</td>
<td>0.1572</td>
<td>0.1572</td>
</tr>
<tr>
<td>$\rho_{S,r^f}$</td>
<td>-0.2957</td>
<td>-0.2957</td>
</tr>
<tr>
<td>$\rho_{v,r^d}$</td>
<td>-0.9000</td>
<td>-0.8500</td>
</tr>
<tr>
<td>$\rho_{v,r^f}$</td>
<td>-0.7500</td>
<td>-0.6000</td>
</tr>
<tr>
<td>$\rho_{r^d,r^f}$</td>
<td>0.6432</td>
<td>0.6432</td>
</tr>
</tbody>
</table>

B.2 Calibrated parameters

The parameters obtained by the described calibration procedure for the two periods are shown in Table B.1. A closer examination of the parameters reveals that the volatility term structure is reflected in the initial and long term stochastic volatility levels. Also, the larger skew in 2014 is reflected in the more negative $\rho_{S,v}$. Other than this, the calibrated parameters seem to be somewhat similar between the two periods.

B.3 Drift over time

The governing drifts $(r^d_t - r^f_t)$ over time for the different models are presented in Figures B.2.
B.4 Model parameters

The full correlation matrix for the process \( (X^1_t, \ldots, X^7_t) = (F^1_t, R^1_t, F^{1.1}_t, R^{1.1}_t, F^{1.2}_t, R^{1.2}_t, F^{1.3}_t, R^{1.3}_t) \) from (5.39), used in the tests of Section 5.6, after regularisation for positive definiteness [82]:

\[
\begin{pmatrix}
1 & -0.3024 & 0.1226 & 0.5815 & -0.0142 & 0.5510 & 0.5351 \\
-0.3024 & 1 & 0.6293 & -0.2577 & 0.6895 & -0.4554 & 0.3188 \\
0.1226 & 0.6293 & 1 & 0.0459 & 0.7453 & -0.3040 & 0.4181 \\
0.5815 & -0.2577 & 0.0459 & 1 & 0.1230 & 0.5490 & -0.0848 \\
-0.0142 & 0.6895 & 0.7453 & 0.1230 & 1 & -0.3015 & 0.3587 \\
0.5510 & -0.4554 & -0.3049 & 0.5490 & -0.3015 & 1 & -0.3260 \\
0.5351 & 0.3188 & 0.4181 & -0.0848 & 0.3587 & -0.3260 & 1 \\
\end{pmatrix}
\]

The FX and interest rate SDEs are driven by the parameters in Table B.2.

<table>
<thead>
<tr>
<th>EUR</th>
<th>( R^j_0 )</th>
<th>( F^i_0 )</th>
<th>( i = 1 ) (US)</th>
<th>( i = 2 ) (GB)</th>
<th>( i = 3 ) (JP)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.8157e-04</td>
<td>1.2470</td>
<td>0.7926</td>
<td>147.53</td>
<td></td>
</tr>
<tr>
<td>( \lambda_d )</td>
<td>0.010</td>
<td>( \gamma^1_f )</td>
<td>0.010</td>
<td>0.0523</td>
<td>0.010</td>
</tr>
<tr>
<td>( \eta_d )</td>
<td>0.0070</td>
<td>( \eta^1_l )</td>
<td>0.0092</td>
<td>0.0104</td>
<td>0.0057</td>
</tr>
</tbody>
</table>

Table B.2: Model parameters.
APPENDIX B. MARKET DATA

As mentioned earlier, the $\Theta_d(t)$, and $\Theta_i^f(t), i = 1, 2, 3$, are calibrated to fit the forward rate curve of the respective markets as seen on 2 December 2014. For this date, the ATM volatilities of the FX rates are used, which can be found in Table B.3.

<table>
<thead>
<tr>
<th>$T$ = 1M</th>
<th>EURUSD(%)</th>
<th>EURGBP(%)</th>
<th>EURJPY(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.852</td>
<td>6.570</td>
<td>10.247</td>
<td></td>
</tr>
<tr>
<td>8.695</td>
<td>6.635</td>
<td>10.245</td>
<td></td>
</tr>
<tr>
<td>8.580</td>
<td>7.350</td>
<td>10.517</td>
<td></td>
</tr>
<tr>
<td>8.605</td>
<td>7.447</td>
<td>10.848</td>
<td></td>
</tr>
<tr>
<td>8.717</td>
<td>7.865</td>
<td>11.580</td>
<td></td>
</tr>
<tr>
<td>8.952</td>
<td>8.068</td>
<td>12.247</td>
<td></td>
</tr>
<tr>
<td>9.635</td>
<td>8.383</td>
<td>13.642</td>
<td></td>
</tr>
</tbody>
</table>

Table B.3: ATM volatilities as seen on 2 December 2014 that are used for bootstrapping the piecewise constant volatility function as explained in Section 5.5.3.

B.5 Numerical parameters

To compute the approximations, we solve one-, two- and three-dimensional PDEs by finite differences on a grid, as detailed in Section 5.4. The grids are non-uniform as in [56], where a sinh transformation is used to place a high density of grid points around the spot values. Moreover, the grid is shifted to include the initial spot value where the non-smooth Dirac delta function for the forward equation is located.

In Table B.4 the computational domain in $F_t$, $R_d^t$ and $R_f^t$ is chosen. The parameter $\xi$ controls the fraction of points that lie close to initial spot values (see [56]).

<table>
<thead>
<tr>
<th></th>
<th>min</th>
<th>max</th>
<th>$\xi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FX rate ($F_t$)</td>
<td>0</td>
<td>$8F_0$</td>
<td>20</td>
</tr>
<tr>
<td>Domestic IR ($R_d^t$)</td>
<td>-0.5</td>
<td>0.8</td>
<td>100</td>
</tr>
<tr>
<td>Foreign IR ($R_f^t$)</td>
<td>-0.5</td>
<td>0.8</td>
<td>100</td>
</tr>
</tbody>
</table>

Table B.4: Finite difference grid parameters.

Moreover, we use upwind differences for large absolute interest rates, i.e., $R \in [-0.5, -0.1] \cup [0.2, 0.8]$, otherwise central differences.

The number of grid points $m_1$, $m_2$, $m_3$ in the three directions is chosen as $m_1 = 2m_2 = 2m_3$; see also Appendix B.7.
For the $F$ grid, an interval $[F_{\text{left}}, F_{\text{right}}] \subseteq [F_{\text{min}}, F_{\text{max}}]$ is defined wherein the grid is uniform and dense, similar to [55], while outside the sinh transformation is used. We set this interval as $[F_{\text{left}}, F_{\text{right}}] = [0.95F_0, 1.02F_0]$.

**B.6 Regression-based Monte Carlo algorithm**

We use constant, linear and bi-linear basis functions

$$\{\psi_j : j = 1, \ldots, 6\} = \{F, R^d, R^f, FR^d, FR^f, R^dR^f\},$$

for the correlation between risk factors. The algorithm determines option values $Y(t) = Y(t; \omega_k)$ along the sample path $\omega_i$, and can be summarised by the following steps:

0. Generate all the paths $(F_t(\omega), R^d_t(\omega), R^f_t(\omega))$ at all the time points $t = 0, \Delta t, \ldots$

1. Start from maturity and set $t \leftarrow T$, $Y_i(T) \leftarrow \phi(F_T(N_{\omega_i}))$.

2. Set $Y_i(t) \leftarrow \exp(-\Delta t R^d_t(\omega)) Y_i(t)$.

3. Solve the linear regression problem

$$\mathbb{E} \left[ \left( \mathbb{E} \left[ Y(t) | X_{t-\Delta t} \right] - \sum_{j=1}^{6} \beta_j \psi_j(X_{t-\Delta t}) \right)^2 \right] \longrightarrow \min \beta$$

by approximating the expectations on the paths $\omega_i$ to find a minimiser $\hat{\beta}$.

4. Set $t \leftarrow T - \Delta t$ and

$$Y_i(t) \leftarrow \hat{\beta}_0 + \hat{\beta}_1 F_t(\omega_i) + \hat{\beta}_2 R^d_t(\omega_i) + \hat{\beta}_3 R^f_t(\omega_i)$$

$$+ \hat{\beta}_4 F_t(\omega_i)R^d_t(\omega_i) + \hat{\beta}_5 F_t(\omega_i)R^f_t(\omega_i) + \hat{\beta}_6 R^d_t(\omega_i)R^f_t(\omega_i)$$

(B.2)

5. Repeat backwards from 2. for all time steps.

**B.7 Finite difference errors and individual terms for Case B**

Here we show the individual FD errors for all terms involved for Case B, compared to a Monte Carlo estimate, for both EE (Table B.5) and EPE (Table B.6).
APPENDIX B. FD ERRORS

<table>
<thead>
<tr>
<th>BS EU</th>
<th>MC SE</th>
<th>m₁ = 40</th>
<th>m₁ = 60</th>
<th>m₁ = 80</th>
<th>m₁ = 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSBS EU-EG</td>
<td>0.19</td>
<td>0.23 (0.20)</td>
<td>0.072</td>
<td>0.13</td>
<td>0.087</td>
</tr>
<tr>
<td>BSBS EU-EJ</td>
<td>0.16</td>
<td>0.22 (0.19)</td>
<td>0.050</td>
<td>0.10</td>
<td>0.057</td>
</tr>
<tr>
<td>BSHW EU-RE</td>
<td>0.15</td>
<td>0.24 (0.15)</td>
<td>0.15</td>
<td>0.20</td>
<td>0.17</td>
</tr>
<tr>
<td>BSHW EU-RU</td>
<td>0.22</td>
<td>0.22 (0.20)</td>
<td>0.23</td>
<td>0.18</td>
<td>0.21</td>
</tr>
<tr>
<td>BSHW EU-RG</td>
<td>0.22</td>
<td>0.22 (0.18)</td>
<td>0.058</td>
<td>0.11</td>
<td>0.071</td>
</tr>
<tr>
<td>BSHW EU-RJ</td>
<td>0.22</td>
<td>0.21 (0.17)</td>
<td>0.061</td>
<td>0.11</td>
<td>0.075</td>
</tr>
<tr>
<td>BSBBS EU-EG-EJ</td>
<td>0.21</td>
<td>0.24 (0.20)</td>
<td>0.088</td>
<td>0.14</td>
<td>0.10</td>
</tr>
<tr>
<td>BSHWHW EU-RE-RU</td>
<td>0.21</td>
<td>0.16 (0.14)</td>
<td>0.14</td>
<td>0.098</td>
<td>0.11</td>
</tr>
<tr>
<td>BSBBSH EU-EG-RE</td>
<td>0.21</td>
<td>0.23 (0.15)</td>
<td>0.16</td>
<td>0.21</td>
<td>0.18</td>
</tr>
<tr>
<td>BSBBSH EU-EJ-RE</td>
<td>0.21</td>
<td>0.22 (0.15)</td>
<td>0.13</td>
<td>0.18</td>
<td>0.15</td>
</tr>
<tr>
<td>BSBBSH EU-EG-RU</td>
<td>0.22</td>
<td>0.24 (0.23)</td>
<td>0.21</td>
<td>0.17</td>
<td>0.19</td>
</tr>
<tr>
<td>BSBBSH EU-EJ-RU</td>
<td>0.22</td>
<td>0.22 (0.21)</td>
<td>0.20</td>
<td>0.16</td>
<td>0.18</td>
</tr>
<tr>
<td>BSBBSH EU-EG-RG</td>
<td>0.22</td>
<td>0.25 (0.22)</td>
<td>0.088</td>
<td>0.14</td>
<td>0.096</td>
</tr>
<tr>
<td>BSBBSH EU-EJ-RJ</td>
<td>0.22</td>
<td>0.23 (0.19)</td>
<td>0.063</td>
<td>0.12</td>
<td>0.074</td>
</tr>
</tbody>
</table>

Table B.5: Finite difference errors of exposures (EE) of a portfolio with three CCYS. The errors are measured in $e_{L2}$ (see equation (5.37)) in percentages for different number of grid points in space. The number of time steps is fixed at 500. Within brackets (for $m₁ = 40$) the error for a FD solution with 1000 timesteps. The Monte Carlo benchmark is computed with $4 \cdot 10^6$ paths and 1000 time steps.

Remark 4. In the following tables, for instance, BSHW EU-RU refers to $V_{(1,5)}(F₁,RF₁)$ from Section 5.6, i.e., a Black-Scholes-Hull-White model where the EURUSD rate $F₁$ follows a Black-Scholes-type model but with Hull-White foreign short rate $R₁$ and deterministic domestic rate, and similarly for the other terms.

The results in Table B.5 (especially) and Table B.6 show that the FD errors are already in the same order of magnitude as the MC standard error for a larger number of samples (i.e., significantly more samples than are used in practice). We therefore use $m₁ = 60$ mesh points and 500 time steps in most of the computations (unless otherwise stated).

In Table B.7, we report the accuracy of the complete approximations for increasing number of mesh points $m₁$. For reference, we also compute the approximation with a standard Monte Carlo estimator.

Remark 5. In the first column of Table B.7, we report the results for an estimator where the same Brownian paths are used for the estimation of all $V_u$ in (5.19) for a given $ΔV_u$, and in brackets the results if the same paths are also used across all $u$. It does not seem generally clear which of the estimators has the smaller variance. Using independent paths for different $u$ results in a summation of the variances of $ΔV_u$. Using the same paths for different $u$ is expected to increase the
Table B.6: Finite difference errors of positive exposures (EPE) of a portfolio with three CCYS. The errors are measured in $e_{L^2}$ (see equation (5.37)) in percentages for different number of grid points in space. The number of time steps is fixed at 500. Within brackets ($m_1 = 40$) the error for a FD solution with 1000 timesteps. The Monte Carlo benchmark is computed with $4 \cdot 10^6$ paths and 1000 time steps.

<table>
<thead>
<tr>
<th></th>
<th>MC SE</th>
<th>$m_1 = 40$</th>
<th>$m_1 = 60$</th>
<th>$m_1 = 80$</th>
<th>$m_1 = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS EU</td>
<td>0.071</td>
<td>0.30 (0.34)</td>
<td>0.099</td>
<td>0.042</td>
<td>0.030</td>
</tr>
<tr>
<td>BSBS EU-EG</td>
<td>0.072</td>
<td>0.23 (0.25)</td>
<td>0.11</td>
<td>0.078</td>
<td>0.071</td>
</tr>
<tr>
<td>BSBS EU-EJ</td>
<td>0.072</td>
<td>0.24 (0.27)</td>
<td>0.070</td>
<td>0.038</td>
<td>0.048</td>
</tr>
<tr>
<td>BSHW EU-RE</td>
<td>0.072</td>
<td>0.31 (0.37)</td>
<td>0.11</td>
<td>0.050</td>
<td>0.041</td>
</tr>
<tr>
<td>BSHW EU-RU</td>
<td>0.094</td>
<td>0.45 (0.42)</td>
<td>0.24</td>
<td>0.17</td>
<td>0.14</td>
</tr>
<tr>
<td>BSHW EU-RG</td>
<td>0.093</td>
<td>0.29 (0.32)</td>
<td>0.090</td>
<td>0.040</td>
<td>0.043</td>
</tr>
<tr>
<td>BSHW EU-RJ</td>
<td>0.093</td>
<td>0.30 (0.34)</td>
<td>0.10</td>
<td>0.054</td>
<td>0.048</td>
</tr>
<tr>
<td>BSBSBS EU-EG-EJ</td>
<td>0.064</td>
<td>0.21 (0.23)</td>
<td>0.090</td>
<td>0.056</td>
<td>0.045</td>
</tr>
<tr>
<td>BSBSHWH EU-RE-RU</td>
<td>0.091</td>
<td>0.53 (0.51)</td>
<td>0.31</td>
<td>0.24</td>
<td>0.21</td>
</tr>
<tr>
<td>BSBSHWH EU-EG-RE</td>
<td>0.069</td>
<td>0.22 (0.25)</td>
<td>0.12</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>BSBSHWH EU-EJ-RE</td>
<td>0.083</td>
<td>0.25 (0.29)</td>
<td>0.080</td>
<td>0.054</td>
<td>0.062</td>
</tr>
<tr>
<td>BSBSHWH EU-EG-RU</td>
<td>0.071</td>
<td>0.33 (0.30)</td>
<td>0.18</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td>BSBSHWH EU-EJ-RU</td>
<td>0.085</td>
<td>0.44 (0.41)</td>
<td>0.26</td>
<td>0.20</td>
<td>0.17</td>
</tr>
<tr>
<td>BSBSHWH EU-EG-RG</td>
<td>0.070</td>
<td>0.26 (0.28)</td>
<td>0.12</td>
<td>0.076</td>
<td>0.058</td>
</tr>
<tr>
<td>BSBSHWH EU-EJ-RJ</td>
<td>0.084</td>
<td>0.29 (0.32)</td>
<td>0.11</td>
<td>0.051</td>
<td>0.030</td>
</tr>
</tbody>
</table>

Table B.7: Errors of decomposed approximations to exposures of a portfolio with three CCYS (see Table 5.2, Case B), computed with different numbers of mesh points. The errors are measured in $e_{L^2}$ (see equation (5.37)) in percentages for different number of grid points in space. The number of time steps is fixed at 500. The Monte Carlo estimator for these approximation errors is computed with $4 \cdot 10^6$ paths and 1000 time steps. Within brackets for MC the error for a MC simulation with identical paths for all estimators, as explained in Remark 5. Within brackets for $m_1 = 40$ the error for a FD computation with 1000 timesteps.

<table>
<thead>
<tr>
<th></th>
<th>MC SE</th>
<th>$m_1 = 40$</th>
<th>$m_1 = 60$</th>
<th>$m_1 = 80$</th>
<th>$m_1 = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D cort.</td>
<td>1.27</td>
<td>1.35 (1.31)</td>
<td>1.41</td>
<td>1.36</td>
<td>1.39</td>
</tr>
<tr>
<td>3D cort.</td>
<td>0.19</td>
<td>0.25 (0.27)</td>
<td>0.20</td>
<td>0.18</td>
<td>0.20</td>
</tr>
<tr>
<td>EPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D cort.</td>
<td>2.39</td>
<td>2.20 (2.21)</td>
<td>2.29</td>
<td>2.32</td>
<td>2.36</td>
</tr>
<tr>
<td>3D cort.</td>
<td>0.41</td>
<td>0.31 (0.29)</td>
<td>0.26</td>
<td>0.26</td>
<td>0.26</td>
</tr>
</tbody>
</table>

The variance if the terms are predominantly positively correlated, and decrease the variance if they are predominantly negatively correlated.
### APPENDIX B. FD ERRORS

#### Table B.8: Contribution of individual corrections for EE and EPE of a portfolio with three CCYS.

The differences are measured in $e_{L_2}$ (see equation (5.37)) in percentages for $m_1 = 60$ of grid points in space.

<table>
<thead>
<tr>
<th></th>
<th>EE</th>
<th>EPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSBS EU-EG</td>
<td>14.58</td>
<td>40.35</td>
</tr>
<tr>
<td>BSBS EU-EJ</td>
<td>8.74</td>
<td>18.40</td>
</tr>
<tr>
<td>BSHW EU-RE</td>
<td>1.40</td>
<td>2.73</td>
</tr>
<tr>
<td>BSHW EU-RU</td>
<td>0.52</td>
<td>0.51</td>
</tr>
<tr>
<td>BSHW EU-RG</td>
<td>0.0003</td>
<td>0.0022</td>
</tr>
<tr>
<td>BSHW EU-RJ</td>
<td>0.000</td>
<td>0.0093</td>
</tr>
<tr>
<td>BSBSBS EU-EG-EJ</td>
<td>0.000</td>
<td>0.50</td>
</tr>
<tr>
<td>BSHWHW EU-RE-RU</td>
<td>1.19</td>
<td>2.37</td>
</tr>
<tr>
<td>BSBSHW EU-EG-RE</td>
<td>0.31</td>
<td>1.28</td>
</tr>
<tr>
<td>BSBSHW EU-EJ-RE</td>
<td>0.18</td>
<td>0.72</td>
</tr>
<tr>
<td>BSBSHW EU-EG-RU</td>
<td>0.000</td>
<td>0.083</td>
</tr>
<tr>
<td>BSBSHW EU-EJ-RU</td>
<td>0.000</td>
<td>0.58</td>
</tr>
<tr>
<td>BSBSHW EU-EG-RG</td>
<td>0.046</td>
<td>0.066</td>
</tr>
<tr>
<td>BSBSHW EU-EJ-RJ</td>
<td>0.12</td>
<td>0.30</td>
</tr>
</tbody>
</table>
Bibliography


BIBLIOGRAPHY 153


BIBLIOGRAPHY


Acknowledgements

First I would like to thank my supervisor Drona Kandhai for the guidance and support during the past four years. When needed, he was always available to give advice, also outside regular office hours. I appreciate how he taught me to set and achieve ambitious goals. Already in my first year he gave me a lot of responsibility by letting me assist in the course Computational Finance and supervise MSc students. Furthermore he made me feel welcome at ING Bank where I could learn firsthand from practitioners. I am also grateful to Peter Sloot for giving me the opportunity to pursue a PhD in the Computational Science Lab. His comments on my writing were very helpful and his encouraging comments during our meetings gave me a lot of confidence. I would also like to thank the PhD committee for their efforts in assessing this thesis and acting as opponents during its defense.

I am thankful to Kees Oosterlee for introducing me to Drona, and thereby showing me the possibility to pursue a PhD. It was a luxury to have his research group at CWI just across the street. Thanks to Qian for our very successful collaboration which really had a great start with our first publication. The user committee of the CVA project sponsored by STW always gave useful advice. Thanks to Anton for all the fun during the conferences we visited together. Your jokes really helped me to put everything in perspective. I would also like to thank John, Sarunas and Geert Jan for our successful projects.

I was lucky to spend 4 months in Oxford to do research at the Mathematical Institute. Thanks to Christoph Reisinger I was warmly welcomed in his research group and it felt great to be part of it. Andrei, Matthieu, you were the best! I really enjoyed our coffee breaks at the Jericho café where we could discuss numerical techniques and share our love for modeling (stocks).

I am very happy to thank Amir. It always surprises me how much we have in common. Thank you for all the help in designing my thesis cover and for being such a good friend. Thanks Juriaan for being who you are: critical, funny and loyal. Also your help with my Dutch summary was very valuable. Thanks to Willem my first few months in Amsterdam were extra nice, I could not have wished for a better personal chef.

At the Computational Science Lab I liked the regular coffee breaks, our sys-
tematic lunch schedule (thanks Rick!) and our nice outings. Sumit and Ioannis came just in time to help me with my introduction, thank you for that! I have noticed that it is important with whom you do a PhD, thanks Rick, Omri, Louis, Philip, Debraj, Jurjen, Eva, Guusje, Brecht, Frederik, Yuki, Roland, Anna, Joris, Hannan, Gabor, Britt, David, Saad, Lampros, Mikolaj, Emiliano, Jaap, Mike and Alfons for being great colleagues.

I owe a lot to my family, therefore I would like to thank my parents for letting me find my own way and supporting me when needed. Also I am grateful to Anne, Julien, Indie and Stella for being so involved despite the distance between us. Thanks Flip for keeping up with me while complaining on the bike. Also Marlène for being a great example for Donja. Finally I want to thank Donja for being so beautiful and funny. Over the past years you have showed me what is important in life. Even at times when the workload was overwhelming, you were always able to put a smile on my face, please keep on doing that.
List of publications

Journal articles


Conference presentations


Contributions of co-authors


Author contributions: de Graaf, Feng, Oosterlee and Kandhai designed the research, de Graaf and Feng performed the tests and analyzed the results under the guidance of Oosterlee and Kandhai, de Graaf, Feng, Kandhai and Oosterlee wrote the paper.


Author contributions: de Graaf and Kandhai designed the research, de Graaf performed the tests and the results were analyzed jointly by de Graaf and Kandhai, de Graaf and Kandhai wrote the paper which was further revised by Sloot.
BIBLIOGRAPHY


Author contributions: The initial code was written by Simaitis under the guidance of de Graaf and further revised and finalized by de Graaf. The numerical tests and analysis have been conducted jointly by Simaitis and de Graaf in close collaboration with Hari and Kandhai. The data and explanations of the market conventions is provided by Hari and Kandhai. The paper is written by Simaitis and de Graaf and further revised by all authors.


Author contributions: de Graaf, Reisinger and Kandhai designed the research, Kandhai provided the data and insights in setting up realistic test cases and numerical experiments, de Graaf coded the FD and MC solvers, de Graaf and Reisinger performed the tests and analyzed the results in collaboration with Kandhai, de Graaf and Reisinger wrote the paper.