Finite Difference Methods for the Non-linear Black-Scholes-Barenblatt Equation

INAUGURAL-DISSERTATION

zur
Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von
Albrecht Budke
aus Quakenbrück

München, 2013

Berichterstatter: Prof. Dr. R. Seydel
PD Dr. habil. P. Heider

Tag der mündlichen Prüfung: 27. November 2013

Acknowledgments

My special acknowledgment is owed to my supervisor Prof. Dr. Seydel for his excellent support and precious advice. His working group offered an eminently suitable research environment. I also have to express my appreciation to Dr. habil. Pascal Heider for acting as referee for the present thesis. Additionally, I would like to thank Prof. Dr. Rainer Schrader for serving as chairman of the disputation examination committee and Dr. Markus Schulz for acting as the corresponding secretary. Thanks are also due to my former colleagues Christian Jonen, Alexander Schröter, and Markus Lücking for helpful advice and fruitful discussions. You all made me enjoy my time at the Department of Mathematics. Thank you.
In addition, deep gratitude is owed to my parents, my siblings, and my grandparents for their support during my studies and my time as a doctoral student at the University of Cologne. You helped me to forge ahead and always provided a place to feel home. Furthermore, I would like to thank Stefan Müller for passionate discussions on mathematical problems and Manuel Dingerkus for motivation and relieving distractions. You both and many other friends I thank for auxiliary remarks as well as helpful suggestions to improve this thesis.

Abstract

The Uncertain Volatility model is a non-linear generalisation of the Black-Scholes model in the sense that volatility and correlation can take arbitrary values in given intervals. The value of an option is then given by a non-linear partial differential equation of Hamilton-Jacobi-Bellman type. For this type of equation the concept of viscosity solution has to be considered since in general no smooth solutions in the classical sense exist.

To assure the convergence of a discrete scheme it has to be consistent, stable and additionally monotone. Starting from a Finite Difference discretisation we first derive general and structural conditions to adequately price options in the Uncertain Volatility model. Additionally, an optimisation problem has to be solved which can either be done exactly or only approximatively where this choice depends on the discretisation. Finally, sufficient conditions are derived for the different discrete schemes to assure their convergence to the viscosity solution.

The obtained theoretical results are finally tested in numerical experiments. The rates of convergence, the effort for excecuting the policy iteration, and the possible gain by using non-uniform grids are analysed.

Zusammenfassung

Das Uncertain-Volatility-Modell stellt eine nicht-lineare Erweiterung des Black--Scholes-Models insofern dar, dass Volatilität und Korrelation beliebige Werte in vorgegebenen Intervallen annehmen können. Der Wert einer Option ist dann durch eine nicht-lineare partielle Differentialgleichung vom Hamilton-Jacobi-Bellman-Typ bestimmt. Da Lösungen im klassischen Sinne für solche im Allgemeinen nicht existieren, wird das Konzept der Viskositätslösungen, für die die Existenz einer Lösung der PDE garantiert werden kann, verwendet.

Um die Konvergenz eines Verfahrens sicherzustellen, muss zusätzlich zu Konsistenz und Stabilität Monotonie nachgewiesen werden. Ausgehend von einem Ansatz mittels finiter Differenzen werden zu erst allgemeine und strukturelle Bedingungen hergeleitet um Optionen im Uncertain-Volatility-Modell zu bewerten. Abhängig von der Diskretisierung der Gleichung gehört dazu auch die exakte oder approximative Lösung von nicht-linearen Optimierungsproblemen. Anschließend werden hinreichende Bedingungen nachgewiesen, um die Konvergenz der einzelnen Verfahren garantieren zu können.

Die entwickelten theoretischen Resultate werden anschließend in numerischen Experimenten praktisch untersucht. Dabei werden insbesondere die Konvergenzraten, der Aufwand für die policy iteration sowie der Nutzen der Verwendung nicht-uniformer Gitter betrachtet.

Preface

In the past decades the Black-Scholes model was the one which helped to price options and to estimate market risks. With certain enhancements it is still the one used most in this area. Financial crisis and collapses unveiled that market risks are not displayed correctly by the model. Within the assumptions of the Black-Scholes model these properties of the market can not be included. Thus, on the one hand extensions to and generalisations of the model have been developed. On the other hand new models have been introduced. The model which is analysed from a numerical point of view in the present thesis is a generalisation of the Black-Scholes model.

In 1995 Avellaneda, Levy, and Paras [ALP95] and Lyons [Lyo95] independently introduced this generalisation. They relaxed the assumption that volatility is constant and instead assumed it to lie in a compact bounded interval. If several assets are modelled also the correlation between two of them can be assumed to take its values in such an interval. Since no additional assumptions like a stochastic process for volatilities and correlation are made the model is also called the Uncertain Volatility model.
The price of an option on assets modelled under these assumptions is then given by a non-linear partial differential equation. This pde is of Hamilton-Jacobi-Bellman type. It was called the Black-Scholes-Barenblatt equation by Avellaneda et al. For options on one as well as on more assets, its structure is very similar to the Black-Scholes equation. The difference lies the diffusion term where volatility occurs. This term is non-linear since volatilities and correlation depend on the second derivatives of the option value.

Non-linear pdes do not have a smooth solution in the classical sense, generally. In the 80's Crandall and Lions [CL83, CIL92] introduced the concept of viscosity solutions. It is constructed via auxiliary functions which enfold the solution and locally approximate its derivatives. Fleming and Soner [FS06] showed that in a financial context the viscosity solution is the meaningful one. Vargiolu [Var01] proved under some assumptions that there exists a unique viscosity solution of the BSB equation. Moreover, he showed that every solution of the BS equation calculated with volatili-

ties and correlation lying in the interval of the Uncertain Volatility model is smaller than the solution of the BSB equation.
To properly solve the BSB equation numerical schemes have to converge to this solution. Barles and Souganidis [BS90] investigated under which conditions this convergence can be guaranteed. They proved that if the pde satisfies a maximum principle, see below (v. i.), and the scheme is consistent, stable, and monotone it converges to the unique viscosity solution of the BSB equation.

In the present thesis we develop a numerical scheme to solve the BSB equation. The focus is put on multi-asset options and on efficiency increase. The task is twofold. On the one hand, a discretisation has to be constructed which possesses all properties to converge to the viscosity solution. On the other hand, algorithms are needed to solve the maximisation problems incorporated in the BSB equation.
A well known scheme is the Kushner-Dupuis scheme [KD01]. Different approaches via a Finite Volume method have been performed by Pooley, Forsyth, and Vetzal [PFV03b], Wang and Forsyth [WF08], and Forsyth and Labahn [FL07] among others. Finite Difference methods have been applied to the one-dimensional Uncertain Volatility model by Heider [Hei10] and Schaeling [Sch10].
The basis for the discretisation in the present thesis are Finite Difference methods. For the two- and three-dimensional Uncertain Volatility model we extend the results already present in the literature and construct a complete algorithm to price options. We prove the convergence of the schemes to the viscosity solution under certain conditions.
In the first chapter of this thesis we give a general introduction to the mathematical framework which is needed for the rest of this thesis. We first consider the Black-Scholes model and its extension the Uncertain Volatility model. Hereafter we give the existence and uniqueness results of Vargiolu and the convergence result of Barles and Souganidis.
The next chapter treats the discretisation of the pde. As representative case the two-dimensional Uncertain Volatility model is examined. The results obtained here for the spatial discretisation either apply for the higher dimensional cases or can easily be transferred to them.
Of special interest is the underlying spatial grid. It influences the monotonicity of the scheme and in which way the occurring optimisation problems can be solved.
The last sections of this chapter focus on the discretisation in time and the pricing of American options.
As last step to solve the BSB equation, optimisation problems have to be solved. Since their structure depends on the spatial grid and the discretisation several cases

have to be analysed. Chapter 3 treats this local optimisation at each grid point. Furthermore, due to the dependence of option value and volatilities and correlation on each other we introduce an iteration presented by Forsyth and Labahn [FL07].
In Chapter 4 we prove the convergence of the considered schemes to the unique viscosity solution of the BSB equation. The results of the one-dimensional schemes are of similar type compared to those given in [Hei10] and [Sch10].
The last chapter concludes with the numerical results for the schemes introduced above. We analyse runtime and speed of convergence as well as efficiency increase due to the use of non-uniform grids.

Munich, October 2013 Albrecht Budke

Contents

List of tables

List of figures

List of algorithms

List of symbols

General symbols

$\partial\mathcal{G}$ . . . . . Discrete or continuous boundary of the set $\mathcal{G}$.

$\mathrm{diag}(v)$ . . . . . $n \times n$-matrix. $\mathrm{diag}(v)_{i,j} = \delta_{i,j} v_i$, $v \in \mathbb{R}^n$.

$\mathcal{L}^n_{BS}$ . . . . . n-dimensional BS differential operator.

$\mathcal{L}^n_{BSB}$ . . . . . n-dimensional BSB differential operator.

Ω_n . . . . . Domain of the transformed BS equation.

$\overline{\Omega}_n$ . . . . . Domain of the untransformed BSB equation.

I_n . . . . . Set of indices, $I_n := \{1, \ldots, n\}$.

J_n . . . . . Set of indices, $J_n := \{0, \ldots, n\}$.

$\mathbb{N}$ . . . . . Set of natural numbers.

$\mathcal{N}_p(x, y)$ . . . . . Standard p-dimensional normal distribution function with mean $x \in \mathbb{R}^p$ and variance $y \in \mathbb{R}^{p \times p}$, $p \in \mathbb{N}$.

$\mathbb{R}^n$ . . . . . n-dimensional set of reell numbers, $n > 0$.

$\mathbb{R}^n_+$ . . . . . n-dimensional set of positive, reell numbers, $n > 0$.

$e_{\overline{j}}$ . . . . . $\overline{j}$-th canonical unit vector.

$\{W_t\}$ . . . . . Standard Wiener process.

Numerical Symbols

$\delta^b_{x_1}$ . . . . . Backward difference quotient for $\frac{\partial}{\partial x_1}$.

$\delta^c_{x_1}$ . . . . . Centered difference quotient for $\frac{\partial}{\partial x_1}$.

$\delta^f_{x_1}$ . . . . . Forward difference quotient for $\frac{\partial}{\partial x_1}$.

$\delta^2_{x_1}$ . . . . . Centered difference quotient for $\frac{\partial^2}{\partial x_1^2}$.

$\delta^+_{x_1x_2}$ . . . . . Difference quotient for $\frac{\partial^2}{\partial x_1 \partial x_2}$.

$\delta^-_{x_1x_2}$ . . . . . Difference quotient for $\frac{\partial^2}{\partial x_1 \partial x_2}$.

$D^{c,i}_{\bar{j},\gamma}\left(v^i, \tau^i, \mathcal{S}^i, \tau^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right)$ Discretisation for time step $\tau^{i-1} \to \tau^i$ from grid $\mathcal{S}^{i-1}$ to $\mathcal{S}^i$ with unknown v^i.

$\mathcal{G}\left(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_n\right)$ . . . . . Grid in Ω_n with L grid points t^i in $[0,T]$ and the spatial grid $\mathcal{S}^i$ at t^i.

τ^i . . . . . i-th grid point in time.

h^i_τ . . . . . Step size at τ^i.

$x^{\nu,i}_\mu$ . . . . . ν-th grid point in μ-direction at τ^i.

$\bar{j}$ . . . . . General lexicographic index for a grid point.

$N(k,l)$ . . . . . Set of indices of grid points adjacent to (x^k_1, x^l_2).

$h^{\nu,i}_\mu$ . . . . . Spatial step size in x_μ-direction at τ^i.

H . . . . . Global maximal spatial step size.

M^i_j . . . . . Number of intervals in j-th spatial direction at τ^i.

M^i . . . . . Number of spatial points at τ^i, $M^i = \prod_{l=1}^{n}(M^i_l + 1)$.

Symbols of models

Θ_n . . . . . Set of controls, $n > 0$.

Γ'_n . . . . . Set of admissible covariance matrices for n assets.

Γ_n . . . . . Set of admissible volatilities for n assets.

$\Xi^d_{i,j}$ . . . . . Lower bound for the ration of the step sizes in spatial direction x_i and x_j, $\Xi^d_{i,j} := \underline{\rho_{i,j}}\frac{\underline{\sigma_i}}{\overline{\sigma_j}}$.

$\Xi^u_{i,j}$ . . . . . Upper bound for the ration of the step sizes in spatial direction x_i and x_j, $\Xi^u_{i,j} := \frac{1}{\underline{\rho_{i,j}}}\frac{\overline{\sigma_i}}{\underline{\sigma_j}}$.

$\rho_{\mu,\nu}$ . . . . . Correlation between the μ-th and the ν-th asset.

$\mathcal{C}_{\mu,\nu}$ . . . . . Set of the admissible correlation for $\rho_{\mu,\nu}$.

$\underline{\rho}_{\mu,\nu}$ . . . . . Lower bound of correlation $\rho_{\mu,\nu}$.

$\overline{\rho_{\mu,\nu}}$ . . . . . Upper bound of correlation $\rho_{\mu,\nu}$.

$\mathcal{D}_2$ . . . . . Discrete set of controls for n = 2.

δ_μ . . . . . Dividend rate of the μ-th asset.

r . . . . . Riskless interest rate.

T . . . . . Maturity of an option.

Ψ_A . . . . . Payoff of an option of type A.

K . . . . . Strike price of an option, $K > 0$.

τ . . . . . Backward time, $\tau := T - t$.

σ_μ . . . . . Volatility of the μ-th asset S_μ.

$\mathcal{V}_\mu$ . . . . . Set of admissible volatilities of the μ-th asset S_μ.

$\underline{\sigma_\mu}$ . . . . . Lower bound of the volatility of the μ-th asset S_μ.

$\overline{\sigma_\mu}$ . . . . . Upper bound of the volatility of the μ-th asset S_μ.

List of abbreviations

General abbreviations

cf. . . . . confer

cp. . . . . compare

i. e. . . . . that is (lat.: id est)

s. . . . . see

s. t. . . . . subject to

t. h. o. . . . . terms of higher order

v. i. . . . . see below (lat.: vide infra)

v. s. . . . . see above (lat.: vide supra)

w. l. o. g. . . . . without loss of generality

Mathematical abbreviations

BDF1 . . . . . 1-step backward differentiation formula

BDF2 . . . . . 2-step backward differentiation formula

BiCGSTAB . . . . . Biconjugate Gradient Stabilized

BS . . . . . Black-Scholes

BSB . . . . . Black-Scholes-Barenblatt

CN . . . . . Crank-Nicolson

EE . . . . . Explicit Euler

FD . . . . . Finite Difference

FV . . . . . Finite Volume

HJB . . . . . Hamilton-Jacobi-Bellman

KD . . . . . Kushner-Dupuis

KKT . . . . . Karush-Kuhn-Tucker

pde . . . . . partial differential equation

1 Introduction

In this chapter we introduce the most basic concepts for the present thesis. In the first section we give an introduction to elementary terms and definitions of modern numerical and stochastical finance. Only those definitions, which are needed for the famous Black-Scholes (BS) model, are brought into focus. We directly address the case of a market consisting of n assets and a money market account. The one-dimensional case is trivially contained. As basic results for the chapters hereafter we introduce the BS equation for European options and the linear complementary problem for American options on n assets each. The section is closed with the introduction of different types of options we want to price in Chapter 5.
In the second section we introduce the non-linear model, which represents the pricing framework for the whole thesis: the Uncertain Volatility model. All necessary definitions and assumptions are given. Finally, we introduce the Black-Scholes-Barenblatt (BSB) equation, while the derivation of the pricing equation itself is omitted.
A general non-linear partial differential equation (pde) does not have to have a smooth solution in the classical sense. Therefore, the viscosity solution has been developed. In the third section this concept is presented in a short introduction. Hereafter, the existence and uniqueness results for the viscosity solution of the BSB equation are given. The last part of the section then focuses on the question which conditions a discrete scheme has to fulfil in order to converge to this viscosity solution.

1.1 The Black-Scholes model

In 1973 Black and Scholes [BS73] and Merton [Mer73] independently developed the model which is now often named the Black-Scholes model after the first two authors. It marks the first important step to model stock market dynamics mathematically. Modern mathematical finance is in many cases based on or at least motivated by this model. Many different modifications and ameliorations have been introduced to improve it. For their work Scholes and Merton received the Nobel Memorial Price in Economic Sciences in 1997.

We assume that the market in the BS model consists of $n \in \mathbb{N}$ assets or stocks $S_1, \ldots, S_n$ and a money market account where money can be borrowed or invested at a constant interest rate $r \geq 0$. The dynamics of the assets are modelled by a stochastic process, which we will introduce in the following.

Definition 1.1 (Wiener process[1])**.**
A stochastic process $\{W_t\}_{0 \leq t \leq T}$ is called Wiener process, if:

1. $W_0 = 0$.
2. W_t *is continuous in time.*
3. *The random variable W_t is normally distributed: $W_t \sim \mathcal{N}_1(0, t)$. That is, it has expected value $\mathbb{E}(W_t) = 0$ and variance $Var(W_t) = t$.*
4. *For $0 \leq t_1 < t_2 \leq t_3 < t_4 \leq T$ the increments $W_{t_1} - W_{t_2}$ and $W_{t_3} - W_{t_4}$ are stochastically independent.*

With the definition of a Wiener process we now define the driving process of the market within the Black-Scholes model.

Definition 1.2 (Geometric Brownian[2] motion)**.**
A stochastic process $\{\overline{S}_t\}_{0 \leq t \leq T} \subset \mathbb{R}^n$ is called a Geometric Brownian motion if it satisfies the stochastic differential equation

$$d\overline{S}_t = a\text{diag}(\overline{S}_t)dt + b\text{diag}(\overline{S}_t d\overline{W}_t,$$

where $\overline{W}_t = (W_{1,t}, \ldots, W_{n,t})^T$ is an n-dimensional Wiener process with stochastically independent scalar Wiener processes $W_{1,t}, \ldots, W_{n,t}$, $a \in \mathbb{R}^n$ a vector, and $b \in \mathbb{R}^{n \times n}$ a matrix.

For a detailed introduction to stochastic processes in finance we refer the reader to the book of Shreve [Shr04].

For each asset S_μ a volatility $\sigma_\mu \in \mathbb{R}_+, \mu \in I_n := \{1, \ldots, n\}$ is known. Furthermore, for two assets S_μ and $S_\nu, \mu \neq \nu$, a correlation $\rho_{\mu,\nu} \in [-1, 1]$ is known. Of course, we have $\rho_{\mu,\mu} = 1, \mu \in I_n$.
The model is based on several assumptions which are restrictive in some ways.

Assumption 1.3.
The market in the BS model is assumed to satisfy the assumptions that

[1] The name is due to the mathematician Norbert Wiener.
[2] The name is due to the botanist Robert Brown who analysed the movement of pollen in water.

- *the interest rate $r \geq 0$ and the dividends $\delta_\mu \geq 0, \mu \in I_n$, are constant*[3],
- *the volatility $\sigma_\mu, \mu \in I_n$ and the correlations $\rho_{\mu,\nu}, \mu \neq \nu, \mu, \nu \in I_n$, are constant,*
- *the market is frictionless that is (i. e.) there are no fees like transaction costs or taxes, every information is accessible for every market participant, every stock can be traded at every time and in every amount, and single transactions do not influence the price of an asset,*
- *the dynamic of each asset is given by Geometric Brownian motion,*
- *there is no arbitrage in the market, i. e. no market participant can expect a positive payment without taking any risk.*

In Assumption 1.3 there obviously are some crucial points that do not match the real world market. For modifications and ameliorations compare the literature mentioned in Chapter 1.2. Especially the assumption of constant volatilities and correlation will be relaxed in this work.

Going on from Assumption 1.3, we introduce the fundamental stochastic differential equation of the BS model. Let $\gamma'_{BS} \in \mathbb{R}^{n\times n}$ be the covariance matrix of the n assets: $(\gamma'_{BS})_{\mu,\nu} = \sigma_\mu \sigma_\nu \rho_{\mu,\nu}$. Since γ'_{BS} is by construction symmetric positive definite, the unique Cholesky decomposition exists. Let $\gamma_{BS}\gamma_{BS}^T = \gamma'_{BS}$.
Then the classical Black-Scholes model for n dividend-paying assets $S_1, \ldots, S_n$ is given by

$$dS_{\mu,t} = (r - \delta_\mu)S_{\mu,t}dt + \sum_{\nu=1}^{n}(\gamma_{BS})_{\mu,\nu}S_{\nu,t}dW_{\nu,t}, \quad \mu \in I_n, \tag{1.1a}$$

$$\Leftrightarrow \quad d\overline{S}_t = (\overline{r} - \overline{\delta})\mathrm{diag}(\overline{S}_t)dt + \gamma_{BS}\mathrm{diag}(\overline{S}_t)d\overline{W}_t, \tag{1.1b}$$

with the vector notation mentioned in Definition 1.2. The vector $\overline{r}$ contains the rate r in each component and the vector $\overline{\delta}$ contains the specific dividend rates δ_μ at the relevant component. We say that the process $\{\overline{S}_t\} \subset \mathbb{R}^n_+$ has drift $(\overline{r} - \overline{\delta})\mathrm{diag}(\overline{S}_t)$ and diffusion $\gamma_{BS}\mathrm{diag}(\overline{S}_t)$ where $\mathrm{diag}(v)$ is a diagonal matrix with the vector v on its diagonal. By construction of (1.1) the process $S_{\mu,t}$ and $S_{\nu,t}$ are correlated with $\rho_{\mu,\nu}$, see for example [Shr04].

The value function of the derivative, we want to price, is denoted by $V(t, \overline{S}_t)$ for $\overline{S}_t \in \mathbb{R}^n_+, t \in [0, T]$. At $t = T$ the value is given by the payoff function $\Psi_A(\overline{S}_T)$,

[3] In the original model the assets were also assumed to be dividend-less. But dividends can easily be included if they are assumed to be at a constant rate.

where Ψ_A is the payoff of the option of type A. Our aim is now to determine the current ($t = 0$) value $V(t, \overline{S}_t)$ of a European style option . That is: only at $t = T$ the owner of the option gets $V(T, \overline{S}_T) = \Psi_A(S_T)$. The option cannot be exercised during its runtime. For different choices of Ψ, see (s.) below.
By constructing a self-financing portfolio whose value in $t = T$ is equal to the derivative's value, we obtain a pde for V by using the Lemma of Itô , cp. [Shr04], and the no-arbitrage assumption. The equation is the famous Black-Scholes equation for n assets:

$$\begin{aligned} \frac{\partial V}{\partial t} + \frac{1}{2}\mathrm{tr}\left(\left(\frac{\partial^2 V}{\partial S_\mu \partial S_\nu}\right)_{\mu,\nu} \mathrm{diag}\left(\overline{S}\right)\sigma\sigma^T \mathrm{diag}(\overline{S})\right) \\ + \sum_{\nu=1}^{n}(r-\delta_\nu)S_\nu\frac{\partial V}{\partial S_\nu} - rV = 0, \quad 0 \le t < T, \\ V(\overline{S}_T, T) = \Psi(\overline{S}_T). \end{aligned} \tag{1.2}$$

The interested reader is referred to the original works of Black and Scholes [BS73] and Merton [Mer73] and the books of Shreve [Shr04] and Levy [Lev04] to get a detailed insight in the derivation of the BS equation (1.2).

For American options it has to be taken into account that at every time $0 \le t_{ex} \le T$ the owner of the option can exercise at t_{ex} and receives the payoff $\Psi_\cdot(\overline{S}_{t_{ex}})$. The value of such an option is determined by a linear complementary problem. If we write

$$\mathcal{L}^n_{BS} := \frac{1}{2}\mathrm{tr}\left(\left(\frac{\partial^2 V}{\partial S_\mu \partial S_\nu}\right)_{\mu,\nu} \mathrm{diag}(\overline{S})\sigma\sigma^T \mathrm{diag}(\overline{S})\right) + \sum_{\nu=1}^{n}(r-\delta_\nu)S_\nu\frac{\partial V}{\partial S_\nu} - rV,$$

it is given by:

$$\begin{aligned} \frac{\partial V}{\partial t} + \mathcal{L}^n_{BS}(V) &\le 0 \\ \left(\frac{\partial V}{\partial t} + \mathcal{L}^n_{BS}(V)\right)(V - \Psi) &= 0 \\ V &\ge \Psi \end{aligned} \tag{1.3}$$

The derivation of (1.3) can be found in the article of Jaillet, Lamberton, and Lapeyre [JLL90] and the book of Lamberton and Lapeyre [LL08]. The approach introduced above is not the only one to price American options via pdes, see for example the algorithm of Brennan and Schwartz [BS77]. But for options with possibly non-convex payoffs or those depending on several assets it is the most workable. For the numerical solution of the complementary problem see Section 2.4.

In the remainder of this section we introduce different types of options, some of which we will analyse in the following chapters. We will consider standard options and several ones of exotic type like butterfly options. For an overview over more exotic payoffs we confer to the books of Deutsch [Deu09] and Wilmott, Dewynne, and Howison [WDH00].

The simplest options on one asset are Vanilla options, the European put, and the European call. They are the standard instruments at a derivative market and are for example used to hedge a position in a stock or to speculate on growing stock prices. Their payoff for a given strike price K is given via

$$\Psi_P(S) := \max\{S-K,0\} \quad \text{and} \quad \Psi_C(S) := \max\{K-S,0\}.$$

For ease of notation we will write $(S)^+ := \max\{S,0\}$.
Furthermore, we will consider some options which are not that standard. The first ones are the butterfly spread and the digital call respectively put. Their payoff functions are defined as follows

$$\Psi_{BfS}(S) := \begin{cases} S-K+\alpha, & K-\alpha \le S < K \\ K+\alpha-S, & K \le S \le K+\alpha \\ 0, & \text{otherwise} \end{cases}, \quad \alpha \in [0,K],$$

$$\Psi_{DiC}(S) := \begin{cases} 0, & S < K \\ 1, & S \ge K \end{cases}, \quad \text{and} \quad \Psi_{DiP}(S) := \begin{cases} 1, & S < K \\ 0, & S \ge K \end{cases}.$$

In Figure 1.1 two payoff functions are plotted.

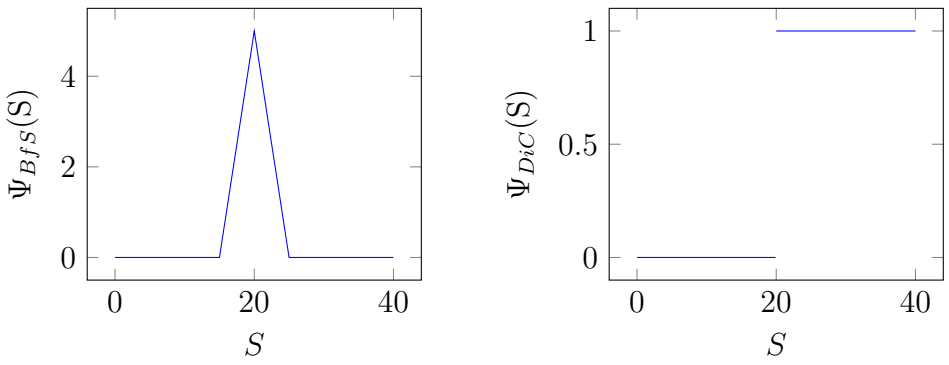

Fig. 1.1: *Payoff of a butterfly spread (left) and a digital call (right) with* $K = 20$ *and* $\alpha = 5$.

The last class of exotic options are barrier options. The value directly depends on the realised development of the stock price. If the price oversteps (up) or undercuts

(down) a predefined barrier B, it comes into existence (in) or its existence is extinguished (out). By combining these four characteristics and the distinction between put and call, there are eight possible barrier options. In Figure 1.2 the payoff of an up-in call and a down-out put are diagramed.

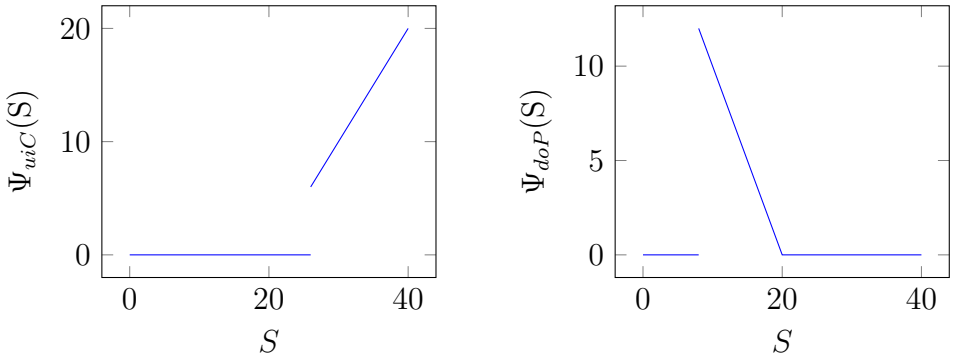

Fig. 1.2: *Payoff of a (knocked in) up-in call (left) with $B = 26$ and a down-out put (not knocked out) (right) with $B = 8$. It is $K = 20$.*

The payoffs of these two options are given by

$$\Psi_{uiC}(S) := \begin{cases} (S-K)^+, & \exists t \in [0,T] : S_t \geq B \\ 0, & otherwise \end{cases}$$

and

$$\Psi_{doP}(S) := \begin{cases} 0, & \exists t \in [0,T] : S_t \leq B \\ (K-S)^+, & otherwise \end{cases}.$$

For the remainder of the six other barrier options see for example the book of Wilmott, Dewynne, and Howison [WDH00].
In opposite to the Vanilla options, the payoffs of the latter ones are not convex and for the digital and the barrier options not even continuous. These properties will make them interesting for the non-linear model we are going to analyse.

There are four different standard types of options on several assets: minimum and maximum options and geometric and arithmetic average options. Each of them can be constructed as a put as well as a call.
For a given strike price K we have

$$\Psi_{MinP}(S_1,\ldots,S_n) := (K - \min(S_1,\ldots,S_n))^+ \quad \text{and}$$
$$\Psi_{MaxC}(S_1,\ldots,S_n) := (\max(S_1,\ldots,S_n) - K)^+$$

for a minimum put and a maximum call, s. also Figure 1.3.
The payoffs of the geometric average put and the arithmetic-average call are given by

$$\Psi_{GeoP}(S_1,\dots,S_n) := \left(K - \left(\prod_{i=1}^{n} S_i^n\right)^{\frac{1}{n}}\right)^+ \quad \text{and}$$
$$\Psi_{AriC}(S_1,\dots,S_n) := \left(\frac{1}{n}\sum_{i=1}^{n} S_i - K\right)^+$$

for a fixed strike price K. For all payoffs a capped version can be constructed by defining

$$\Psi_{cA}(S_1,\dots,S_n) = \min\{\Psi_A(S_1,\dots,S_n),\alpha\}, \quad \alpha \in \mathbb{R}_+,$$

s. also Figure 1.3.

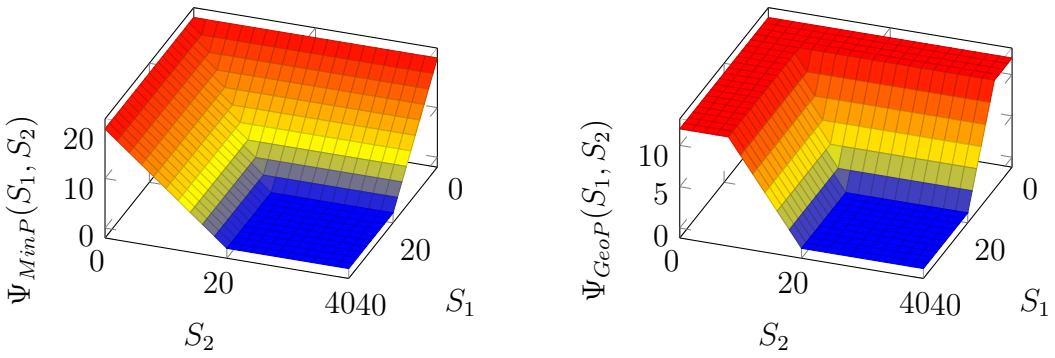

Fig. 1.3: *Payoff of a minimum put (left) and a capped minimum put (right) with $K = 20$, $\alpha = 12$. (Notice the rotation of the axis.)*

The remaining payoffs can be constructed in an analogous way.
Another type of basket option is the extension of the butterfly spread for two underlying assets. For a strike price K and a constant $\alpha \in \mathbb{R}_+$ its payoff is given by:

$$\Psi_{BfS}(S_1,S_2) := (\max(S_1,S_2) - K + \alpha)^+ + (\max(S_1,S_2) - K - \alpha)^+ \\ -2(\max(S_1,S_2) - K)^+$$

which is plotted in Figure 1.4.

All introduced payoffs share the property that they are not differentiable. For the barrier options the payoff is not even continuous.

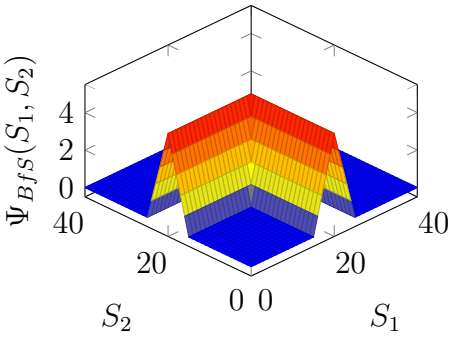

Fig. 1.4: *Payoff of a butterfly spread on two assets with* $K = 20$ *and* $\alpha = 5$.

1.2 Uncertain Volatility model

It is well known that the model developed by Black, Scholes [BS73], and Merton [Mer73] is based on assumptions that do not take into account some of the basic dynamics of the market. In the model volatility is assumed to be independent of time, asset value, etc. Instead it is a commonly known effect that the implied volatility calculated of market data is not constant. But it varies depending among others on the strike price which gives the so called volatility smile, confer (cf.) [CF02, Shi93]. Furthermore, modelling correlation is in general a difficult task. It often is only modelled for different asset classes, cf. [LS95].
Therefore, many different approaches have been developed in order to include non-constant volatility. One important example is the Heston model [Hes93]. Here the dynamic of volatility itself is modelled by an Ornstein-Uhlenbeck process. Another example is the ansatz of Dupire [Dup94]. He modelled volatility as a deterministic function of strike and maturity of derivatives traded at the market. The function itself is the solution of a pde.

In this chapter we introduce the Uncertain Volatility model developed independently by Avellaneda, Levy and Paras [ALP95] and Lyons [Lyo95] in 1995. Here no assumptions about a driving process or a deterministic function for the volatility are made. The only point, we will assume, is that each volatility and correlation lies somewhere in between given specific upper and lower bounds, respectively. Considered this way, the model is quite general. In the resulting pricing equation, the actual volatilities and correlations depend on different derivatives of the value function.
Since classical solutions of the derived pricing equations do not have to exist, we shortly introduce the concept of viscosity solution. Finally, we remind the reader which requirements schemes for the discrete pricing equation have to fulfil in order to converge to the desired solution.

Assume that there are given n assets $S_1, \ldots, S_n$ with volatility $\sigma_\mu, \mu \in I_n$, and correlations $\rho_{\mu,\nu}, \mu, \nu \in I_n$, between two assets S_μ and S_ν. Of course $\rho_{\mu,\mu} = 1, \mu \in I_n$. For all of the above parameters no assumptions on driving processes or other dependencies are made. We only assume that we know intervals in which the volatilities and correlations can take values. Let $\mathcal{V}_\mu := [\underline{\sigma_\mu}, \overline{\sigma_\mu}] \subset \mathbb{R}_+, \mu \in I_n$, and $\mathcal{C}_{\mu,\nu} := [\underline{\rho_{\mu,\nu}}, \overline{\rho_{\mu,\nu}}] \subset [-1,1], \mu, \nu \in I_n$, be the intervals for the volatilities and the correlations, respectively:

$$\sigma_\mu \in \mathcal{V}_\mu, \quad \mu \in I_n, \quad \text{and} \quad \rho_{\mu,\nu} \in \mathcal{C}_{\mu,\nu}, \quad \mu, \nu \in I_n \tag{1.4}$$

To handle all possible combinations of volatilities and correlations, we will define the so called set of admissible covariance matrices.

Definition 1.4 (set of admissible correlation matrices)**.**
The set of admissible covariance matrices Γ'_n is defined as the set of all correlation matrices which is obtained by choosing $\sigma_\mu \in \mathcal{V}_\mu, \mu \in I_n$, and $\rho_{\mu,\nu} \in \mathcal{C}_{\mu,\nu}, \mu, \nu \in I_n$.

$$\Gamma'_n := \left\{ \gamma' \in [-1,1]^{n \times n} \mid \gamma' = (\sigma_\mu \sigma_\nu \rho_{\mu,\nu})_{\mu,\nu}, \sigma_\mu \in \mathcal{V}_\mu, \rho_{\mu,\nu} \in \mathcal{C}_{\mu,\nu}, \mu, \nu \in I_n \right\} \tag{1.5}$$

For each element $\gamma' \in \Gamma'_n$ we now define the set of admissible volatilities as the set of Cholesky decompositions of all admissible correlation matrices. Since these are positive definite [Kwo08] by construction they can be decomposed in the required way [GvL96].

Definition 1.5 (set of admissible volatilities)**.**
The set of admissible volatilities is defined as

$$\Gamma_n := \left\{ \gamma \in \mathbb{R}^{n \times n} \mid \gamma \gamma^T = \gamma', \gamma' \in \Gamma'_n \right\}. \tag{1.6}$$

Starting with these preliminaries, we can now introduce the Uncertain Volatility model. Finally, we receive a non-linear pde which can be used to value financial derivatives in this model.
Let $\{\sigma_t\} = \{(\sigma_t^{\mu,\nu})\}, \mu, \nu = 1, \ldots, n, t \geq 0$, be an arbitrary stochastic process defined over $\mathcal{A}(\Gamma)$ which is the smallest sigma algebra (cf. [Fel08]) defined by the set of admissible volatilities Γ_n. We model the dynamics of the n assets by Geometric Brownian motion where $\{\sigma_t\}$ represents volatility. Let $\{W_{\mu,t}\}, \mu \in I_n$, be n stochastic independent Wiener processes [Fel08]. The single process for the μ-th asset will be denoted by $S_{\mu,t}, \mu \in I_n$. Then our model is given by the stochastic differential

equations

$$d\overline{S}_t = r\overline{S}_t dt + \operatorname{diag}(\overline{S}_t)\sigma_t d\overline{W}_t, \tag{1.7a}$$

$$\Leftrightarrow \quad dS_{\mu,t} = rS_{\mu,t}dt + \sum_{\nu=1}^{n} \sigma_t^{\mu,\nu} S_{\nu,t} dW_{\nu,t}, \quad \mu \in I_n. \tag{1.7b}$$

Up to the volatility $\{\sigma_t\}$, equation (1.7) is identical to the multi-dimensional BS model.

Using a super-replication approach via delta-hedging (see also Subsection 1.3.1) and the lemma of Itô, compare (cp.) [Shr04], the following pricing equation can be derived for the value of a European option during time [ALP95, Lyo95].

Problem 1.6 (Valuing European options)**.**
Solve

$$\begin{aligned} &\frac{\partial V}{\partial t} + \frac{1}{2}\max_{\sigma\in\Gamma_n}\left\{\operatorname{tr}\left(\left(\frac{\partial^2 V}{\partial S_\mu \partial S_\nu}\right)_{\mu,\nu} \operatorname{diag}(\overline{S})\sigma\sigma^T\operatorname{diag}(\overline{S})\right)\right\} \\ &\qquad\qquad + \sum_{\nu=1}^{n}(r-\delta_\nu)S_\nu\frac{\partial V}{\partial S_\nu} - rV = 0, \quad 0 \le t < T, \\ &V(T,\overline{S}_T) = \Psi(\overline{S}_T) \end{aligned} \tag{BSB_n}$$

on $\overline{\Omega}_n := [0,T]\times\mathbb{R}^n_+$.

In the literature the equation (BSB_n) is also known as the Black-Scholes-Barenblatt equation[4]. It is a non-linear partial differential equation of Hamilton-Jacobi-Bellman type. For results on the existence and the uniqueness of the viscosity solution of (BSB_n) and its properties see Section 1.3.1.

Example 1.7 (Black-Scholes-Barenblatt equation for one asset)**.**
In the case of one asset the volatility σ_1 *is assumed to lie in the interval* $[\underline{\sigma_1}, \overline{\sigma_1}]$. *The pricing equation for the option value* $V(t,S_{1,t})$, $(t,S_{1,t}) \in \Omega_1$, *for a European vanilla option is then determined by*

$$\begin{aligned} &\frac{\partial V}{\partial t} + \frac{1}{2}S_1^2\sigma_1^2\left(\frac{\partial^2 V}{\partial S_1^2}\right)\frac{\partial^2 V}{\partial S_1^2} + (r-\delta_1)S_1\frac{\partial V}{\partial S_1} - rV = 0, \quad 0 \le t < T \\ &V(T,S_{1,T}) = \Psi(S_{1,T}) \end{aligned} \tag{BSB_1}$$

[4]This denomination stems from Avellaneda, Levy and Lyons [ALP95] who chose the name after the physicist G.I. Barenblatt. He introduced an equation with similar structure compared to the Black-Scholes-Barenblatt-equation (BSB_n) [Bar96].

where

$$\sigma_1(x) := \begin{cases} \overline{\sigma_1}, & \operatorname{sign}(x) > 0 \\ \underline{\sigma_1}, & \text{else} \end{cases}$$

holds, cf. [ALP95].
Up to the volatility function $\sigma(\cdot)$ the structure of equation (BSB_1) is strongly reminiscent to the BS equation.

For options of American style we assume that the value is given by a complementary problem [JLL90, WDH00].

Problem 1.8 (Valuing American options)**.**
Solve

$$\begin{aligned} \frac{\partial V}{\partial t} - F_{BSB}\left(t, x, V, D_x V, D_x^2 V\right) &\leq 0, \\ V - \Psi &\geq 0, \\ \left(\frac{\partial V}{\partial t} - F_{BSB}\left(t, x, V, D_x V, D_x^2 V\right)\right)(V - \Psi) &= 0, \end{aligned}$$

Here

$$\begin{aligned} F_{BSB_n}(t, x, u, a, A) := -\frac{1}{2} \max_{\sigma \in \Gamma} \left\{ \operatorname{tr}\left(A \operatorname{diag}(x) \sigma \sigma^T \operatorname{diag}(x)\right) \right\} \\ - \sum_{\nu=1}^{n} (r - \delta_\nu) x_\nu a_\nu + ru, \end{aligned}$$

$D_x V$, and $D_x^2 V$ respectively denote the gradient and the Hesse matrix of V with respect to x.

Remark 1.9.
For the linear version of Problem 1.8 it is a known result that the price of an American option is given by the linear complementary problem, s. for example the paper of Jaillet, Lamberton, and Lapeyre [JLL90]. For the BSB model we do not know about a similar result and thus we have to assume that the non-linear complementary problem is the right representation.

1.3 Viscosity solutions

Non-linear partial differential equation do not have a classical solution in the sense that it is continuously differentiable. Therefore, the concept of viscosity solution

has been developed in the 90s.
This section is subdivided into three smaller parts. In the first one we give a short introduction to viscosity solutions. Only the principals and the results we will need in the sections hereafter are reviewed. The second part treats properties like existence and uniqueness of the viscosity solutions for the BSB equation. It possesses a nice financial interpretation as we will see later on.
In the third part we consider the viscosity solution from a numerical point of view. To ensure convergence of a given scheme to this solution, it has to satisfy certain assumptions. The main result in this field stems from Barles and Souganidis [BS90]. We introduce the necessary notation for the schemes and finally state their result.

The following example illustrates, that non-linear pdes do not have a solution in the classical sense in general, cf. [Eva10].

Example 1.10 ([Eva10, FS06])**.**
We are looking for a solution of the non-linear pde of Eikonal[5] *type*

$$|u'(t)| = 1, \quad t \in (-1,1), \quad u(-1) = u(1) = 0 \tag{1.8}$$

If we assume a classical solution to exist, then by the mean value theorem of differential calculus, there has to be a $\xi \in (-1,1)$ *for which* $u'(\xi) = 0$ *holds true. Thus, there is no classical solution.*

In order to circumvent this fact, the concept of viscosity solution has been developed by Crandall and Lions in the 90s, cf. [CL83]. A nice survey article on this subject has been published by Crandall, Ishii, and Lions [CIL92]. More basic information can be found in the books of Bardi and Dolcetta [BD97], Kushner and Dupuis [KD01], and Fleming and Soner [FS06]. Especially in the financial context the viscosity solution is a meaningful solution [DL97, FS06, Chapter X].
The idea of viscosity solutions is to construct the solution of a pde via test functions. Loosely speaking, the viscosity solution itself will be enfolded by these from above and below. The properties they possess can be transferred to the viscosity solution of the pde.

Consider the general pde

$$\frac{\partial u}{\partial t} - F\left(t, x, u, D_x u, D_x^2 u\right) = 0, \quad (t,x) \in \Omega \tag{1.9}$$

on a compact domain Ω where F is continuous and $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times S(n) \to \mathbb{R}$

[5] Eikonal equations are a special type of non-linear pdes which are for example used in physics.

and $S(n) \subset \mathbb{R}^{n\times n}$ is the set of symmetric matrices.
We now introduce some notation which will be needed to formulate the definition of the viscosity solution.

Notation 1.11.
Let $A = (a_{i,j})_{i,j}, B = (b_{i,j})_{i,j} \in \mathbb{R}^{n\times n}, i, j \in I_n$, be matrices. We write

$$A > B,\ A \geq B,\ A < B,\ A \leq B$$

if for any pair of indices (i, j) the condition $a_{i,j} > b_{i,j}$, $a_{i,j} \geq b_{i,j}$, $a_{i,j} < b_{i,j}$, $a_{i,j} \leq b_{i,j}$ holds true, respectively.
The same notation is used for vectors.

In order to give a meaningful definition of viscosity solutions, F should satisfy the following assumptions, cp. [CIL92].

Assumption 1.12.
Assume that for every $u, v \in \mathbb{R}$, $u \leq v$,

$$F(t, x, u, a, A) \geq F(t, x, v, a, A), \tag{1.10}$$

holds true where $(t, x) \in \Omega$, $a \in \mathbb{R}^n$, and $A \in \mathbb{R}^{n\times n}$.
Furthermore, assume that for every $A, B \in \mathbb{R}^{n\times n}$, $A \leq B$,

$$F(t, x, u, a, A) \leq F(t, x, u, a, B), \tag{1.11}$$

where $(t, x) \in \Omega$, $u, a \in \mathbb{R}^n$, is satisfied.

If F fulfils (1.10), it is called *proper* and *elliptic* if (1.11) is satisfied. Assumption 1.12 as a whole is called *monotonicity assumption.*

Remark 1.13.
In the publication [CIL92] these assumptions are formulated for the elliptic case

$$F(x, u, D_x u, D_x^2 u) = 0.$$

The parabolic case (1.9) can be transformed to this one by considering the $n + 1$-dimensional problem with state variable $(t, x) \in \mathbb{R}^{n+1}$, cf. [Lio83, CIL92].

We now define the viscosity solution of the pde (1.9) in two steps. First, we define viscosity sub- and supersolutions in

Definition 1.14 (viscosity sub- and supersolution of a pde)**.**
Let a pde of the form (1.9) be given. Let F satisfy Assumption 1.12. A function v is called viscosity supersolution of equation (1.9) if for all test functions $u \in C^{1,2}(\Omega)$, where $(x^*, t^*) \in \Omega$ is a minimum of $v - u$, the inequality

$$\frac{\partial}{\partial t}u(t^*, x^*) - F\left(t^*, x^*, u(t^*, x^*), D_x u(t^*, x^*), D_x^2 u(t^*, x^*)\right) \geq 0 \tag{1.12}$$

is satisfied.
A function is called viscosity subsolution of equation (1.9) if we replace "minimum" by "maximum" and "$\geq$" by "$\leq$" in (1.12) of the above definition.

Now, we can define the viscosity solution for pde of type (1.9).

Definition 1.15 (viscosity solution of a pde)**.**
A function v is called a viscosity solution of the pde (1.9) if it is both a viscosity sub- and a viscosity supersolution.

Example 1.16 (Continuation of Example 1.10)**.**
Although equation (1.8) has no classical solution, $v(t) = 1 - |t|$ is a viscosity solution. For all $t \in [-1, 1]\backslash\{0\}$ v is differentiable and solves (1.8). Thus, for every test function $u \in C^1([-1, 1])$ where $v - u$ has a maximum / minimum at $t_0 \in [-1, 1]\backslash\{0\}$ the inequality (1.12) holds true with equality.
The interesting point is $t_0 = 0$. Let t_0 be a maximum of $v - u$ with $u(t_0) = v(t_0)$. We have $-|x| \leq v(x) - v(0) \leq u(x) - u(0)$. By this inequality we obtain $|u'(t_0)| \leq 1$ in the limit. Therefore, v is a subsolution of (1.8) for all $t \in [-1, 1]$. On the other hand, if we assume that t_0 is a minimum, we do not obtain the same limits for the derivative from the right and the left, respectively. This means that there are no test functions touching v from below and v trivially is a viscosity supersolution for all $t \in [-1, 1]$. Collectively, v is a viscosity solution of equation (1.8).

Now that we have defined the viscosity solution for the parabolic pde (1.9) the interesting question is what types of equations satisfy Assumption 1.12. In the publication [CIL92] a couple of examples are given. One of them is a general Hamilton-Jacobi-Bellman (HJB) pde which includes the BSB equation (BSB_n).

1.3.1 Viscosity solution of the Black-Scholes-Barenblatt equation

First we state that the BSB equation (BSB_n) satisfies Assumption 1.12. Then we will restate a result proven by Vargiolu [Var01].

Lemma 1.1.
The Black-Scholes-Barenblatt equation (BSB_n) satisfies

1. *inequality* (1.10) *if* $r \geq 0$ *holds true and*
2. *inequality* (1.11) *if* $\sigma\sigma^T$ *is positive definite.*

Thus, Assumption 1.12 is satisfied if 1. and 2. hold.

The proof of the lemma is left to the reader.

Remark 1.17.
The converse of statement 1.1. in Lemma 1.1 is also true [ALP95].

Remark 1.18.
In an economical context the first statement of Lemma 1.1 is negligible in a normal economical environment. On the contrary, the second one is crucial. The set of admissible correlation matrices Γ' *(s. Definition 1.4) also could contain matrices where* $\gamma' \geq 0$*, is violated. Thus, the Black-Scholes-Barenblatt operator is not elliptic for all choices of correlation matrices.*
To ensure the operator to be elliptic and thus have a well defined concept of viscosity solution, we will restrict ourselves to positive definite covariance matrices in the following.

The remaining open question for the rest of this section is: when does a viscosity solution of the Black-Scholes-Barenblatt equation (BSB_n) exist? One answer to this question is the result proven by Vargiolu [Var01].

The following assumption summarises the necessary properties of model (1.7) which have to be fulfilled to apply the results of Vargiolu.

Assumption 1.19.
As in the stochastic process (1.7), let $f(\overline{S}, \sigma) = r\overline{S}$ *and* $g(\overline{S}, \sigma) = \mathrm{diag}(\overline{S})\sigma$ *on* $\mathbb{R}^n \times \Gamma_n$ *denote the convection and the diffusion terms, respectively.*
It is assumed that

1. *f, g are continuous,*
2. *for $\sigma \in \Gamma_n$ fix $f(\cdot, \sigma)$, $g(\cdot, \sigma)$ are continuous differentiable,*
3. *there are constants c_1, $c_2 \in \mathbb{R}_+$ so that $|D_{\overline{S}} f|, |D_{\overline{S}} g| \leq c_1$ and $|f(0,\sigma)| + |g(0,\sigma)| < c_2$,*
4. *if $\tilde{\Psi}$ is a continuous extension of Ψ on $\mathbb{R}^n$, then $|\tilde{\Psi}(\overline{S})| \leq c_3(1 + |\overline{S}|^k)$ for $c_3 \in \mathbb{R}_+, k > 0$.*

It is clear that Assumption 1.19 is satisfied by model (1.7). We now state the existence and uniqueness result for the viscosity solution of equation (BSB_n).

Theorem 1.2.
Let the set of admissible volatilities Γ_n be compact. Furthermore, let Assumption 1.19 hold and let the payoff $\Psi(\overline{S})$ be continuous on $\mathbb{R}^n$.
Then there exists a unique viscosity solution $V(\overline{S}, t)$ of the Black-Scholes-Barenblatt equation (BSB_n) on $[0,T] \times \mathbb{R}^n_+$ with $V(\overline{S}, T) = \Psi(\overline{S})$.

For further details on how the value function $V(\overline{S}, t)$ can be constructed via delta-hedging arguments compare the paper of Vargiolu.

As already mentioned earlier the viscosity solution of the BSB equation is constructed via a super-replication strategy. It is impossible to exactly replicate the option value by a hedging portfolio, since the market is incomplete due to the unknown volatility.
As for the BS model, a self-financing portfolio $\Pi_t(m_t, D_S V(S,t))$ consisting of a certain number of units m_t in the money market account and $\left(D_{\overline{S}} V(\overline{S}, t)\right)_\mu$ units of asset $S_\mu, \mu \in I_n$, is constructed at every time $t \in [0,T]$. Here $V(\overline{S}, t)$ denotes the viscosity solution of equation (BSB_n).
The following theorem has been proven by Romagnoli and Vargiolu [RV00].

Theorem 1.3.
If $V(\overline{S}, t) \in C^{1,2}(\mathbb{R}^n_+ \times [0,T])$ is a solution of the Black-Scholes-Barenblatt equation (BSB_n), then the above strategy is a super-hedging strategy. That is $\Pi_t \geq V(\overline{S}, t)$, for all $t \in [0,T]$. Furthermore, there is no super-hedging strategy for a $V'(\overline{S}, t)$ solving (BSB_n) with $V' < V$.

The theorem above shows that the viscosity solution of the BSB equation is the minimal super-hedging strategy. At $t = 0$ the super-hedging portfolio Π_t opens up the possibility to make arbitrage by holding a short position in the option and buying the

portfolio. At time $t = T$ an investor would receive the amount $P_T - V(\overline{S}_T, T) \geq 0$. Thus, the value $V(\overline{S}, t), t \in [0, T]$ gives an upper bound for arbitrage free prices of the contingent claim.
Similarly, if we solve the BSB equation with a minimisation instead of a maximisation, we would obtain a value giving a lower bound for the arbitrage-free prices of the contingent claim.
Summing up, all option prices $V_{BS}(\overline{S}, t)$ calculated in the BS model with a fixed correlation matrix $\gamma \in \Gamma_n$ lie in between the values $V_{BSB}^{min}(\overline{S}, t)$ and $V_{BSB}^{max}(\overline{S}, t)$ calculated in the BSB model with minimisation and maximisation, respectively.

It has also been proven in the publication of Romagnoli and Vargiolu that for the log-price version of the Black-Scholes-Barenblatt equation and for regular $\gamma \in \Gamma_n$ the smoothness conditions of Theorem 1.3 are fulfilled, cp. [RV00, Theorem 5]. The transformed Black-Scholes-Barenblatt equation is given by

$$\frac{\partial v}{\partial \tau} - \frac{1}{2} \max_{\substack{\sigma_\mu \in \mathcal{V}_\mu \\ \rho_{\mu,\nu} \in \mathcal{C}_{\mu,\nu} \\ \mu,\nu \in I_n}} \left\{ \sum_{\mu=1}^{n} \sigma_\mu^2 \left(\frac{\partial^2 v}{\partial x_\mu^2} - \frac{\partial v}{\partial x_\mu} \right) + \sum_{\substack{\mu,\nu=1 \\ \mu \neq \nu}}^{n} 2\sigma_\mu \sigma_\nu \rho_{\mu,\nu} \frac{\partial^2 v}{\partial x_\mu \partial x_\nu} \right\} - \sum_{\mu=1}^{n} (r - \delta_\mu) \frac{\partial v}{\partial x_\mu} + rv = 0,$$

for details see Section 2.1. Here $\tau = T - t$ denotes backward time.
Thus, under the assumption that the correlation matrices are positive definite, we always will assume this property.

1.3.2 Numerical convergence to the viscosity solution

To solve the BSB equation (BSB_n) numerically for a given n, we transfer it into a discretisation scheme which is solved instead. In this section we examine abstractly under which conditions convergence of the discrete solution to the continuous is given if the scheme is refined. The findings presented here serve as basis for all following convergence results. Otherwise we cannot guarantee convergence form a theoretical point of view. Practical examples where certain schemes do not converge to the viscosity solution can be found in [Poo03]. On the other hand, there are schemes whose convergence cannot be proven but nevertheless seem to converge [Hei09].

Additionally to the classical conditions consistency and stability, an important property is the monotonicity of the scheme (s. Definition 1.23).

Let us consider a general pde of the form (1.9)

$$\frac{\partial v}{\partial t} - F\left(t, x, v, D_x v, D_x^2 v\right) = 0, \quad (t,x) \in \Omega \tag{1.13}$$

and assume that a unique viscosity solution of this equation exists.
The discretisation of this pde is constructed on a grid $\mathcal{G}$. In order to properly denote a grid, we use

Notation 1.20 ($(n+1)$-dimensional grid)**.**
Consider a given space Ω_n. Let $L \in \mathbb{N}$ be the number of grid points in time. Each time level is denoted by $t^i, i \in J_L$, where $J_n := \{0, \ldots, n\}$. At every time level a specific spatial grid is given. For each the number of nodes is $M^0, \ldots, M^L \in \mathbb{N}$, respectively. For the spatial grid points we use the notation $\overline{x}^i_{\overline{j}} \in \mathbb{R}^n, \overline{j} \in J_{M^i}, i \in J_L$. For a regular grid in several spatial dimensions the index $\overline{j}$ is the lexicographical index, cp. also Notation 2.5. If $M^i_1, \ldots, M^i_n \in \mathbb{N}$ denote the number of spatial grid points at time level i for the different spatial directions, we have $M^i = \prod_{j=1}^n M^i_j$.
By $\mathcal{S}^i := \left\{x^i_j \mid j \in J_{M^i}\right\} \subset \mathbb{R}^n$ we denote the set of all grid points in space at time level i.
The complete, possibly non-uniform grid in Ω_n is then denoted by

$$\mathcal{G}\left(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_n\right) = \{(t^i, x^i_j) \in [0,T] \times \mathbb{R}^n \mid j \in J_{M^i}, i \in J_L\}.$$

By Δt and H we denote the maximal step width in time and space, respectively.
If the spatial grid does not change across time levels, i. e. $M^0 = \ldots = M^L$, we denote the grid by $\mathcal{G}\left(L, \{t^i\}, \mathcal{S}, \Omega_n\right)$ and only write M for ease of notation. In the same way we omit the superindex i if the structure of the grid is time independent.

Let a grid $\mathcal{G}\left(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_n\right)$ be given on which we solve the discretised pde (1.13). Let one time step from t^{i-1} to t^i be described by the discrete equation

$$D^i\left(v^i; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right) = 0 \tag{1.14}$$

where $v^i \in \mathbb{R}^{M^i}$ is a vector with the entries $v^i_{\overline{j}}$ denoting the approximation to $V(t^i, x^i_{\overline{j}}), \overline{j} \in J_{M^i}$ on the grid $\mathcal{S}^i$.

Example 1.21.
One time step of the Explicit Euler (EE) scheme on a uniform grid for the pde (1.9) has the form

$$D^i\left(v^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right) = \\ v^i - v^{i-1} - (t^i - t^{i-1})\left(F_{\overline{j}}(t^{i-1}, x^{i-1}_{\overline{j}}, v^{i-1}, D^x v^{i-1}_{\overline{j}}, D^2_x v^{i-1}_{\overline{j}})\right)_{\overline{j}\in J_M}$$

where $D_x v^{i-1}_{\overline{j}}$ and $D^2_x v^{i-1}_{\overline{j}}$ denote discrete approximations to the first respectively second derivative at the j-th node of the grid at time level t^i with respect to x.

In order to properly define the monotonicity of a discrete scheme, we will need some more notation for the discretisation (1.14).

Notation 1.22.
Let $x \in \mathbb{R}^n$ be a vector. Then by $[x]_i$ we denote the vector in $\mathbb{R}^n$ where the i-th component of x is set to zero and the remaining ones stay unchanged: $([x]_i)_j = x_j, j \in I_n, j \neq i$ and $([x]_i)_i = 0$.
For all time levels $i \in I_L$ at every grid point $(t^i, x^i_{\overline{j}}) \in \mathcal{G}(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_n), \overline{j} \in J_{M^i}$, we denote the discretisation by

$$D^i_{\overline{j}}\left(\left[v^i\right]_{\overline{j}}, v^i_{\overline{j}}; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right) = 0. \tag{1.15}$$

Using this notation, we are able to consider the value at the current grid point and at the remaining points separately.
With (1.15) we have

$$D^i\left(v^i; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right) = \\ \left(D^i_{\overline{j}}\left(\left[v^i\right]_{\overline{j}}, v^i_{\overline{j}}; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right)\right)_{\overline{j}\in J_{M^i}}.$$

We now give the definition of a monotone scheme. In order to keep the definition short we a priori agree on a sign convention, s. Remark 1.24.

Definition 1.23 (Monotonicity of a scheme)**.**
Let $D^i(v^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}) = 0$, $i \in I_L$, be a discretisation of a pde of the form (1.13) over Ω_n following Notation 1.22. Furthermore, let $\mathcal{G}(L, \{t^i\}, \mathcal{S}, \Omega_n)$ be a uniform grid as introduced in Notation 1.20.
We call a discretisation $D^i(\cdot) = 0$ monotone if the following conditions hold for all time levels $i \in I_L$ and grid all points $(t^i, x^i_{\overline{j}}) \in \mathcal{G}, \overline{j} \in J_{M^i}$:

1. *For all* $\varepsilon^{i-1}, \varepsilon^i \in \mathbb{R}_+^M$ *the following inequality is valid:*

$$D_{\bar{j}}^i\left(\left[v^i+\varepsilon^i\right]_{\bar{j}}, v_{\bar{j}}^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}+\varepsilon^{i-1}\right) \leq D_{\bar{j}}^i\left(\left[v^i\right]_{\bar{j}}, v_{\bar{j}}^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right) \tag{1.16a}$$

2. *For all* $\varepsilon_{\bar{j}}^i \in \mathbb{R}_+$ *the following inequality holds:*

$$D_{\bar{j}}^i\left(\left[v^i\right]_{\bar{j}}, v_{\bar{j}}^i+\varepsilon_{\bar{j}}^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right) \geq D_{\bar{j}}^i\left(\left[v^i\right]_{\bar{j}}, v_{\bar{j}}^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right) \tag{1.16b}$$

Remark 1.24.
Let α *be the coefficient in front of* $v_{\bar{j}}^i$ *if we write out discretisation* $D_{\bar{j}}^i(\cdot)$ *in Definition 1.23 in full. We agree on the convention* $\alpha \geq 0$*. Otherwise the symbols "*$\leq$*" and "*$\geq$*" in the inequalities (1.16a) and (1.16b) must be changed to "*$\geq$*" and "*$\leq$*", respectively.*

If the spatial grid is time-dependent, Definition 1.23 has to be extended to non-uniform spatial grids.

Two additional definitions are needed to give a complete formulation of the convergence theorem we want to formulate. We have to define when a scheme is consistent and when it is stable.

Definition 1.25 (Stability of a scheme)**.**
A numerical scheme $D^i\left(v^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right) = 0$ *for a pde of form* (1.9) *is stable regarding the maximum norm if for all time levels* $i \in I_L$ *the solution* v^i *is bounded independently of* Δt *and* H*. That is, there exists a constant* $C \in \mathbb{R}_+$ *independent of* Δt *and* H *so that*

$$\|v^i\|_\infty \leq C, \quad i \in J_L.$$

Definition 1.26 (Consistency of a scheme)**.**
A numerical scheme $D^i\left(v^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right) = 0$ *for a pde of form* (1.9) *is consistent if*

$$\left|\frac{\partial v}{\partial t} - F\left(t, x, v, D_x v, D_x^2 v\right) - D^i\left(v^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right)\right| \to 0$$

for $\Delta t, H \to 0$*.*

The definitions given above are standard in the numerical literature and can for example be found in the book [SB02].
Being prepared this way, it is possible to guarantee the convergence of a discretisation to the viscosity solution of a pde. The main result has been established by Barles and Souganidis [BS90, BDR95]. The result is as follows

Theorem 1.4 (Convergence to the viscosity solution)**.**
Let a pde of the form (1.13) with unique viscosity solution v be given. Furthermore, the pde is discretised on a grid $\mathcal{G}(L, \{t^i\}, \mathcal{S}, \Omega_n)$ by a numerical scheme which is given by the discretisation

$$D^i_{\bar{j}}\left(\left[v^i\right]_j, v^i_j; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right) = 0 \tag{1.17}$$

at a grid point $(t^i, x_{\bar{j}}) \in \mathcal{G}$. For $\Delta t, H \to 0$ the approximation $v^i, i \in J_L$, solving (1.17) converges to the unique viscosity solution v if the scheme is

1. *consistent in the sense of Definition 1.26,*
2. *stable in the sense of Definition 1.25,*
3. *monotone in the sense of Definition 1.23,*

and the equation satisfies the Strong Comparison Principle.

The proof of this theorem can be found for example in [BS90].

Theorem 1.4 gives a path to construct a scheme which converges towards the right solution in the financial context. The first and the second condition will be no severe difficulty in constructing such a scheme. The most challenging task is to maintain monotonicity of the scheme and prove this property for a maximal set of model parameters.

Remark 1.27.
From Definition 1.23 it is easy to see that at a grid point $(t^i, x^i_{\bar{j}}) \in \mathcal{G}$ the coefficients in front of the variables $v^i_{\bar{j}}$ should be non-negative and additionally those in front of $v^i_m, m \in I_M, m \neq \bar{j}$, should be non-positive. In the literature schemes with these features are known as positive coefficient[6] schemes. They have been investigated in several publications. Among these are [WF08, ZFV01, ZFV03, FL07].

[6] In the cited publications the positive coefficient scheme is defined with inverse signs compared to the postulation above. The sign convention is demanded for a matrix A, where $B = I - A$ is the iteration matrix. In this work we require the signs of B to fulfil this convention.

Lemma 1.5.
Let $A^i \in \mathbb{R}^{n\times n}, i \in J_L$, be a matrix with non-positive off-diagonal entries and non-negative diagonal entries, where all entries could possibly depend on t and x. The index i displays the dependency of A^i of the current time level.
Then the discretisation scheme of a pde with one time step on a grid $\mathcal{G}(L, \{t^i\}, \mathcal{S}, \Omega_n)$ given by

$$D^i\left(v^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right) = A^i v^i - v^{i-1} - b^i = 0, \tag{1.18}$$

where $b^i \in \mathbb{R}^n$ is a vector independent of v^{i-1} and v^i, is monotone.

Proof.
We will directly verify the conditions of Definition 1.23.
Let $\varepsilon^{i-1}, \varepsilon^i \in \mathbb{R}_+^M$ be two vectors. Then for every grid point $(t^i, x_{\bar{j}}^i) \in \mathcal{G}$ we have

$$\begin{aligned}
& D_{\bar{j}}^i\left(\left[v^i + \varepsilon^i\right]_{\bar{j}}, v_{\bar{j}}^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1} + \varepsilon^{i-1}\right) \\
&= \quad A_{\bar{j}}^i v^i + \underbrace{A_{\bar{j}}^i \left[\varepsilon^i\right]_{\bar{j}}}_{\leq 0} - v_{\bar{j}}^{i-1} - \varepsilon_{\bar{j}}^{i-1} - b_{\bar{j}}^i \\
&\overset{a_{j,j}^i \varepsilon_{\bar{j}}^i = 0}{\leq} \quad A_{\bar{j}}^i v^i - v_{\bar{j}}^{i-1} - b_{\bar{j}}^i \\
&= \quad D_{\bar{j}}^i\left(\left[v^i\right]_{\bar{j}}, v_{\bar{j}}^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right),
\end{aligned}$$

where $A_{\bar{j}}^i$ denotes the j-th row of A^i. Thus condition (1.16a) for monotonicity holds. Moreover, for $\varepsilon_{\bar{j}}^i \in \mathbb{R}_+$ we have

$$\begin{aligned}
& D_{\bar{j}}^i\left(\left[v^i\right]_{\bar{j}}, v_{\bar{j}}^i + \varepsilon_{\bar{j}}^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right) \\
&= \quad A_{\bar{j}}^i v^i + \varepsilon_{\bar{j}}^i A_{\bar{j}}^i e_{\bar{j}} - v_{\bar{j}}^{i-1} - b_{\bar{j}}^i \\
&\overset{a_{j,j}^i \varepsilon_{\bar{j}}^i \geq 0}{\geq} \quad A_{\bar{j}}^i v^i - v_{\bar{j}}^{i-1} - b_{\bar{j}}^i \\
&= \quad D_{\bar{j}}^i\left(\left[v^i\right]_{\bar{j}}, v_{\bar{j}}^i; t^i, \mathcal{S}, t^{i-1}, \mathcal{S}, v^{i-1}\right)
\end{aligned}$$

Thus also condition (1.16b) holds and the discretisation (1.18) is monotone.

□

2 Monotone discretisations

In the previous chapter we have seen three criterions a discretisation scheme has to satisfy in order to converge to the unique viscosity solution of a non-linear partial differential equation. As we will see the one which is the most restrictive one is monotonicity.
In this chapter we thoroughly analyse the applicability of general Finite Difference schemes which we will formulate for rectangular nonuniform grids. Sufficient conditions to guarantee monotonicity are derived for different versions and dimensionality of the pricing equation. A necessity is the distinction between one- and multi-dimensional Finite Difference schemes since the criterions differ. Exemplarily for the latter one, we consider the discretisation for the two-dimensional pricing problem. The results obtained here can easily be transferred to higher dimensional pricing problems.
In all cases we prove that we can construct a scheme which satisfies monotonicity under specific assumptions and procedures. The grid on which the Finite Difference schemes will be based is one key point in this context. In order to build a grid with many sampling points in the area of interest without increasing their total number, we develop a method for building non-uniform grids. The discretisation matrix associated with the resulting schemes possesses an immanent property which helps proving their convergence.

In the first section we transform the Black-Scholes-Barenblatt equation for multi-dimensional pricing problems. It is considered in backward time and with log-prices. The main part of this chapter is the second section in which the Finite Difference discretisation in space is performed. We commence with the two-dimensional pricing problem in transformed coordinates and the properties of the corresponding grid. We derive a general condition for the volatility set which has to hold so that the desired properties can at all be preserved by skilled construction. In the following section we focus on sufficient conditions for the spatial step sizes regarding monotonicity.
Hereafter, we generalise the results towards the pricing of options on more than two assets. In the last subsection of Section 2.2 we focus on the one-dimensional pricing problem.

The schemes for the discretisation in time are introduced in Section 2.3. We finally formulate the complete Finite Difference schemes for pricing European options in a general form.
In the fourth section M-matrices are introduced and we address the M-Matrix property of the discretisation matrix.
At last, the fifth section focusses on the problem of pricing American options and introduces the well known penalty iteration to approach the associated complementary problem.

2.1 Transformation of the Black-Scholes-Barenblatt equation

The Black-Scholes-Barenblatt equation (BSB_n) is not discretised directly, but will at first be transformed to log-prices and backward time. We define new variables

$$\tau = T - t,\ t \in [0,T], \quad \text{and} \quad x_i = \log\left(\frac{S_i}{K}\right),\ i \in I_n, \tag{2.1}$$

where K is the strike price of the option. This means that we change the original domain Ω_n to $\tilde{\Omega}_n := [0,T] \times \mathbb{R}^n$.
By transformation (2.1) we obtain the following Hamilton-Jacobi-Bellman equation for an option on n assets

$$\begin{aligned} \frac{\partial v}{\partial \tau} - \frac{1}{2} \max_{\substack{\sigma_i \in \mathcal{V}_i \\ \rho_{k,l} \in \mathcal{C}_{k,l} \\ i,k,l \in I_n}} \left\{ \sum_{i=1}^{n} \sigma_i^2 \left(\frac{\partial^2 v}{\partial x_i^2} - \frac{\partial v}{\partial x_i} \right) + \sum_{\substack{k,l=1 \\ k \neq l}}^{n} 2\sigma_k \sigma_l \rho_{k,l} \frac{\partial^2 v}{\partial x_k \partial x_l} \right\} \\ - \sum_{i=1}^{n} (r - \delta_i) \frac{\partial v}{\partial x_i} + rv = 0, \end{aligned} \tag{2.2}$$

compare the notations introduced in Section 1.2.
The existence of a solution v of equation (2.2) in the viscosity sense is guaranteed by the used transformation, cp. the book of Bardi and Dolcetta [BD97], since a viscosity solution of equation (BSB_n) exists.

In the following text we will often use the term “controls” to comprise volatilities and correlations. It stems of the context of Hamilton-Jacobi-Bellman equations.

Notation 2.1.
For the set of admissible controls we write

$$\Theta_n := \bigotimes_{i=1}^{n} \mathcal{V}_i \times \bigotimes_{i,j=1, i\neq j}^{n} \mathcal{C}_{i,j} \subset \mathbb{R}^{n+0.5(n-1)n}.$$

A single control is denoted by $\gamma \in \Theta_n$.

Now, we formulate the three pricing problems which are of basic interest for the remainder of this thesis. We denote with v_x and v_{xx} the first and second order continuous derivatives of v in x-direction, respectively.

Problem 2.2 (European options on one underlying asset)**.**
Solve the equation

$$\begin{aligned} v_\tau - \frac{1}{2} \max_{\sigma_1 \in \Theta_1} \left\{ \sigma_1^2 \left(v_{xx} - v_x \right) \right\} - r v_x + r v = 0 \\ v(0, x) = \Psi(x) \end{aligned} \tag{2.3}$$

on $\tilde{\Omega}_1$.

Problem 2.3 (European options on two underlying assets)**.**
Solve the equation

$$\begin{aligned} v_\tau - \frac{1}{2} \max_{(\sigma_1, \sigma_2, \rho_{1,2}) \in \Theta_2} \left\{ \sigma_1^2 \left(v_{x_1 x_1} - v_{x_1} \right) + 2\rho_{1,2} \sigma_1 \sigma_2 v_{x_1 x_2} + \sigma_2^2 \left(v_{x_2 x_2} - v_{x_2} \right) \right\} \\ - r \left(v_{x_1} + v_{x_2} \right) + r v = 0 \\ v(0, x_1, x_2) = \Psi(x_1, x_2) \end{aligned} \tag{2.4}$$

on $\tilde{\Omega}_2$.

Pricing problems for options on more than two assets can be formulated analogously. In short notation we rewrite (2.2) as

$$\frac{\partial}{\partial \tau} v + \mathcal{L}_{BSB}^n v = 0, \quad n \in \mathbb{N},$$

where $\mathcal{L}_{BSB}^n$ is defined accordingly.
The value of an American option is then assumed to be deterined by a non-linear complementary problem.

Problem 2.4 (American Options on n underlying assets)**.**
Solve

$$\begin{aligned} \frac{\partial v}{\partial t} - \mathcal{L}_{BSB}^n v &\leq 0, \\ v - \Psi &\geq 0, \\ \left(\frac{\partial v}{\partial t} - \mathcal{L}_{BSB}^n v\right)(v - \Psi) &= 0, \end{aligned}$$

on Ω_n.

2.2 Spatial discretisation

The idea of Finite Differences is to approximate the derivatives of v via difference quotients at every sampling point of a given grid. This proceeding results in a system of equations which has to be solved at every spatial grid point and successively in time.

We will approach the pricing problem with a general rectangular grid. We now introduce the non-equidistant difference quotients for the first and second derivative in a single direction. They can all be derived by Taylor expansion.
As usual we do not solve the pricing equation (2.2) on $\tilde{\Omega}_n = [0,T] \times \mathbb{R}^n$, but we solve (2.3) on a truncated domain $\Omega_1^D := [0,T] \times [x^{min}, x^{max}]$ and analogously (2.4) on $\Omega_2^D := [0,T] \times [x_1^{min}, x_1^{max}] \times [x_2^{min}, x_2^{max}]$.

Notation 2.5.
Let a grid $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_n) \subset \Omega_n^D$ *as in Notation 1.20 be given. The grid points in* x_μ*-direction,* $\mu = 1, \ldots, d$*, will be denoted by* $x_\mu^{\nu,i}, \nu \in J_{M_\mu}$*. With our usual notation we have*

$$(\tau^i, \overline{x}_{\overline{j}}) = (\tau^i, x_1^{\nu_1,i}, \ldots, x_d^{\nu_d,i}) \in \mathcal{G},$$

where $\overline{j}$ *denotes the lexicographical ordering at each time level:*

$$\overline{j} = j(\nu_1, \ldots, \nu_d) := \sum_{k=1}^{d} \prod_{l=1}^{k-1} (M_l^i + 1) + \nu_k.$$

With $h_\mu^{\nu,i} := x_\mu^{\nu,i} - x_\mu^{\nu-1,i}, \nu \in I_{M_\mu}$*, the spatial step sizes of the grid are denoted. Without loss of generality (w. l. o. g.) we assume* $h_\mu^{\nu,i} < H$*. The step width for the discretisation in time is denoted by* $h_\tau^i := \tau^i - \tau^{i-1}$.

The difference quotients are introduced in

Notation 2.6.
Let us assume that a two-dimensional grid is given. We write $v^i_{k,j}$ for the approximation of $v(\tau_i, x_1^k, x_2^j)$. Then the spatial derivatives can be approximated as follows.

1. *For the approximation first derivative we write*

$$\delta^f_{x_1} v^i_{k,l} := \frac{1}{h_1^{k+1}} \left(v^i_{k+1,l} - v^i_{k,l} \right), \tag{2.5a}$$

$$\delta^b_{x_1} v^i_{k,l} := \frac{1}{h_1^{k}} \left(v^i_{k,l} - v^i_{k-1,l} \right), \text{ and} \tag{2.5b}$$

$$\delta^c_{x_1} v^i_{k,l} := \frac{1}{h_1^k + h_1^{k+1}} \left(\frac{h_1^k}{h_1^{k+1}} v^i_{k+1,l} - \left(\frac{h_1^k}{h_1^{k+1}} - \frac{h_1^{k+1}}{h_1^k} \right) v^i_{k,l} - \frac{h_1^{k+1}}{h_1^k} v^i_{k-1,l} \right) \tag{2.5c}$$

   for the forward, backward, and centered difference quotient, respectively.

2. *For the approximation of the second derivative in x_1-direction we write*

$$\delta^2_{x_1} v^i_{k,l} := \frac{2}{h_1^{k+1} + h_1^k} \left(\frac{v^i_{k-1,l}}{h_1^k} - \left(\frac{1}{h_1^k} + \frac{1}{h_1^{k+1}} \right) v^i_{k,l} + \frac{v^i_{k+1,l}}{h_1^{k+1}} \right). \tag{2.6}$$

3. *In a two-dimensional example the mixed derivative is approximated by*

$$\begin{aligned} \delta^+_{x_1x_2} v^i_{k,l} := & \frac{1}{2h_1^k h_2^l} \left(v^i_{k-1,l-1} - v^i_{k,l} + h_2^l \delta^{\cdot}_{x_2} v^i_{k,l} + h_1^k \delta^{\cdot}_{x_1} v^i_{k,l} \right) \\ & - \frac{1}{2h_1^k h_2^l} \left(\frac{1}{2} \left(h_2^l \right)^2 \delta^2_{x_2} v^i_{k,l} + \frac{1}{2} \left(h_1^k \right)^2 \delta^2_{x_1} v^i_{k,l} \right) \\ & + \frac{1}{2h_1^{k+1} h_2^{l+1}} \left(v^i_{k+1,l+1} - v^i_{k,l} - h_2^{l+1} \delta^{\cdot}_{x_2} v^i_{k,l} - h_1^{k+1} \delta^{\cdot}_{x_1} v^i_{k,l} \right) \\ & - \frac{1}{2h_1^{k+1} h_2^{l+1}} \left(\frac{1}{2} \left(h_2^{l+1} \right)^2 \delta^2_{x_2} v^i_{k,l} + \frac{1}{2} \left(h_1^{k+1} \right)^2 \delta^2_{x_1} v^i_{k,l} \right) \end{aligned} \tag{2.7a}$$

   and

$$\begin{aligned} \delta^-_{x_1x_2} v^i_{k,l} := & - \frac{1}{2h_1^k h_2^{l+1}} \left(v^i_{k-1,l+1} - v^i_{k,l} - h_2^{l+1} \delta^{\cdot}_{x_2} v^i_{k,l} + h_1^k \delta^{\cdot}_{x_1} v^i_{k,l} \right) \\ & + \frac{1}{2h_1^k h_2^{l+1}} \left(\frac{1}{2} \left(h_2^{l+1} \right)^2 \delta^2_{x_2} v^i_{k,l} + \frac{1}{2} \left(h_1^k \right)^2 \delta^2_{x_1} v^i_{k,l} \right) \\ & - \frac{1}{2h_1^{k+1} h_2^l} \left(v^i_{k+1,l-1} - v^i_{k,l} + h_2^l \delta^{\cdot}_{x_2} v^i_{k,l} - h_1^{k+1} \delta^{\cdot}_{x_1} v^i_{k,l} \right) \\ & + \frac{1}{2h_1^{k+1} h_2^l} \left(\frac{1}{2} \left(h_2^l \right)^2 \delta^2_{x_2} v^i_{k,l} + \frac{1}{2} \left(h_1^{k+1} \right)^2 \delta^2_{x_1} v^i_{k,l} \right) \end{aligned} \tag{2.7b}$$

The difference quotients in 1. and 2. are used in an analogue way if we approximate derivatives in x_2-direction. We use the notation $\delta^{\cdot}_{x_1}$ to clarify that all three difference quotients introduced for the first derivative can be used.

The difference quotient $\delta^+_{x_1x_2}$ for the mixed derivative can be constructed by choosing a convex combination of the Taylor expansion of $v^i_{k+1,l+1}$ and $v^i_{k-1,l-1}$ and omitting terms of higher order. In the same way the second one can be constructed by considering $v^i_{k-1,l+1}$ and $v^i_{k+1,l-1}$. They are used in different situations, v. i.. The difference quotients stated here can be found for example in the paper by Ikonen and Toivanen [IT07]. If we work on a uniform grid, all difference quotients simplify to the standard ones which can be found in the book of Seydel [Sey12], for example.

The accuracy of the single approximations can be derived by Taylor expansion. We just state the results. The forward and backward difference quotients are first order accurate:

$$\delta^f_{x_1} v^i_{k,l} - \frac{\partial}{\partial x_1} v(t^i, x^k_1, x^l_2) = \frac{1}{2} h^{k+1}_1 \frac{\partial}{\partial x_1} v(t^i, \xi) + \mathcal{O}\left(\left(h^{k+1}_1\right)^2\right)$$
$$\delta^b_{x_1} v^i_{k,l} - \frac{\partial}{\partial x_1} v(t^i, x^k_1, x^l_2) = \frac{1}{2} h^{k}_1 \frac{\partial}{\partial x_1} v(t^i, \xi) + \mathcal{O}\left(\left(h^{k}_1\right)^2\right)$$

For the centered difference quotient we have the error estimation

$$\delta^c_{x_1} v^i_{k,l} - \frac{\partial}{\partial x_1} v(t^i, x^k_1, x^l_2) = \frac{1}{6} h^k_1 h^{k+1}_1 \frac{\partial^3}{\partial x^3_1} v(t^i, \xi) + \text{t. h. o.},$$

with the abbreviation t. h. o. used for terms of higher order. For the approximation of the second partial derivative we obtain

$$\delta^2_{x_1} v^i_{k,l} - \frac{\partial^2}{\partial (x_1)^2} v(t^i, x^k_1, x^l_2) = \frac{1}{6}\left(h^k_1 - h^{k+1}_1\right) \frac{\partial^3}{\partial x^3_1} v(t^i, \xi) + \frac{\partial^4}{\partial x^4_1} v(t^i, \xi) \frac{\left(h^k_1\right)^3 + \left(h^{k+1}_1\right)^3}{24(h^k_1 + h^{k+1}_1)} + \text{t. h. o.}$$

Here $\xi \in \mathbb{R}^2$ always denotes an appropriately chosen point.

We now start to discretise equation (2.4).

2.2.1 Options on two assets

This section is the basis for all sections on spatial discretisation to follow. We introduce notation which can be used analogously in the one- and general n-dimensional, $n > 2$, case. The results which will be derived also apply to discretisations in different dimensions, cp. Section 2.2.3 and Section 2.2.4. The pricing problem is given by Problem 2.3.

In their book [KD01] Kushner-Dupuis (KD) analyse the so called KD scheme for a general class of HJB equations. These are discretised on a completely rectangular and uniform spatial grid. In opposite to the present approach the KD scheme uses one-sided difference quotients for the convection terms. They prove the monotonicity of the scheme if all correlation matrices of the set Γ_n are diagonally dominant. The approach in this thesis differs in the way the first derivatives are discretised. As we will see, a consequence is that we have to take a close look at the spatial step sizes. But we also might gain a scheme of higher order. Furthermore, we can prove for the present scheme that it is applicable to a more general set of parameters.

Let a grid $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_2) \subset \tilde{\Omega}_2^D$ as introduced in Notation 2.5 be given. For the discretisation of equation (2.4) we assume a non-zero correlation $\rho_{1,2}$ for the time being. For the case $\rho_{1,2} = 0$ see Remark 2.14.

In Chapter 1.3.2 we pointed out that the off-diagonal elements of the discretisation matrix have to be non-positive and the diagonal elements non-negative. This necessary property reduces the number of possible answers to the question on how to deal with convection dominance. That is, the coefficient multiplied to the convection term is bigger than the one multiplied to the diffusion term. A sum of these coefficients amount to the elements of the discretisation matrix and thus determines its sign. We will handle this task by an appropriate choice of the difference quotients. It is important to keep in mind that convection dominance can occur "suddenly" due to the uncertain character of volatility.
In order to do a complete discretisation of equation (2.4), we also will have to discretise the control $(\sigma_1, \sigma_2, \rho_{1,2}) \in \Theta_2$. In the discrete setup, we must determine this triple at every grid point. In this subsection we assume the controls to be constant, but always keep in mind the more general case by adjusted notation. The subscript γ indicates this dependence. The discretisation of the controls and the final determination of the "right" or "optimal" control is the topic of Chapter 3.

Notation 2.7.
Let a grid $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_2)$ with M^i grid points in space at each time level as introduced in Notation 2.6 be given. With $\gamma^i_{k,l} = (\sigma^i_{1,\bar{j}}, \sigma^i_{2,\bar{j}}, \rho^i_{1,2,\bar{j}}) \in \Theta_2$ we denote the discrete control at the grid point (τ^i, x_1^k, x_2^l). To keep notation simple, we do not use a time index at every single control. By the matrix $\gamma^i = (\gamma^i_{k,l})_{k,l} \in \Theta_n^{M^i}$ we denote all controls on the grid in space at time level i, where $\Theta_n^{M^i} := (\Theta_n)^{M^i \times n}$.
The subscript γ combined with matrices, discretisations, etc. will denote their dependence on the control.

Now we start discretising equation (2.4). For sake of simplicity we omit the time index i since we consider a fixed time level τ^i. For the time being let $\gamma \in \Theta_n^M$ be fix. We commence by looking at the discrete version of the equation at an inner grid point with index $(\cdot, k, l)$, $k = 1, \ldots, M_1 - 1, l = 1, \ldots, M_2 - 1$. The time discretisation is not yet performed.

Without loss of generality we will assume $\mathcal{C}_{1,2} \subset \mathbb{R}_+$ and use the approximation $\delta^+_{x_1x_2}$ for the mixed derivative in (2.4). For the other cases see Remark 2.18. At an arbitrary inner grid point we obtain the following equation by rearranging terms and using (2.7a):

$$\begin{aligned} v_\tau + d^{k,l}_{1,\gamma}\delta^2_{x_1}v^i_{k,l} + d^{k,l}_{2,\gamma}\delta^2_{x_2}v^i_{k,l} + c^{k,l}_{1,\gamma}\delta^{\cdot}_{x_1}v^i_{k,l} + c^{k,l}_{1,\gamma}\delta^{\cdot}_{x_2}v^i_{k,l} + (r - a^{k,l}_{1,\gamma} - a^{k,l}_{2,\gamma})v^i_{k,l} \\ + a^{k,l}_{1,\gamma}v^i_{k+1,l+1} + a^{k,l}_{2,\gamma}v^i_{k-1,l-1} = 0, \end{aligned} \tag{2.8a}$$

where

$$d^{k,l}_{1,\gamma} = -\frac{1}{2}\sigma^2_{1,\bar{j}} + \frac{1}{4}\rho_{1,2,\bar{j}}\sigma_{1,\bar{j}}\sigma_{2,\bar{j}}\left(\frac{h^k_1}{h^l_2} + \frac{h^{k+1}_1}{h^{l+1}_2}\right), \tag{2.8b}$$

$$d^{k,l}_{2,\gamma} = -\frac{1}{2}\sigma^2_{2,\bar{j}} + \frac{1}{4}\rho_{1,2,\bar{j}}\sigma_{1,\bar{j}}\sigma_{2,\bar{j}}\left(\frac{h^l_2}{h^k_1} + \frac{h^{l+1}_2}{h^{k+1}_1}\right), \tag{2.8c}$$

$$c^{k,l}_{1,\gamma} = \frac{1}{2}\sigma^2_{1,\bar{j}} - r + \delta_1 - \frac{1}{2}\rho_{1,2,\bar{j}}\sigma_{1,\bar{j}}\sigma_{2,\bar{j}}\left(\frac{1}{h^l_2} - \frac{1}{h^{l+1}_2}\right), \tag{2.8d}$$

$$c^{k,l}_{2,\gamma} = \frac{1}{2}\sigma^2_{2,\bar{j}} - r + \delta_2 - \frac{1}{2}\rho_{1,2,\bar{j}}\sigma_{1,\bar{j}}\sigma_{2,\bar{j}}\left(\frac{1}{h^k_1} - \frac{1}{h^{k+1}_1}\right), \tag{2.8e}$$

$$a^{k,l}_{1,\gamma} = -\frac{1}{2}\rho_{1,2,\bar{j}}\sigma_{1,\bar{j}}\sigma_{2,\bar{j}}\frac{1}{h^{k+1}_1 h^{l+1}_2}, \text{ and} \tag{2.8f}$$

$$a^{k,l}_{2,\gamma} = -\frac{1}{2}\rho_{1,2,\bar{j}}\sigma_{1,\bar{j}}\sigma_{2,\bar{j}}\frac{1}{h^k_1 h^l_2} \tag{2.8g}$$

In order to obtain a monotone scheme, our aim is to construct a seven or nine point finite difference stencil for the space discretisation, which results in a matrix with non-positive off-diagonal entries and non-negative diagonal-entries. We introduce the following

Notation 2.8.
At a grid point with index $(\cdot, k, l)$ the difference stencil is denoted by:

$$\begin{bmatrix} 0 & a^{k,l+1}_{k,l,\gamma} & a^{k+1,l+1}_{k,l,\gamma} \\ a^{k-1,l}_{k,l,\gamma} & a^{k,l}_{k,l,\gamma} & a^{k+1,l}_{k,l,\gamma} \\ a^{k-1,l-1}_{k,l,\gamma} & a^{k,l-1}_{k,l,\gamma} & 0 \end{bmatrix} \tag{2.9}$$

The lower index (k, l, γ) denotes the grid point in consideration and the dependence on the control. The upper indices tag the index of the corresponding option value

which is multiplied to the entries.
In the discretisation matrix the coefficients are placed in row $j(k,l)$.
For the indices of the adjacent grid points to (x_1^k, x_2^l) we write

$$N(k,l) := \{(k-1,l),(k+1,l),(k,l-1),(k,l+1),(k-1,l-1),(k+1,l+1)\}.$$

Notice that if the correlation is negative equation (2.8a) and the stencil would be of slightly modified form.

Going on from (2.8), the monotone scheme can be developed. The coefficients in (2.8f) and (2.8g) represent the off-diagonal elements in the discretisation matrix. Therefore, they have to be non-positive. Thus, due to the sign of $v^i_{k+1,l+1}$ and $v^i_{k-1,l-1}$ in $\delta^+_{x_1,x_2}$, we use this approximation for positive correlation.

Non-negativity condition. The next step is to assure the non-negativity of the diffusion coefficients (2.8b) and (2.8c). In the remainder of this section we clarify why we start from this point. For (2.8b) it is

$$d^{k,l}_{1,\gamma} \leq 0 \quad \Leftrightarrow \frac{h_1^k}{h_2^l} + \frac{h_1^{k+1}}{h_2^{l+1}} \leq \frac{\sigma_{1,\bar{j}}}{\sigma_{2,\bar{j}}\rho_{1,2,\bar{j}}} + \frac{\sigma_{1,\bar{j}}}{\sigma_{2,\bar{j}}\rho_{1,2,\bar{j}}}$$

which is satisfied for

$$h_1^k \leq \frac{\sigma_{1,\bar{j}}}{\sigma_{2,\bar{j}}\rho_{1,2,\bar{j}}} h_2^l \quad \text{and} \quad h_1^{k+1} \leq \frac{\sigma_{1,\bar{j}}}{\sigma_{2,\bar{j}}\rho_{1,2,\bar{j}}} h_2^{l+1}. \tag{2.10a}$$

Analogously, for

$$h_2^l \leq \frac{\sigma_{2,\bar{j}}}{\sigma_{1,\bar{j}}\rho_{1,2,\bar{j}}} h_1^k \quad \text{and} \quad h_2^{l+1} \leq \frac{\sigma_{2,\bar{j}}}{\sigma_{1,\bar{j}}\rho_{1,2,\bar{j}}} h_1^{k+1} \tag{2.10b}$$

the inequality $d^{k,l}_{2,\gamma} \leq 0$ holds true.
At every grid point the four inequalities in (2.10) have to be fulfilled. Moreover, we have to take account of the fact that volatilities and correlations are determined by a bandwidth and not a single figure. For fixed controls (2.10a) and (2.10b) hold true if

$$\frac{\sigma_{1,\bar{j}}\rho_{1,2,\bar{j}}}{\sigma_{2,\bar{j}}} h_2^l \;\leq\; h_1^k \;\leq\; \frac{\sigma_{1,\bar{j}}}{\sigma_{2,\bar{j}}\rho_{1,2,\bar{j}}} h_2^l,$$

for all k, l. We now also take the constraints on volatility and correlation into account. The inequalities to hold become

$$\begin{aligned} \Xi^d_{1,2} h_2^l \;\leq\; h_1^k \;\leq\; \Xi^u_{1,2} h_2^l \quad \text{with} \\ \Xi^u_{1,2} := \frac{\underline{\sigma_1}}{\overline{\sigma_2}\,\overline{\rho_{1,2}}} \quad \text{and} \quad \Xi^d_{1,2} = \frac{\overline{\sigma_1}\,\overline{\rho_{1,2}}}{\underline{\sigma_2}} \end{aligned} \tag{2.11}$$

for all $k \in I_{M_1}, l \in I_{M_2}$.

A consequence of (2.11), which a priori has to be fulfilled by the model itself, is

Assumption 2.9.
For the coefficients $\Xi^d_{1,2}$ and $\Xi^u_{1,2}$ depending on the bounds of the control we assume the inequality

$$\Xi^d_{1,2} \leq \Xi^u_{1,2}$$

to hold true.

With the same argumentation presented above we obtain

$$\frac{1}{\Xi^u_{1,2}} \leq \frac{h_2^l}{h_1^k} \leq \frac{1}{\Xi^d_{1,2}}, k \in I_{M_1}, l \in I_{M_2}, \quad \Rightarrow \quad d^{k,l}_{2,\gamma} \leq 0.$$

Thus, Assumption 2.9 is sufficient for both spatial directions.

It is the only condition which cannot be handled by an appropriate choice of the numerical parameters like step sizes. For a better understanding of Assumption 2.9 we illustrate it in the following

Example 2.10.
The following three cases display three different situations:

1. *If the set of diffusion matrices Γ_2' only contains strictly diagonally dominant matrices, Assumption 2.9 is always satisfied. We have*

$$\sigma_1^2 > \sigma_1\sigma_2\rho_{1,2} \Rightarrow 1 < \frac{\sigma_1}{\rho_{1,2}\sigma_2} \quad \textit{and} \quad \sigma_2^2 > \sigma_1\sigma_2\rho_{1,2} \Rightarrow 1 > \frac{\rho_{1,2}\sigma_1}{\sigma_2}$$

   for all choices of controls and thus especially for $\Xi^u_{1,2}$ and $\Xi^d_{1,2}$. For the more general case see Assumption 2.23 and Example 2.25.

2. *The set $\mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{C}_{1,2} = [0.3, 0.4] \times [0.1, 0.2] \times [0.6, 0.8]$ fulfils Assumption 2.9. Although the corresponding correlation matrices are not diagonally dominant.*

3. *For* $\mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{C}_{1,2} = [0.2, 0.4] \times [0.1, 0.3] \times [0.6, 0.8]$ *the criterion is not satisfied. We have* $\Xi^d_{1,2} = \frac{32}{30} \geq \frac{25}{30} = \Xi^u_{1,2}$*, although it will occur in a financial context.*

The deliberations and the example above lead us to the following

Definition 2.11.
Let Assumption 2.9 be satisfied. A two-dimensional grid $\mathcal{G}(L, \{t^i\}, \mathcal{S}, \Omega_2)$ *is called monotone if it satisfies condition* (2.12).

In Figure 2.1 the influence of correlation on the maximal intervals for volatility is displayed. Here, a uniform grid is assumed with $h^k_1 = h^l_2, k, l \in I_M$. For all sets Θ_2 lying in the coloured area the uniform grid is monotone. The bigger the maximal correlation gets the smaller the possible choices for Θ_2 become. Finally, for $\overline{\rho_{1,2}} \approx 1$ no meaningful choice might be possible.

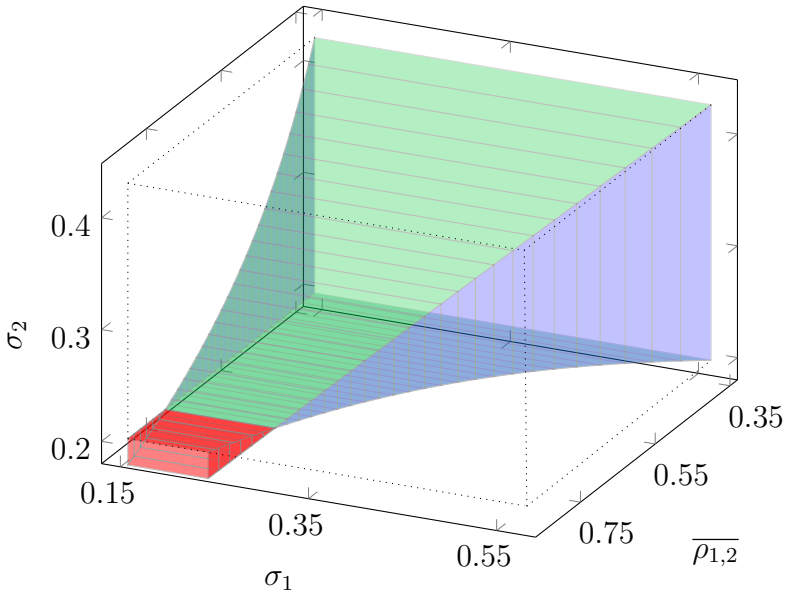

Fig. 2.1: *Example of the dependence of correlation and volatility on each other. The values* $\underline{\sigma_1} = 0.15$ *and* $\overline{\sigma_2} = 0.2$ *are fixed. The upper bounds are given by* (2.11)*. The blue and green surfaces display the lower and upper bound for* σ_1 *and* σ_2*, respectively. In the red area there is no possible choice.*

Since for a fixed $l \in I_{M_2}$ and all $k \in I_{M_1}$ and vice versa the four constraints (2.11) have to hold true, we combine them to

$$\Xi^d_{1,2} \max_{l \in J_{M_2}} \{h^l_2\} \leq h^k_1 \leq \Xi^u_{1,2} \min_{l \in J_{M_2}} \{h^l_2\}, \quad k \in I_{M_1}, \quad \text{and} \tag{2.12a}$$

$$\frac{1}{\Xi^u_{1,2}} \max_{k \in J_{M_1}} \{h^k_1\} \leq h^l_2 \leq \frac{1}{\Xi^d_{1,2}} \min_{k \in J_{M_1}} \{h^k_1\}, \quad l \in I_{M_2}. \tag{2.12b}$$

We conclude the construction with the following remarks.

Remark 2.12.
If we had not chosen the transformed version of the Black-Scholes-Barenblatt equation (2.4), the constants $\Xi^u_{1,2}$ and $\Xi^d_{1,2}$ could have become impractical to handle. Due to the structure in (BSB_n) the quotient $\frac{S_1}{S_2}$ would have been a part of them. Especially in the upper left and the lower right corner of the grid they would have been quite big respectively small. Thus, by inequality (2.12), the effort to obtain a monotone scheme would have been comparatively high.

Remark 2.13 (Assumption 2.9)**.**
Assumption 2.9 may not be satisfied for all sets of parameters being financially meaningful. Thus, the discretisation introduced so far does not have to be monotone.
Nevertheless, the smaller the bounds of the correlation are the bigger the set of volatilities we can deal with gets.

Remark 2.14.
If the correlation is zero, i. e. $\mathcal{C}_{1,2} = \{0\}$, the construction so far can be neglected. Just the following explanation how to combine the approximations for the second and first derivatives is needed to construct a monotone scheme.

Choosing the difference quotients. Up to now, the contribution of the second order and the mixed derivative to the off-diagonal entries is non-positive and the one to the diagonal entry is non-negative by construction of the grid and the structure of the difference quotient (2.6). In the next step we explain how to discretise the first derivatives in x_1-direction and x_2-direction. Since for the second derivative we always use (2.6), we have to ensure by the choice of the discretisation for the first derivative that the signs of the elements so far do not change. We only consider the variable x_1. For x_2 the discretisation is completely analogue.
If we use $\delta^f_{x_1} v^i_{k,l}$, the inequality

$$\frac{2}{h_1^{k+1} + h_1^k} d^{k,l}_{1,\gamma} + c^{k,l}_{1,\gamma} \leq 0 \tag{2.13a}$$

has to be satisfied to guarantee a non-positive coefficient $a^{k+1,l}_{k,l}$ of $v^i_{k+1,l}$ in the stencil (2.9). The factors in front of $d^{k,l}_{1,\gamma}$ and $c^{k,l}_{1,\gamma}$ result from the difference quotients (2.5a) and (2.6). Analogously, to use $\delta^b_{x_1} v^i_{k,l}$ the inequality

$$\frac{2}{h_1^{k+1} + h_1^k} d^{k,l}_{1,\gamma} - c^{k,l}_{1,\gamma} \leq 0 \tag{2.13b}$$

has to hold true. In both cases the coefficient $a_{k,l}^{k,l}$ of $v_{k,l}^{i}$ is by construction non-negative if $r \geq 0$, cp. (2.18).
We use central differencing in x_1-direction if

$$h_1^k c_{1,\gamma}^{k,l} + 2d_{1,\gamma}^{k,l} \leq 0 \quad \text{and} \quad -h_1^{k+1} c_{1,\gamma}^{k,l} + 2d_{1,\gamma}^{k,l} \leq 0 \tag{2.13c}$$

both are fulfilled. The following lemma gives a sufficient condition for (2.13c).

Lemma 2.1.
Let the conditions (2.13a) and (2.13b) hold true. Then the inequalities

$$Ca_{k,l,\gamma}^{k+1,l} = \frac{h_1^k}{h_1^{k+1}} c_{1,\gamma}^{k,l} + \frac{2}{h_1^k} d_{1,\gamma}^{k,l} \leq 0, \tag{2.14a}$$

$$Ca_{k,l,\gamma}^{k-1,l} = -\frac{h_1^{k+1}}{h_1^k} c_{1,\gamma}^{k,l} + \frac{2}{h_1^{k+1}} d_{1,\gamma}^{k,l} \leq 0, \quad \textit{and} \tag{2.14b}$$

$$Ca_{k,l,\gamma}^{k,l} = -\left(\frac{h_1^k}{h_1^{k+1}} - \frac{h_1^{k+1}}{h_1^k}\right) c_{1,\gamma}^{k,l} - \left(\frac{1}{h_1^k} + \frac{1}{h_1^{k+1}}\right) d_{1,\gamma}^{k,l} \geq 0 \tag{2.14c}$$

with $C = h_1^{k+1} + h_1^k$ are valid and centered difference quotients can be applied.

The inequalities (2.14a) and (2.14b) stem from the combination of $\delta_{x_1}^c v_{k,l}^i$ and $\delta_{x_1}^2 v_{k,l}^i$ representing the sign of $v_{k-1,l}^i$ and $v_{k+1,l}^i$, respectively. The last inequality (2.14c) represents the sign requirement for the coefficient $a_{k,l,\gamma}^{k,l}$ of $v_{k,l}^i$ if central differencing is utilised.

Proof.
With the choice of the forward difference quotient we have

$$(2.13a) \Rightarrow -\frac{2}{Ch_1^{k+1}} d_{1,\gamma}^{k,l} \geq \frac{h_1^k}{C} \frac{1}{h_1^{k+1}} c_{1,\gamma}^{k,l} \Rightarrow (2.14a),$$

since $d_{1,\gamma}^{k,l} \leq 0$, $\frac{h_1^k}{C} < 1$, and $C \in \mathbb{R}_+$.
The implication (2.13b) $\Rightarrow$ (2.14b) can be shown analogously. The converse statements do not have to be valid.
We can rewrite (2.14c) as two summands where each one of them is non-positive by (2.14a) and (2.14b) respectively.

□

To assure the maximal use of central differencing, we proceed as follows. If (2.14) is fulfilled, then central differencing is used. Otherwise, depending on the validity of (2.13a) and (2.13b) forward or backward differencing respectively is used.

The inequalities in x_2-direction can be derived by the same considerations as before. For completeness they are listed below. We use forward differencing if

$$\frac{2}{(h_2^l + h_2^{l+1})} d_{2,\gamma}^{k,l} + c_{2,\gamma}^{k,l} \leq 0, \tag{2.15a}$$

backward differencing if

$$\frac{2}{(h_2^l + h_2^{l+1})} d_{2,\gamma}^{k,l} - c_{2,\gamma}^{k,l} \leq 0, \quad \text{and} \tag{2.15b}$$

centered differencing if

$$h_2^l c_{2,\gamma}^{k,l} + 2d_{2,\gamma}^{k,l} \leq 0 \quad \text{and} \quad -h_2^{l+1} c_{2,\gamma}^{k,l} + 2d_{x,\gamma}^{k,l} \leq 0 \tag{2.15c}$$

hold true, respectively.
An analogue result to Lemma 2.1 holds true, too. Thus, the difference quotients are chosen in the same manner as in x_1-direction.

According to the results above we define the areas where the different difference quotients will be used. This is necessary since the choice of the optimal control and the used difference quotients depend on each other.

Definition 2.15 (Domain of backward / forward / centered differencing)**.**
Let a grid $\mathcal{G}(L, \{t^i\}, \mathcal{S}, \Omega_2) \subset \Omega_2^D$ be given. For a grid point (x_1^k, x_2^l) we define the set

$$\Theta_{cc}^{k,l} := \left\{ (\sigma_1, \sigma_2, \rho_{1,2}) \in \Theta_2 \mid 2d_{a,\gamma}^{k,l} \pm (h_a^b + h_a^{b+1}) c_{a,\gamma}^{k,l} \leq 0, (a,b) \in \{(1,k),(2,l)\} \right\} \tag{2.16}$$

as the set of controls where only central differencing is used. Analogously, the sets $\Theta_{bb}^{k,l}$ and $\Theta_{ff}^{k,l}$ are defined with "-" and "+" instead of "±" in the inequality, respectively.
For the use of different difference quotients in the space dimensions we set for example

$$\Theta_{bc}^{k,l} := \left\{ (\sigma_1, \sigma_2, \rho_{1,2}) \in \Theta_2 \mid 2d_{1,\gamma}^{k,l} - (h_1^k + h_1^{k+1}) c_{1,\gamma}^{k,l} \leq 0, 2d_{2,\gamma}^{k,l} \pm (h_2^l + h_2^{l+1}) c_{2,\gamma}^{k,l} \leq 0 \right\} \tag{2.17}$$

All other sets $\Theta_{bf}^{k,l}, \Theta_{cf}^{k,l}, \Theta_{cb}^{k,l}, \Theta_{fb}^{k,l}$, and $\Theta_{fc}^{k,l}$ for the different combinations of different quotients are defined in a similar way by guaranteeing the appropriate combination of inequalities.

Altogether, we have nine sets which form a partition of Θ. Of course, some of these sets may be nil. In Section 3 we will need these sets again for determining the optimal controls and for guaranteeing convergence of the methods. In Figure 2.2 an exemplary illustration of the different domains is plotted.

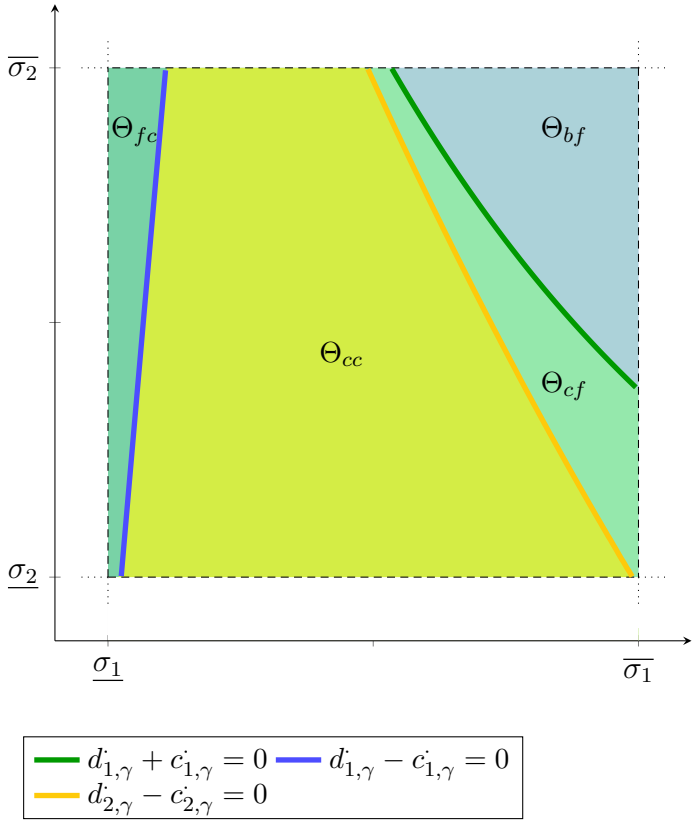

Fig. 2.2: *Domains of forward, backward, and centered differencing. For* $r = 0.035$, $\delta_1 = 0.019, \delta_2 = 0.01$, $\rho = 0.7$, $\mathcal{V}_1 = [0.16, 0.26]$, *and* $\mathcal{V}_2 = [0.23, 0.27]$.

Now we can give a complete definition of the seven point stencil (2.9) at an inner grid point $(\tau^i, x_1^k, x_2^l), k = 1, \ldots, M_1 - 1, l = 1, \ldots, M_2 - 1$. It is

$$a_{k,l,\gamma}^{k+1,l} = \frac{2d_{1,\gamma}^{k,l}}{(h_1^k + h_1^{k+1})h_1^{k+1}} + \begin{cases} \dfrac{c_{1,\gamma}^{k,l}}{h_1^{k+1}}, & (2.13a),\ \text{not } (2.13b), \\ 0, & (2.13b),\ \text{not } (2.13a) \\ \dfrac{c_{1,\gamma}^{k,l} h_1^k}{(h_1^k + h_1^{k+1})h_1^{k+1}}, & (2.13c), \end{cases} \tag{2.18a}$$

$$a_{k,l,\gamma}^{k-1,l} = \frac{2d_{1,\gamma}^{k,l}}{(h_1^k + h_1^{k+1})h_1^k} - \begin{cases} 0, & (2.13a),\ \text{not } (2.13b) \\ \dfrac{c_{1,\gamma}^{k,l}}{h_1^k}, & (2.13b),\ \text{not } (2.13a) \\ \dfrac{c_{1,\gamma}^{k,l} h_1^{k+1}}{(h_1^k + h_1^{k+1})h_1^k}, & (2.13c), \end{cases} \tag{2.18b}$$

$$a_{k,l,\gamma}^{k,l+1} = \frac{2d_{2,\gamma}^{k,l}}{(h_2^l + h_2^{l+1})h_2^{l+1}} + \begin{cases} \dfrac{c_{2,\gamma}^{k,l}}{h_2^{l+1}}, & (2.15a),\ \text{not } (2.15b), \\ 0, & (2.15b),\ \text{not } (2.15a) \\ \dfrac{c_{2,\gamma}^{k,l} h_2^l}{(h_2^l + h_2^{l+1})h_2^{l+1}}, & (2.15c), \end{cases} \tag{2.18c}$$

$$a_{k,l,\gamma}^{k,l-1} = \frac{2d_{2,\gamma}^{k,l}}{(h_2^l + h_2^{l+1})h_2^l} - \begin{cases} 0, & (2.15a),\ \text{not } (2.15b), \\ \dfrac{c_{2,\gamma}^{k,l}}{h_2^l}, & (2.15b),\ \text{not } (2.15a), \\ \dfrac{c_{2,\gamma}^{k,l} h_2^{l+1}}{(h_2^l + h_2^{l+1})h_2^l}, & (2.15c), \end{cases} \tag{2.18d}$$

$$a_{k,l,\gamma}^{k+1,l+1} = a_1^\gamma, \quad a_{k,l,\gamma}^{k-1,l-1} = a_2^\gamma, \quad \text{and} \tag{2.18e}$$

$$a_{k,l,\gamma}^{k,l} = r - a_{k,l,\gamma}^{k+1,l} - a_{k,l,\gamma}^{k-1,l} - a_{k,l,\gamma}^{k,l+1} - a_{k,l,\gamma}^{k,l-1} - a_{k,l,\gamma}^{k+1,l+1} - a_{k,l,\gamma}^{k-1,l-1}. \tag{2.18f}$$

The validity of the last equality for $a_{k,l,\gamma}^{k,l}$ can be checked by summing all coefficients for the distinct usage of forward, backward and central differencing.

Boundary conditions. For the boundary conditions we use the ansatz which is for example explained in the book of Randall and Tavella [RT00]. Accordingly, we assume that there is no contribution from the mixed derivatives at the grid points on the boundary of Ω_2^D. In this case equation (2.8a) reduces to

$$\begin{aligned} v_\tau - \frac{1}{2}\sigma_{1,\bar{j}}^2 \delta_{x_1}^2 v_{k,l}^i + \left(\frac{1}{2}\sigma_{1,\bar{j}}^2 - r + \delta_1\right) \delta_{x_1}^{\cdot} v_{k,l}^i \\ -\frac{1}{2}\sigma_{2,\bar{j}}^2 \delta_{x_2}^2 v_{k,l}^i + \left(\frac{1}{2}\sigma_{2,\bar{j}}^2 - r + \delta_2\right) \delta_{x_2}^{\cdot} v_{k,l}^i + r v_{k,l}^i = 0, \end{aligned} \tag{2.19}$$

where either $k = 0, M_1$ and $l \in J_{M_2}$ or $l = 0, M_2$ and $k \in J_{M_1}$. Since the boundary conditions depend on time, we will use the index i for time again.
We will suppose Neumann boundary conditions of the form

$$\frac{\partial v}{\partial x_j} = \beta_{k,l}^{i,j}, \tag{2.20}$$

where $j = 1, 2$ denotes the direction and the index (i, k, l) as usual the considered grid point. Our aim is to choose the difference quotients in (2.19) so that the

monotonicity of the discretisation is maintained. We will demonstrate that this can be achieved by equidistant central difference quotients across the boundary.
Let us take a closer look at a grid point with index $k = 0$ and $l = 1, \ldots, M_2 - 1$. By discretising the first derivative, we obtain

$$\frac{v^i_{1,l} - v^i_{-1,l}}{2h^1_1} = \beta^{i,1}_{0,l} \quad \Rightarrow \quad v^i_{-1,l} = v^i_{1,l} - 2h^1_1 \beta^{i,1}_{0,l}.$$

Here $v^i_{-1,l}$ is the value of the option at an artificial point left of the left boundary. The approximation of the second derivative becomes

$$\frac{v^i_{1,l} - 2v^i_{0,l} + v^i_{-1,l}}{\left(h^1_1\right)^2} = \frac{2}{\left(h^1_1\right)^2}\left(v^i_{1,l} - v^i_{0,l}\right) - \frac{2\beta^{i,1}_{0,l}}{h^1_1}.$$

By this construction the coefficients in the finite difference stencil (2.9) at the grid point (τ^i, x^0_1, x^l_2) are given by

$$\begin{aligned} &a^{1,l}_{0,l,\gamma} = 0, \\ &a^{0,l}_{0,l,\gamma} = r - a^{1,l}_{0,l,\gamma} - a^{1,l}_{0,l,\gamma} - a^{0,l+1}_{0,l,\gamma} - a^{0,l-1}_{0,l,\gamma}, \quad \text{and} \\ &a^{1,l+1}_{0,l,\gamma} = 0, \end{aligned}$$

where the unlisted coefficients $a^{0,l+1}_{0,l,\gamma}$ and $a^{0,l-1}_{0,l,\gamma}$ are defined as in (2.18) with zero correlation. The coefficients $a^{-1,l}_{0,l,\gamma}$ and $a^{-1,l-1}_{0,l,\gamma}$ of the artificial grid points are treated as zero. As we can see, the off-diagonal entries are non-positive and the diagonal entry is non-negative. The constant parts of the difference quotients are collected in a vector, see below.
For all other grid points of the grid the boundary conditions can be included in the same way. At the four corners of the grid the above performance is done in both directions. The off-diagonal entries are always non-positive and the diagonal entries are non-negative.
We define a vector $b^i_\gamma := \left(b^i_j\right)_j \in \mathbb{R}^M$ to completely incorporate the boundary conditions via

$$b^i_{k(M_2+1)+l} := \begin{cases} \left(2\frac{d^{0,l}_{1,\gamma,0}}{h^1_1} - c^{0,l}_{1,\gamma,0}\right)\beta^{i,1}_{0,l}, & k = 0, l = 1, \ldots, M_2 - 1, \\ -\left(2\frac{d^{M_1,l}_{1,\gamma,0}}{h^{M_1}_1} + c^{M_1,l}_{1,\gamma,0}\right)\beta^{i,1}_{M_1,l}, & k = M_1, l = 1, \ldots, M_2 - 1, \\ \left(2\frac{d^{k,0}_{2,\gamma,0}}{h^1_2} - c^{k,0}_{2,\gamma,0}\right)\beta^{i,2}_{k,0}, & k = 1, \ldots, M_1 - 1, l = 0, \\ -\left(2\frac{d^{k,M_2}_{2,\gamma,0}}{h^{M_2}_2} + c^{k,M_2}_{2,\gamma,0}\right)\beta^{i,2}_{k,M_2}, & k = 1, \ldots, M_1 - 1, l = M_2 - 1 \end{cases} \tag{2.21a}$$

on the edges except the skews,

$$b_0^i := \left(2\frac{d_{1,\gamma,0}^{0,0}}{h_1^1} - c_{1,\gamma,0}^{0,0}\right)\beta_{0,0}^{i,1} + \left(2\frac{d_{2,\gamma,0}^{0,0}}{h_2^1} - c_{2,\gamma,0}^{0,0}\right)\beta_{0,0}^{i,2}, \tag{2.21b}$$

$$b_{M_2}^i := \left(2\frac{d_{1,\gamma,0}^{0,M_2}}{h_1^1} - c_{1,\gamma,0}^{0,M_2}\right)\beta_{0,M_2}^{i,1} - \left(2\frac{d_{2,\gamma,0}^{0,M_2}}{h_2^{M_2}} + c_{2,\gamma,0}^{0,M_2}\right)\beta_{0,M_2}^{i,2}, \tag{2.21c}$$

$$b_{M-M_2-1}^i := -\left(2\frac{d_{1,\gamma,0}^{M_1,0}}{h_1^{M_1}} + c_{1,\gamma,0}^{M_1,0}\right)\beta_{M_1,0}^{i,1} + \left(2\frac{d_{2,\gamma,0}^{M_1,0}}{h_2^1} - c_{2,\gamma,0}^{M_1,0}\right)\beta_{M_1,0}^{i,2}, \tag{2.21d}$$

$$b_{M-1}^i := -\left(2\frac{d_{1,\gamma,0}^{M_1,M_2}}{h_1^{M_1}} + c_{1,\gamma,0}^{M_1,M_2}\right)\beta_{M_1,M_2}^{i,1} - \left(2\frac{d_{2,\gamma,0}^{M_1,M_2}}{h_2^{M_2}} + c_{2,\gamma,0}^{M_1,M_2}\right)\beta_{M_1,M_2}^{i,2}, \tag{2.21e}$$

on the skews and

$$b_{\bar{j}}^i = 0 \tag{2.21f}$$

at all inner grid points. Here we denote with $c_{1,\gamma,0}^{k,l}$, etc. the coefficients defined in (2.18) with the correlation set to zero.

For the Betas we fall back on the Deltas of the options which can be determined analytically for certain types of options. For the maximum and minimum options the $\beta_{k,l}^{i,1}$ and $\beta_{k,l}^{i,2}$ for $(\tau^i, x_1^k, x_2^l) \in \partial\mathcal{G}$ are calculated via the analytic formulas of the multi-dimensional Black-Scholes model. In the paper of Stulz [Stu82] formulas of this type are given for the option value. For options on more than two assets the results of Stulz have been extended by Johnson [Joh87]. Beginning with these Reiß and Wystup [RW01] derive closed form expressions for the Greeks of these options. The Delta is given by

$$\frac{\partial V(t,S_1,S_2)}{\partial S_j} = \phi\exp(-\delta_j(T-t))\mathcal{N}_2\left(\phi d_{p_1}, \eta d_{p_2}, \phi\eta d_{p_3}\right), \quad p_1,p_2,p_3 \in \{1,\ldots,6\}$$

where $\mathcal{N}_2(x,y,z)$ is the distribution function of the multi-variate normal distribution with mean zero, variances one, and correlation z. The constants $d_p, p = 1,\ldots,6$, depend on the particular choice of the option. For more details regarding those and the numerical evaluation of the distribution function $\mathcal{N}_2$ see Appendix A.2.
The option's Delta is essentially influenced by volatility and correlation which are incorporated into the constants d_p. Thus, every $\beta_{k,l}^{i,j}$ is calculated with the corresponding control $(\sigma_{1,\bar{j}}, \sigma_{2,\bar{j}}, \rho_{1,2,\bar{j}})$. The Neumann boundary conditions for maximum

and minimum options are now realised by setting

$$\beta_{k,l}^{i,j} = \frac{\partial V(t^i, S_1, S_2)}{\partial S_j} S_j, \quad j = 1, 2. \tag{2.22}$$

The price of a Geometric Average call option equals the price of a Vanilla call in the standard Black-Scholes model with modified volatility $\tilde{\sigma}$ and dividend rate $\tilde{\delta}$:

$$\tilde{\sigma}^2 := \frac{1}{4} \sum_{i,j=1}^{n} \sigma_i \sigma_j \rho_{i,j} \quad \text{and} \quad \tilde{\delta} := \frac{1}{2} \sum_{i=1}^{n} \left(\delta_i + \frac{1}{2} \sigma_i^2 \right) - \frac{1}{2} \tilde{\sigma}^2, \tag{2.23}$$

cf. [LO08]. The one-dimensional space variable is given by:

$$y = \prod_{i=1}^{n} \exp\left(\frac{S_i}{n} \right)$$

Applying the chain rule of differentiation gives a formula for the Delta of Geometric Average call option

$$\frac{\partial V(t, S_1, S_2)}{\partial S_j} = \phi \exp(-\tilde{\delta}(T-t)) \frac{y}{n} \mathcal{N}_1(\phi d_1),$$

where d_1 depends on the volatility and the dividend rate given in (2.23). The constant $\phi \in \{-1, 1\}$ is chosen according to the valuation of a call or a put, respectively. For more details see section A.3.
Algorithms for the evaluation of the (multi-variate) standard normal distribution function are given in Section A.1.

Remark 2.16.
The constants $d_{p_1}, d_{p_2}, d_{p_3}$ with $p_1, p_2, p_3 \in \{1, \ldots, 6\}$ depend on the control, cp. Appendix A.1. During the determination of the optimal control, cp. Chapter 3, the actual value of the optimal control will vary. To assure that the determination of the controls converges, we do not recalculate the boundary conditions for every variation, cp. the proof of Theorem 3.1. Instead we use the controls of the previous time step.

In Table 2.1 an overview of the different boundary conditions we use in the numerical experiments at the different boundaries is given.

Here c_0, c_1, c_2 denote the appropriately chosen discount factors and $\mathcal{N}_2(\cdot)$ the two-dimensional cumulative normal distribution function with corresponding arguments for the specific option. With Neumann boundary conditions the option value is not

Tab. 2.1: *Overview of the boundary conditions for different options on two assets.*

option	$\beta^{i,1}_{0,l}$	$\beta^{i,1}_{M_1,l}$	$\beta^{i,2}_{k,0}$	$\beta^{i,2}_{k,M_2}$
Butterfly Spr.	0	0	0	0
Min. put	$-c_1\mathcal{N}_2(\cdot)$	0	$-c_2\mathcal{N}_2(\cdot)$	0
Min. call	0	$c_1\mathcal{N}_2(\cdot)$	0	$c_2\mathcal{N}_2(\cdot)$
Min. put cap.	0	0	0	0
Max. put	$-c_1\mathcal{N}_2(\cdot)$	0	$-c_2\mathcal{N}_2(\cdot)$	0
Max. call	0	$c_1\mathcal{N}_2(\cdot)$	0	$c_2\mathcal{N}_2(\cdot)$
Max. call cap.	0	0	0	0
Geom. Av. put	$-c_0\frac{y}{2}\mathcal{N}_1(-d_1)$	$-c_0\frac{y}{2}\mathcal{N}_1(-d_1)$	$-c_0\frac{y}{2}\mathcal{N}_1(-d_1)$	$-c_0\frac{y}{2}\mathcal{N}_1(-d_1)$
Geom. Av. call	$c_0\frac{y}{2}\mathcal{N}_1(d_1)$	$c_0\frac{y}{2}\mathcal{N}_1(d_1)$	$c_0\frac{y}{2}\mathcal{N}_1(d_1)$	$c_0\frac{y}{2}\mathcal{N}_1(d_1)$

invariant to translations by a constant. To fix this issue we use Dirichlet boundary conditions at one edge of the computational domain Ω_2^D where the payoff is zero. This is possible for all types of options considered in this thesis. Numerically the Dirichlet boundary condition is easily implemented by setting $(A^i_\gamma)_{\overline{j}} = e_{\overline{j}}$ where $\overline{j}$ is the lexicographic index of the corresponding edge.

Matrix representation. The representation of equation (2.8) for all grid points can now be constructed as a tridiagonal block matrix $A^i_\gamma \in \mathbb{R}^{M\times M}$:

$$A^i_\gamma := \begin{pmatrix} A^c_{0,\gamma} & A^u_{0,\gamma} & 0 & & & \cdots & 0 \\ A^d_{1,\gamma} & A^c_{1,\gamma} & A^u_{1,\gamma} & 0 & & & \\ 0 & \ddots & \ddots & \ddots & & & \\ & & & & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & 0 \\ & & & 0 & A^d_{M_1-1,\gamma} & A^c_{M_1-1,\gamma} & A^u_{M_1-1,\gamma} \\ 0 & & \cdots & & 0 & A^d_{M_1,\gamma} & A^c_{M_1,\gamma} \end{pmatrix}, \tag{2.24a}$$

where each matrix $A^c_{k,\gamma}, A^d_{k,\gamma}, A^u_{k,\gamma} \in \mathbb{R}^{(M_2+1)\times(M_2+1)}$ itself is of tridiagonal structure:

$$A^c_{k,\gamma} := \begin{pmatrix} a^{k,0}_{k,0,\gamma} & a^{k,1}_{k,0,\gamma} & 0 & & & \cdots & 0 \\ a^{k,0}_{k,1,\gamma} & a^{k,1}_{k,1,\gamma} & a^{k,2}_{k,1,\gamma} & 0 & & & \\ 0 & \ddots & \ddots & \ddots & & & \\ & & & & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & 0 \\ & & & 0 & a^{k,M_2-2}_{k,M_2-1,\gamma} & a^{k,M_2-1}_{k,M_2-1,\gamma} & a^{k,M_2}_{k,M_2-1,\gamma} \\ 0 & & \cdots & & 0 & a^{k,M_2-1}_{k,M_2,\gamma} & a^{k,M_2}_{k,M_2,\gamma} \end{pmatrix}, k \in J_{M_1}, \tag{2.24b}$$

$$A^u_{k,\gamma} := \begin{pmatrix} a^{k+1,0}_{k,0,\gamma} & a^{k+1,1}_{k,0,\gamma} & 0 & & \cdots & 0 \\ 0 & a^{k+1,1}_{k,1,\gamma} & a^{k+1,2}_{k,1,\gamma} & 0 & & \\ 0 & \ddots & \ddots & \ddots & & \\ & & & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ & & & 0 & a^{k+1,M_2-1}_{k,M_2-1,\gamma} & a^{k+1,M_2}_{k,M_2-1,\gamma} \\ 0 & \cdots & & 0 & 0 & a^{k+1,M_2}_{k,M_2,\gamma} \end{pmatrix}, \; k \in J_{M_1-1} \tag{2.24c}$$

$$A^d_{k,\gamma} := \begin{pmatrix} a^{k-1,0}_{k,0,\gamma} & 0 & & & \cdots & 0 \\ a^{k-1,0}_{k,1,\gamma} & a^{k-1,1}_{k,1,\gamma} & 0 & & & \\ 0 & \ddots & \ddots & \ddots & & \\ & & & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \\ & & 0 & a^{k-1,M_2-2}_{k,M_2-1,\gamma} & a^{k-1,M_2-1}_{k,M_2-1,\gamma} & 0 \\ 0 & \cdots & & 0 & a^{k-1,M_2-1}_{k,M_2,\gamma} & a^{k-1,M_2}_{k,M_2,\gamma} \end{pmatrix}, \; k \in I_{M_1} \tag{2.24d}$$

We remind the reader that for ease of notation we omit the indices i in the tridiagonal matrices.

By the construction performed before we have proven the following

Lemma 2.2.

Let $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_2)$ be a grid where Assumption 2.9 is satisfied. Let the space discretisation be given by (2.18) and the corresponding boundary conditions (2.20)*. Then the matrix A^i_γ in (2.24) has non-positive off-diagonal entries and non-negative diagonal entries for all $\gamma^i \in \Theta_2^{M^i}$ and $i \in I_L$, if $r \geq 0$ holds.*

Remark 2.17.

The matrix A^i_γ is a sparse matrix. In the case of non-zero correlation it has seven respectively nine diagonals unequal to zero. Otherwise there are only five diagonals unequal to zero, s. (2.24)*.*

This fact can be exploited for the implementation of the final algorithm. It also influences the choice of the solver for the system of equations. A direct solver would possibly fill up the zero entries of the matrix and thus require much more memory. Iterative solvers instead take advantage of the sparse structure and are therefore more efficient.

Remark 2.18.

For the discretisation of the BSB equation (BSB_n) we assumed w. l. o. g. that $\mathcal{C}_{1,2}$ is

a subset of the positive real figures. Analogously to the construction above, we would also obtain a discretisation matrix with seven diagonals for the case $\mathcal{C}_{1,2} \subset \mathbb{R}_-$.
In the most general case $\mathcal{C}_{1,2} \subset \mathbb{R}$*, we possibly have to switch between the different quotients* $\delta^+_{x_1,x_2}$ *and* $\delta^-_{x_1,x_2}$*. As a consequence the discretisation matrix might have up to nine non-zero diagonals. But they still fulfil the requirements regarding the sign and Lemma 2.2 thus still applies.*

2.2.2 Choice of step sizes and construction of non-uniform monotone grids

In this section we analyse how the choice of the grid a priori influences the use of the different difference quotients in the discretisation. We consider grids with uniform spatial step sizes in every space direction. Furthermore, we point out the special case for zero correlation.
This analysis is important for the final algorithm, since we learn how to determine a priori which approach we have to choose for the optimisation part, cp. Section 3.2. Therefore we are interested in bounds as close as possible.
The second part of this section gives a procedure how a monotone non-uniform grid can be constructed. If it is thoroughly constructed the accuracy of the resulting option value may be increased meanwhile keeping the number of sampling points constant.

Bounds for the step widths. In the following all results are formulated for only one spatial dimension. Of course, they are of general nature and can be applied for any direction with the obvious changes. The basis for the following lemmas and corollaries are the conditions for the combination of difference quotients given in Section 2.2.1.

Lemma 2.3.
Let a monotone grid $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_2)$ *with uniform spatial step sizes in each direction be given, that is* $h_j^1 = \ldots = h_j^{M_j}, j = 1, 2$.
Then only centered difference quotients can be used for the spatial discretisation if for

i) $r \leq \delta_1$ *the inequality*

$$h_1^i \leq \min_{x \in \mathcal{V}_1} \left\{ \frac{x^2 - x\Lambda}{\frac{1}{2}x^2 - r + \delta_1} \right\}, \tag{2.25a}$$

ii) $r > \delta_1$ *the inequality*

$$h_1^{\cdot} \leq \min_{x \in B} \left\{ \frac{x^2 - x\Lambda}{\max\{\left|\frac{1}{2}x^2 - r + \delta_1\right|, \xi\}} \right\}, \quad \textit{where} \tag{2.25b}$$

$$B := \{\underline{\sigma_1}, \overline{\sigma_1}\} \cup \left(\{z_1, \sqrt{2(r-\delta_1)^+} \pm 0.02(\Lambda)^+\} \cap \mathcal{V}_1 \cap B_1 \right),$$

$$B_1 := \left[\sqrt{2(r-\delta_1)^+} + 0.02(\Lambda)^+, \sqrt{2(r-\delta_1)^+} - 0.02(\Lambda)^+ \right],$$

respectively, holds true.
The values in the above inequalities are given by:

$$\Lambda := \frac{\Delta}{2}\overline{\rho_{1,2}\sigma_2}, \qquad \Delta := \frac{h_1^1}{h_2^1},$$

$$z_1 := \frac{\Lambda}{2},$$

$$\xi := \frac{1}{h_1^1} \min\left\{ \underline{\sigma_1}^2 - \Lambda\underline{\sigma_1}, \overline{\sigma_1}^2 - \Lambda\overline{\sigma_1}, 2\Lambda \right\}$$

Proof.
Since the grid is assumed to be monotone we have $d_{1,\gamma}^{k,l} \leq 0$*, cp. Section 2.2.1. The proof is done by showing that* (2.13c) *holds true for the conditions* (2.25)*. We start the analysis by noting that for all* $\sigma_2 \in \mathcal{V}_2$ *and* $\rho_{1,2} \in \mathcal{C}_{1,2}$ *we have*

$$h_1^{\cdot}\left(\frac{1}{2}\sigma_1^2 - r + \delta_1\right) - \sigma_1^2 + \frac{\Delta}{2}\overline{\rho_{1,2}\sigma_2}\sigma_1 \leq 0$$

$$\Rightarrow \quad h_1^{\cdot}\left(\frac{1}{2}\sigma_1^2 - r + \delta_1\right) - \sigma_1^2 + \frac{\Delta}{2}\rho_{1,2}\sigma_2\sigma_1 \leq 0$$

since the grid is uniform. Which is why we focus on the first inequality.
Let us start with case i). And let an index k *be fixed. We have* $c_{1,\gamma}^{k,l} = \frac{1}{2}\sigma_1^2 - r + \delta_1 > 0$ *for all* $\sigma_1 \in \mathcal{V}_1$*. Thus, the inequality*

$$h_1^{\cdot} \leq \frac{\sigma_1^2 - \frac{\Delta}{2}\overline{\rho_{1,2}\sigma_2}\sigma_1}{\frac{1}{2}\sigma_1^2 - r + \delta_1}$$

has to hold. Thus, we directly obtain the desired appraisal.

Case ii) is a little more elaborate. We cannot a priori determine the sign of $c_{1,\gamma}^{k,l}$*. But the conditions* (2.13c) *are symmetric in the sense that we may pass over to*

$$h_1^{\cdot}\left|\frac{1}{2}\sigma_1^2 - r + \delta_1\right| - \sigma_1^2 + \frac{\Delta}{2}\rho_{1,2}\sigma_2\sigma_1 \leq 0$$

Moreover, dividing by $|c_{1,\gamma}^{k,l}|$ would include a dispensable pole into our considerations. If ξ is defined as above we have a non-negative lower bound to $|c_{1,\gamma}^{k,l}|$ for which the above inequality holds. We thus pass over to

$$h_1^{\cdot} \max\left\{\left|\frac{1}{2}\sigma_1^2 - r + \delta_1\right|, \xi\right\} - \sigma_1^2 + \frac{\Delta}{2}\rho_{1,2}\sigma_2\sigma_1 \leq 0$$

and obtain as upper bound for $h_1^{\cdot}$, which depends on σ_1,

$$h_1^{\cdot} \leq \frac{\sigma_1^2 - \frac{\Delta}{2}\overline{\rho_{1,2}\sigma_2}\sigma_1}{\max\left\{\left|\frac{1}{2}\sigma_1^2 - r + \delta_1\right|, \xi\right\}}$$

The minimum for the right hand side of the above inequality can possibly be attained on $\partial\mathcal{V}_1$. Or it is attained, for ξ larger than $|c_{1,\gamma}^{k,l}|$, at the point z_1 if $z_1 \in B_1$ or at the boundaries of B_1 given they lie in $\mathcal{V}_1$. Which gives us the result we looked for.

□

Lemma 2.3 gives us a simple criterion to guarantee the exclusive use of centered difference quotients with just the knowledge of the ratio of the step sizes towards each other. The ratio Δ can be appraised from above by the constant $\Xi_{1,2}^u$. Furthermore ξ can be reduced by replacing h_1^0 by the larger constant H. Thus, the bounds can be made completely independent from any monotone grid.

As a direct consequence of Lemma 2.3 we can derive

Corollary 2.4.
Let $\mathcal{C}_{1,2} = \{0\}$ and let a monotone grid $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_2)$ be given. The exclusive use of central differencing in the discretisation given by (2.18), (2.24), and (2.21) results in a discretisation with non-positive diagonal entries and non-negative off-diagonal entries, if for $\mu = 1, 2$ for all $\nu \in I_{M_\mu}$

$$h_\mu^\nu \leq \begin{cases} \min\limits_{x\in\{\underline{\sigma_\mu},\overline{\sigma_\mu}\}} \dfrac{x^2}{\frac{1}{2}x^2 - r + \delta_\mu}, & r > \delta_\mu \\ \min\limits_{x\in\{\underline{\sigma_\mu},\overline{\sigma_\mu}\}} \dfrac{x^2}{\max\left\{\left|\frac{1}{2}x^2 - r + \delta_\mu\right|, \frac{{\sigma_\mu}^2}{h_\mu^\nu}\right\}}, & \textit{else} \end{cases} \tag{2.26}$$

holds true.

Remark 2.19.
The above lemma can be considered as a generalisation of a similar result proven by Schaeling and Heider [SH11] for non-linear one-dimensional equations on an

equidistant grid, compare also section 2.2.4.

For a general non-uniform grid we can deduce a similar result, which due to the more general approach is not as pleasant as the one of Lemma 2.3.

Lemma 2.5.
Let a non-uniform monotone grid $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_2)$ be given.
Then only centered difference quotients can be used in x_1-direction, if the following inequality is fulfilled for all $k \in J_{M_1-1}$:

$$h_1^k \leq \min_{x \in C} \frac{x^2 - x\Lambda^k}{\max\{\left|\frac{1}{2}x^2 - r + \delta_1\right|, \xi\} + x\Upsilon}, \quad \textit{with} \tag{2.27}$$

$$C := \left\{\underline{\sigma_1}, \overline{\sigma_1}\right\} \cup \left(\{z_1, \sqrt{2(r-\delta_1)^+} \pm 0.02(\Lambda)^+\} \cap \mathcal{V}_1 \cap C_1\right),$$

$$C_1 := \left[\sqrt{2(r-\delta_1)^+} + 0.02(\Lambda)^+, \sqrt{2(r-\delta_1)^+} - 0.02(\Lambda)^+\right],$$

where

$$\Lambda^k := \frac{\Delta^k}{2}\overline{\rho_{1,2}\sigma_2}, \qquad \Delta^k := \max_{l \in J_{M_2-2}} \left\{\frac{h_1^k}{h_2^l} + \frac{h_1^{k+1}}{h_2^{l+1}}\right\},$$

$$\Lambda := \max_k \Lambda^k \qquad \Delta := \max_k \Delta^k$$

$$\Upsilon := \frac{1}{2}(\overline{\rho_{1,2}\sigma_2})^+ \max_{l \in J_{M_2-1}} \left|\frac{1}{h_2^l} - \frac{1}{h_2^{l+1}}\right|, \ \textit{and}$$

z_1 and ξ as in Lemma 2.3.

Proof.
Centered difference quotients are used, if (2.13c) *is assured. Due to the dependence on the step widths in x_2- direction we start by appraising the coefficients $d_{1,\gamma}^{k,l}$ and $c_{1,\gamma}^{k,l}$. It is*

$$\begin{aligned} d_{1,\gamma}^{k,l} &= -\frac{1}{2}\sigma_1^2 + \frac{1}{4}\rho_{1,2}\sigma_1\sigma_2\left(\frac{h_1^k}{h_2^l} + \frac{h_1^{k+1}}{h_2^{l+1}}\right) \\ &\leq -\frac{1}{2}\sigma_1^2 + \frac{1}{4}\sigma_1\overline{\rho_{1,2}\sigma_2} \max_{l \in J_{M_2-2}} \left\{\frac{h_1^k}{h_2^l} + \frac{h_1^{k+1}}{h_2^{l+1}}\right\} \leq 0 \quad \textit{and} \\ \left|c_{1,\gamma}^{k,l}\right| &= \left|\frac{1}{2}\sigma_1^2 - r + \delta_1 - \frac{1}{2}\rho_{1,2}\sigma_1\sigma_2\left(\frac{1}{h_2^l} - \frac{1}{h_2^{l+1}}\right)\right| \\ &\leq \left|\frac{1}{2}\sigma_1^2 - r + \delta_1\right| + \frac{1}{2}\sigma_1\overline{\rho_{1,2}\sigma_2} \max_{l \in J_{M_2-2}} \left|\frac{1}{h_2^l} - \frac{1}{h_2^{l+1}}\right| \end{aligned}$$

We still have $d_{1,\gamma}^{k,l} \leq 0$ *due to the appraisal* $\frac{h_1^k}{h_2^l} \leq \Xi_{1,2}^u$ *for all step widths.*
Using the above estimates we conclude analogously to the proof of Lemma 2.3, that

$$h_1^k \leq \frac{\sigma_1^2 + \Lambda^k \sigma_1}{\left|\frac{1}{2}\sigma_1^2 - r + \delta_1\right| + \sigma_1 \Upsilon}$$

The minimum value of the right hand side either is attained at the boundaries of $\mathcal{V}_1$*, at* z_1 *for* $\mathcal{V}_1 \cap C_1$*, or at the boundaries of* C_1*, if they lie in* $\mathcal{V}_1$*. Which concludes the argumentation.*

□

The bounds for h_1^k given in the above lemma are not as simple as those of Lemma 2.3. The maximum term in Δ can be estimated from above by $2\Xi_{1,2}^u$. But for the term contained in Υ it is difficult to give an appropriate bound a priori. Nevertheless, given a monotone non-uniform grid we can verify if only centered difference quotients can be used. Especially for the case $\mathcal{C}_{1,2} = \{0\}$ the bound (2.27) becomes much more less restrictive and non-uniform grids can easily be used, cp. Chapter 5.

Non-uniform monotone grid. The aim of this passage is to develop a method for constructing a non-uniform monotone grid. Here we perform the construction for $d = 2$. For higher dimensional problems the method can be generalised in an intuitive way.
The sampling points should be distributed in such a way that near a specified point, mostly the strike price, they lie dense and moving towards the boundary, their distance increases. The structure of the grid will be as follows

1. small, constant step sizes near the strike price,
2. increasing steps moving towards the boundary, and
3. larger, constant step sizes close to the boundary.

The latter point should be satisified in order to preserve the monotonicity of the grid, that is to assure inequality (2.12) to hold true, see below.

Before we start, we assume w. l. o. g. $x_\mu^{min} \leq 0 \leq x_\mu^{max}$.

Let a set of indices J_{M_μ} be given. The idea of constructing a non-uniform grid is to assign to each index $\nu \in J_{M_\mu}$ a grid point x_μ^ν via a function $\xi_\mu(\nu) = x_\mu^\nu, \nu \in J_{M_\mu}$. The set $\{x_\mu^\nu\}_{\nu \in J_{M_\mu}}$ forms the non-uniform grid in x_μ-direction. Let minimal step

sizes h_1^{min}, h_2^{min} which satisfy

$$\begin{aligned} &\Xi_{1,2}^d h_2^{max} \leq h_1^{min} \leq h_1^{max} \leq \Xi_{1,2}^u h_2^{min} \quad \text{and} \\ &\frac{1}{\Xi_{1,2}^u} h_1^{max} \leq h_2^{min} \leq h_2^{max} \leq \frac{1}{\Xi_{1,2}^d} h_1^{min} \end{aligned} \tag{2.28}$$

be given, cp. Section 2.2.2.
The following example illustrates two possible choices.

Example 2.20.
The inequalities (2.28) *are satisfied for*

1. $h_1^{min} = h_2^{min}$ *chosen freely and* $h_1^{max} = \Xi_{1,2}^u h_2^{min}$, $h_2^{max} = \frac{1}{\Xi_{1,2}^d} h_1^{min}$
2. h_1^{min} *chosen freely and* $h_1^{max} = \frac{\Xi_{1,2}^u}{\Xi_{1,2}^d} h_1^{min}$. *We then deduce from the second inequality of* (2.28) *that* $h_2^{min} = h_2^{max} = \frac{1}{\Xi_{1,2}^d} h_1^{min}$.

Example 2.20 illustrates in which strong manner the choice of spatial step sizes depends on the ratio $\frac{\Xi_{1,2}^u}{\Xi_{1,2}^d}$. We saw that for a two-dimensional grid the (restricted) freedom of choice can either be concentrated on only one spatial dimension or be distributed equally on both.

To obtain the desired structure of the grid mentioned above, we divide $[x_\mu^{min}, x_\mu^{max}]$ into subintervals :

$$[x_\mu^{min}, x_\mu^{max}] = [x_\mu^{min}, -dx_\mu^+] \cup [-dx_\mu^+, -x_\mu^+] \cup [-x_\mu^+, x_\mu^+] \cup [x_\mu^+, dx_\mu^+] \cup [dx_\mu^+, x_\mu^{max}],$$

with $\frac{2x_\mu^+}{h_\mu^{min}} \in \mathbb{N}$ and $d \in \mathbb{R}, d > 1$. The number of grid points is hereby counted cumulatively starting with N_μ^1 for the most left interval up to $N_\mu^5 = M_\mu$ for the complete interval $[x_\mu^{min}, x_\mu^{max}]$.

For the central interval we define:

$$\xi_\mu(\nu) := h_\mu^{min}(\nu - N_\mu^2) - x_\mu^+, \quad N_\mu^2 \leq \nu \leq N_\mu^3,$$

where N_μ^2 and N_μ^3 will be defined below.
For the increasing step sizes on $[x_\mu^+, dx_\mu^+]$ we choose the $N_\mu^4 - N_\mu^3$ sampling points according to

$$\xi_\mu(\nu) = x_\mu^+ + h_\mu^{min} \left(\nu - N_\mu^3 + \xi_c \sum_{\iota=1}^{\nu - N_\mu^3} \iota \right), \quad N_\mu^3 < \nu \leq N_\mu^4,$$

where ξ_c is chosen to satisfy for all $N_\mu^3 < \nu \leq N_\mu^4$

$$x_\mu^\nu - x_\mu^{\nu-1} = h_\mu^{min}\left(\nu - N_\mu^3 + \xi_c \sum_{\iota=1}^{\nu-N_\mu^3} \iota\right) - h_\mu^{min}\left(\nu - 1 - N_\mu^3 + \xi_c \sum_{\iota=1}^{\nu-1-N_\mu^3} \iota\right) \overset{!}{\leq} h_\mu^{max}$$
$$\Rightarrow \quad \xi_c \overset{!}{\leq} \left(\frac{h_\mu^{max}}{h_\mu^{min}} - 1\right) \frac{1}{N_\mu^4 - N_\mu^3 + 1}$$

As a result we obtain the next part of the sampling point function $N_\mu^3 < \nu \leq N_\mu^4$

$$\xi_\mu(\nu) := x_\mu^+ + h_\mu^{min}\left(\nu - N_\mu^3 + \xi_c \sum_{\iota=1}^{\nu-N_\mu^3} \iota\right)$$

Next, for the most right interval we apply the approach

$$\xi_\mu(\nu) = h_\mu^{max}\left(\nu - N_\mu^4\right) + \xi_\mu(N_\mu^4), \quad N_\mu^4 < \nu \leq N_\mu^5.$$

Finally, the number of grid points for each subinterval has to be calculated. We have

$$\begin{aligned} N_\mu^1 &= \left\lfloor \frac{-dx_\mu^+ - x_\mu^{min}}{h_\mu^{max}} \right\rfloor + 1 & N_\mu^5 - N_\mu^4 &= \left\lfloor \frac{x_\mu^{max} - dx_\mu^+}{h_\mu^{max}} \right\rfloor + 1 \\ N_\mu^3 - N_\mu^2 &= \frac{2x_\mu^+}{h_\mu^{min}} & N_\mu^2 - N_\mu^1 &= N_\mu^4 - N_\mu^3 = 2\left\lceil \frac{(d-1)x_\mu^+}{h_\mu^{max} - h_\mu^{min}} \right\rceil, \end{aligned} \tag{2.29}$$

where the last equation is derived from the requirement to span the whole interval $[x_\mu^+, dx_\mu^+]$. Which is guaranteed if the inequality

$$dx_\mu^+ - x_\mu^+ \leq h_\mu^{min}\left(n + \xi_c \frac{n(n+1)}{2}\right)$$

where n is the quantity in question, holds.
For the intervals $[-dx_\mu^+, -x_\mu^+]$ and $[x_\mu^{min}, -dx_\mu^+]$ we use the analogous construction as for their positive counterparts. Summing up, a sampling point function for building

a non-uniform grid in x_μ-direction is given by:

$$\xi_\mu(\nu) = \begin{cases} -h_\mu^{max}\left(N_\mu^1 - \nu\right) + \xi_\mu(N_\mu^1), & 0 \le \nu < N_\mu^1 \\ -x_\mu^+ - h_\mu^{min}\left(N_\mu^2 - \nu + \xi_c \sum_{\iota=1}^{N_\mu^2-\nu} \iota\right), & N_\mu^1 \le \nu < N_\mu^2 \\ h_\mu^{min}(\nu - N_\mu^2) - x_\mu^+, & N_\mu^2 \le \nu \le N_\mu^3 \\ x_\mu^+ + h_\mu^{min}\left(\nu - N_\mu^3 + \xi_c \sum_{\iota=1}^{\nu-N_\mu^3} \iota\right), & N_\mu^3 < \nu \le N_\mu^4 \\ h_\mu^{max}\left(\nu - N_\mu^4\right) + \xi_\mu(N_\mu^4), & N_\mu^4 < \nu \le N_\mu^5 \end{cases} \tag{2.30}$$

Remark 2.21.
The above formula for a non-uniform grid might yield grid points which might not coincide with x_μ^{min} and x_μ^{max}. This mismatch can easily be handled by adjusting the appropiate step sizes.

Remark 2.22.
The non-uniform grid has a nice property by construction. If h_1^{min} is halved, then all other step sizes are also scaled in the same way. This fact can be seen directly by the definition of ξ_μ in (2.30)*. Accordingly, the number of sampling points is doubled, cp.* (2.29)*.*
Of course, if h_1^{min} is halved, this also holds for the step sizes in x_2-direction, cp. Example 2.20.

In order to actually build the grid, we have to provide several parameters. One possible triple is the number of steps on minimal size $N_\mu^3 - N_\mu^2$, the boundary of the central interval x^+, and the growth factor of the intervals d. All other parameters are then intrinsically determined by the choice of the maximal step size and the construction of the grid.

2.2.3 Options on three and more assets

In this section we summarise the most important facts for the discretisation of the BSB equation (BSB_n) for at least three space dimensions. The results given here can be derived from those of Section 2.2.1, since the task can be decomposed into several two-dimensional ones for which the same argumentation applies.
The construction performed for the two-dimensional case in Section 2.2.1 has now

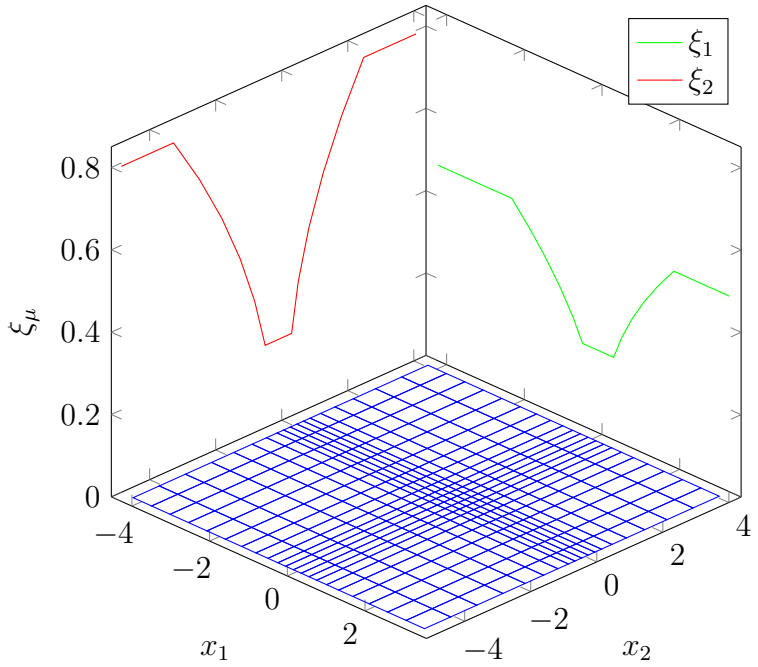

Fig. 2.3: *Example of a two-dimensional non-uniform monotone grid with $M_1 = 21$ and $M_2 = 18$ points. The bounds are $\Xi^u_{1,2} = 2.5$ and $\Xi^d_{1,2} = 0.24$. The sampling point functions ξ_1 and ξ_2 are plotted on the back of the graph.*

to be done for every tuple $(x_i, x_j), i \neq j, 1 \leq i, j \leq n$ of space variables. Accordingly, we have to make

Assumption 2.23.
For every tuple $(i, j), i \neq j, 1 \leq i, j \leq n$, we assume that either

1. *for a uniform grid the correlation matrix is diagonally dominant, or*

2. *for an arbitrary non-uniform grid the inequality*

$$(n-1)\Xi^d_{i,j} \leq \frac{h^\mu_i}{h^\nu_j} \leq \frac{1}{(n-1)}\Xi^u_{i,j}, \quad i < j \tag{2.31}$$

holds for all $\mu \in J_{M_i-1}, \nu \in J_{M_j-1}$.

Remark 2.24.
The first point of the above assumption is the one which can be found for example in the book of Kushner and Dupuis [KD01].

The following example shows that these two assumptions are not equivalent to each other.

Example 2.25.
In the first case point 1 of Assumption 2.23 does not hold but point 2. In the second case we have the reverse situation.

1. *Consider the intervals* $\mathcal{V}_1 = [0.2, 0.25]$, $\mathcal{V}_2 = [0.3, 0.33]$, $\mathcal{V}_3 = [0.25, 0.28]$, *and* $\mathcal{C}_{1,2} = [0.3, 0.4]$, $\mathcal{C}_{1,3} = [0.25, 0.3]$, $\mathcal{C}_{2,3} = [0.2, 0.25]$.
   For the tuples $(i, j) = (1, 2), (1, 3), (2, 3)$ *we have* $2\,\Xi^d_{i,j} \leq 0.5\,\Xi^u_{i,j}$ *and the selection* $h_1^{\cdot} = 0.7\,h_2^{\cdot} = 0.7\,h_3^{\cdot}$ *fulfils the second point of Assumption 2.23. But since* $\underline{\sigma_1}^2 < \underline{\sigma_1}\overline{\sigma_2\rho_{1,2}} + \underline{\sigma_1}\overline{\sigma_3\rho_{1,3}}$ *holds at least one matrix in* Γ_3' *is not diagonally dominant.*

2. *Let the volatilities* $\sigma_1 = 0.32$, $\sigma_2 = 0.25$, *and* $\sigma_3 = 0.27$ *be constant and the correlations be uncertain:* $\mathcal{C}_{1,2} = [0.2, 0.35]$, $\mathcal{C}_{1,3} = [0.4, 0.55]$, *and* $\mathcal{C}_{2,3} = [0.2, 0.25]$.
   It can easily be verified that the set Γ_3' *only consists of diagonally dominant matrices. But we have* $2\,\Xi^d_{1,3} \approx 1.3037 > 1.0774 \approx 0.5\,\Xi^u_{1,2}$.

The discretisation matrix $A^i_\gamma \in \mathbb{R}^{M^i \times M^i}$ now has $2n^2 - 2n + 3$ non-zero diagonals in the case of strictly negative or positive correlation, respectively. Otherwise, there are $2n^2 + 1$ diagonals. The boundary conditions at each boundary of Ω_3^D are implemented as in the two-dimensional case. The results of Reiß and Wystup [RW01] can be transferred and an algorithm for the evaluation of the three-dimensional cumulative normal distribution function is given in the book of Haug [Hau06].

The next step is to prove that we can guarantee the non-positivity of the associated off-diagonal entries by an adequate choice of forward, backward, and centered difference quotients.

Lemma 2.6.
Let a monotone grid $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_3) \subset \Omega_3^D$ *be given and let Assumption 2.23 be satisfied. Then, for every* $\gamma \in \Theta_n$ *and by choosing between forward, backward and centered difference quotients as described in Section 2.2.1 for each spatial direction, the discretisation matrix has only non-positive off-diagonal and non-negative diagonal entries, if* $r > 0$.

Proof.
We prove the above statement for the second case of Assumption 2.23. We start by

showing $d^{\bar{j}}_{\mu,\gamma} \leq 0, \mu \in I_d$*, for an inner grid point with lexicographic index* $\bar{j}$*. We have*

$$\begin{aligned} d^{\bar{j}}_{\mu,\gamma} &= -\frac{1}{2}\sigma^2_{\mu,\bar{j}} + \frac{1}{4}\sum_{\nu=1,\nu\neq\mu}^{n} \rho_{\mu,\nu,\bar{j}}\sigma_{\mu,\bar{j}}\sigma_{\nu,\bar{j}} \left(\frac{h^k_\mu}{h^{\iota_\nu}_\nu} + \frac{h^{k+1}_\mu}{h^{\iota_\nu+1}_\nu}\right) \\ &\leq -\frac{1}{2}\sigma^2_{\mu,\bar{j}} + \frac{1}{2(n-1)}\sum_{\nu=1,\nu\neq\mu}^{n} \rho_{\mu,\nu,\bar{j}}\sigma_{\mu,\bar{j}}\sigma_{\nu,\bar{j}}\Xi^u_{\mu,\nu} \\ &\leq -\frac{1}{2}\sigma^2_{\mu,\bar{j}} + \frac{1}{2}\underline{\sigma_\mu}\sigma_{\mu,\bar{j}} \leq 0 \end{aligned}$$

where $\bar{j}$ *is the lexicographic index to the indices* $k \in J_{M_\mu-1}$*, and* $\iota_\nu \in J_{M_\nu-1}$*.*
From the constructions of Section 2.2.1, especially Lemma 2.1, we can assure (including the maximal use of central difference quotients) that the off-diagonal entries corresponding to the μ*-th spatial direction of the grid point with lexicographic index* $\bar{j}$ *are non-positive.*
At the boundary points we use, as before, difference quotients with equidistant step sizes. Here the off-diagonal entries are non-positive by construction.
Furthermore, since the diagonal entry for each grid point is the negative sum of all off-diagonal entries plus the interest rate r*, it is non-negative.*

□

The results of the previous section for the choice of the spatial step sizes also apply in this more general setting since they were derived independent of any dimension. The set of admissible volatilities is now possibly divided into at most 3^n subsets $\Theta_{\dots}$. Each of them representing one combination of the chosen difference quotients for each of the different spatial directions.

2.2.4 Options on one asset

In this section we shortly transfer the results of the Sections 2.2.1 and 2.2.2 to the discretisation of the non-linear Black-Scholes equation (2.3). The pricing problem is given by Problem 2.2. A similar Finite Difference discretisation has been performed by Heider [Hei10] for European and Schaeling [Sch10] for the pricing of American options. A similar transformation as (2.1) was applied by them, but the value function was additionally scaled by the underlying's value. Their results are calculated on an equidistant grid in space and in time.
In the present thesis the discretisation is performed on a non-equidistant grid. The foregoing and the results are similar to the ones of the previous sections and the mentioned publications. Thus, we keep this section short and refer to the mentioned papers and the previous sections.

Let a grid $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_1)$ be given. We use the notations introduced in Notation 1.20 and 2.7. At the i-th time level the $M^i + 1$ grid points are denoted by x_1^k, $k \in J_{M^i}$ and the discrete value $v(\tau^i, x_1^k)$ are denoted by v_k^i. (We omit the time index in order to keep notation simple.) For the time being, let the volatility $\sigma_1^i \in \Omega_1$ be fix.
Without loss of generality, we now consider (2.3) at an inner grid point (τ_i, x_1^k). By rearranging terms we obtain

$$v_\tau + d_{1,\gamma}^k \delta_{x_1}^2 v_k + c_{1,\gamma}^k \delta_{x_1}^{\cdot} v_k + r v_k = 0, \tag{2.32}$$

where

$$d_{1,\gamma}^k = -\frac{1}{2}\sigma_{1,\bar{j}}^2 \quad \text{and}$$
$$c_{1,\gamma}^k = \frac{1}{2}\sigma_{1,\bar{j}}^2 - r + \delta_1.$$

For the spatial discretisation we use the standard Finite Difference stencil for one dimension. Our aim is again to choose central differencing as often as possible. For the inner grid points the choice of forward, backward, and central differencing is thus done as described in Lemma 2.1. We obtain:

$$v_\tau + a_k^{k+1} v_{k+1} + a_k^k v_k + a_k^{k-1} v_{k-1} = 0,$$

where

$$a_k^{k+1} = \frac{2d_{1,\gamma}^k}{(h_1^k + h_1^{k+1})h_1^{k+1}} + \begin{cases} \dfrac{c_{1,\gamma}^k}{h_1^{k+1}}, & (2.13a),\ \text{not } (2.13b), \\ 0, & (2.13b),\ \text{not } (2.13a), \\ \dfrac{c_{1,\gamma}^k h_1^k}{(h_1^k + h_1^{k+1})h_1^{k+1}}, & (2.13c), \end{cases}$$

$$a_k^{k-1} = \frac{2d_{1,\gamma}^k}{(h_1^k + h_1^{k+1})h_1^k} - \begin{cases} 0, & (2.13a),\ \text{not } (2.13b), \\ \dfrac{c_{1,\gamma}^k}{h_1^k}, & (2.13b),\ \text{not } (2.13a), \\ \dfrac{c_{1,\gamma}^k h_1^{k+1}}{(h_1^k + h_1^{k+1})h_1^k}, & (2.13c), \end{cases} \tag{2.33}$$

$$a_k^k = r - a_k^{k+1} - a_k^{k-1}$$

The equations (2.33) describe the three cases for the choice of forward, backward, and central differencing, respectively.
The discretisation is completed by prescribing the behaviour of v at the boundary of the grid. We will use a combination of Neumann and Dirichlet boundary conditions. An alternative choice for Vanilla options is to use their asymptotic behaviour for $x \to \pm\infty$ like it was done in [Hei10, Sch10]. For more advanced options like Barriers,

Dirichlet conditions are inevitable at the barrier. On the other boundaries we use Neumann conditions given by the options‘ Delta.
At the boundary points we use central difference quotients. The Neumann boundary conditions are incorporated by assuring

$$\delta^c_{x_1} v^i_k = \beta^i_k, \quad k = 0, M^i \tag{2.34}$$

for a given constant $\beta^k_i \in \mathbb{R}$, cp. Table 2.2. A calculation analogue to the two asset case, cp. Section 2.2.1, yields the modified coefficients for the grid‘s boundary points and those adjacent to them. Collectively, we obtain the discretisation matrix

$$A^i_\gamma = \begin{pmatrix} a^0_0 & a^1_0 & 0 & & & \dots & 0 \\ a^0_1 & a^1_1 & a^2_1 & 0 & & & \\ 0 & \ddots & \ddots & \ddots & & & \\ \vdots & & & & & & \vdots \\ & & & \ddots & \ddots & \ddots & 0 \\ & & & 0 & a^{M_1-2}_{M_1-1} & a^{M_1-1}_{M_1-1} & a^{M_1}_{M_1-1} \\ 0 & & \dots & & 0 & a^{M_1-1}_{M_1} & a^{M_1}_{M_1} \end{pmatrix} \tag{2.35}$$

and the vector $b^i = (b^i_j)_j \in \mathbb{R}^{M^i_1+1}$

$$\begin{gathered} b^i_1 = \left(2\frac{d^0_{1,\gamma}}{h^0_1} - c^0_{1,\gamma}\right)\beta^i_0, \qquad b^i_{M^i_1} = -\left(2\frac{d^{M^i_1}_{1,\gamma}}{h^{M^i_1}_1} + c^{M^i_1}_{1,\gamma}\right)\beta^i_{M^i_1}, \text{ and} \\ b^i_j = 0, j \in I_{M^i_1-1}. \end{gathered} \tag{2.36}$$

Dirichlet boundary conditions are incorporated by setting

$$v^i_k = \beta^i_k, \quad k = 0, M^i_1. \tag{2.37}$$

where β^i_k is the given value at $(\tau_i x^k_1)$. To include them into the system of equations we modify the first or the last row of A^i_γ, respectively.

Remark 2.26.
Here we handle the controls implicitly contained in the boundary conditions as described in Remark 2.16.

The following table gives an overview of the two different boundary conditions for different types of options.

By construction we have a one-dimensional analogon to Lemma 2.2.

Tab. 2.2: *Overview of the boundary conditions for different options on one asset.*

	β_0^i		$\beta_{M^i}^k$	
Option	Dirichlet	Neumann	Dirichlet	Neumann
Call	0	-	-	$\mathcal{N}_1(d_1)S_{M_1}$
Put	-	$-\mathcal{N}_1(-d_1)S_0$	0	-
Butterfly sp.	0	-	0	-
DaO put	0	-	0	-
DaO call	0	-	-	$\mathcal{N}_1(d_1)S_{M_1}$
UaO call	0	-	0	-

Corollary 2.7.
Let $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_1)$ be a grid. Let the space discretisation be given by (2.33) and (2.36) *and the corresponding boundary conditions by* (2.34) *or* (2.37)*.*
Then, for $r > 0$, the matrix A_γ^i in (2.35) has non-positive off-diagonal entries and non-negative diagonal entries for all $i \in I_L$ and $\gamma^i \in \Theta_1$.

Proof.
The coefficient $d_{1,\gamma}^k$ is always negative for $\underline{\sigma_1} \neq 0$. Thus, the result holds for the inner grid points by the construction done in Section 2.2.1. For the boundary points the result holds by definition for the Dirichlet conditions and by construction for the Neumann conditions.

□

From Section 2.2.1 we know that if the correlation is zero it is always possible to guarantee the use of central differencing by choosing the steps in each space-dimensions small enough, cp. Lemma 2.4. In the one-dimensional case the terms brought into by correlation do not occur and we thus can apply this lemma to compute an upper bound for the step width h_1.

Corollary 2.8.
Let a grid be given and let the discretisation be as in the previous Corollary. Then, only centered difference quotients can be used for all $\gamma \in \Theta_1$ at every grid point, if

$$h_1^k \leq \begin{cases} \min\limits_{x \in \{\underline{\sigma_1}, \overline{\sigma_1}\}} \dfrac{x^2}{\frac{1}{2}x^2 - r + \delta_1}, & r > \delta_\mu \\ \min\limits_{x \in \{\underline{\sigma_1}, \overline{\sigma_1}\}} \dfrac{x^2}{\max\left\{\left|\frac{1}{2}x^2 - r + \delta_1\right|, \frac{{\sigma_1}^2}{h_1^k}\right\}}, & 0 \leq r \leq \delta_\mu \end{cases} \tag{2.38}$$

holds.

The above result is equivalent to the one given in [Sch10, Hei10]. The slight differ-

ences are due to the distinct transformation we applied.

Completing the spatial discretisation, we have a system of (non-linear) coupled ordinary differential equations

$$v_\tau + A_\gamma^{\cdot} v = b^{\cdot}.$$

for every time step. This system is of the same structure as are those for the multi-dimensional cases before. Therefore, the discretization is performed as in these cases, cp. Section 2.3.

2.3 Discretisation of time and general properties

In this section we complete the discretisation and analyse their properties. These will be of great importance for the final algorithm.

Discretisation of time. So far we have performed a semi-discretisation of the BSB equation (2.8)

$$\frac{\partial}{\partial \tau} v(\tau) + A_\gamma v(\tau) = b, \quad \gamma \in \Theta_d \text{ (fixed)} \tag{2.39}$$

for options on one, two, and more than two assets, respectively. Now we do the last step and execute the discretisation for the time variable. For the moment we will focus on an equidistant grid in time. The treatment of the more general case can be found in the following chapters.
We will consider the following three time-stepping methods: 1-step backward differentiation formula (BDF1), 2-step backward differentiation formula (BDF2), and Crank-Nicolson (CN). They belong to the standard instruments for time integration in financial methods. For further details we refer to the book of Seydel [Sey12].

Let a grid $\mathcal{G}(L, \{t^i\}, \mathcal{S}, \Omega_2)$ with equidistant sampling points $\tau^i, i \in J_L$, be given. We consider a general time step from level τ^{i-1} to $\tau^i, i \in I_{L-1}$. The initial condition for all options is given by their specific payoff

$$v_{\bar{j}}^0 = \Psi(x_{\bar{j}}).$$

The BDF1 method results if we discretise the time derivative by the backward difference quotient. If we apply both forward and backward difference quotient and combine them we obtain the CN method. For ease of notation we use the common

θ-notation

$$(I + \theta h_\tau A_\gamma^i)v^i = (I - (1-\theta)h_\tau A_\gamma^{i-1})v^{i-1} + h_\tau \left(\theta b^i + (1-\theta)b^{i-1}\right). \tag{2.40a}$$

For $\theta = 1$ we obtain BDF1 time-stepping with discretisation error in time of $\mathcal{O}(h_\tau)$. For $\theta = \frac{1}{2}$ we get CN time-stepping with an (theoretical) error of $\mathcal{O}(h_\tau^2)$ in time, cf. [Sey12]. The EE scheme would result for $\theta = 0$. But it is of no further interest, cp. Remark 4.2.
For ease of notation we write

$$B_{\gamma,\theta}^i = (I + \theta h_\tau A_\gamma^i), \tag{2.40b}$$
$$C_{\gamma,\theta}^{i-1} = (I - (1-\theta)h_\tau A_\gamma^{i-1}), \quad \text{and} \tag{2.40c}$$
$$r_\theta^i = h_\tau \left(\theta b^i + (1-\theta)b^{i-1}\right). \tag{2.40d}$$

The BDF2 scheme uses the difference quotient $\frac{1}{h_\tau}(1.5v^i - 2v^{i-1} + 0.5v^{i-2})$ to approximate the first order derivative. Thus, it (theoretically) is of second order. The complete discretisation then reads

$$\left(\frac{3}{2} + h_\tau A_\gamma^i\right) v^i = -2v^{i-1} + \frac{1}{2}v^{i-2} + r^i \tag{2.41}$$

for constant $h\tau$. The schee can also be generalised to non-uniform grids in τ.
Analogously to (2.40) we redefine $B_{\gamma,\theta}^i$, $C_{\gamma,\theta}^{i-1}$, and r_γ^i in an appropriate way and set $D := 0.5I$. The three schemes are then combined into one notation by

$$B_{\gamma,\theta}^i v^i = C_{\gamma,\theta}^{i-1} v^{i-1} + Dv^{i-2} + r_\theta^i. \tag{2.42}$$

Remark 2.27.
Obviously, the BDF2 scheme is not monotone. Due to the term $-2v^{i-1}$ on the right hand side of equation (2.41) *condition 1. of Definition 1.23 can in no case be fulfilled.*

From a theoretical point of view the CN and the BDF2 scheme only are of second order if the value function is continuous differentiable in time. This condition is not satisfied in general for option pricing problems. From the numerical point of view this problem can be improved if we use Rannacher timestepping [Ran84] for the first few time steps. That is, the first few time steps are done with the BDF1 method. Thereafter, we use the prescribed scheme. Some numerical evidence on this topic can be found in [GC06, PVF03]. The Rannacher approach also provides the first two points for the initialisation of the BDF2 method.

With the results from the previous sections we could expect a scheme of second order for CN and BDF2 schemes if only centered difference quotients are used. Since it is not guaranteed for all choices of difference quotients to be applicable permanently, we generally could not expect the space discretisation to be of second order. But by using them as much as possible we hope to not only obtain first order errors (for a continuous option value).
For the BDF1 scheme we could only expect first order convergence for the whole scheme.
For all schemes the influence of changing controls may have an impact on their accuracy.

General properties. Having completed the discretisation of the BSB equation in a general manner, we now analyse the properties of the discrete equation.
Let us introduce the concept of M-matrices. We only give a short introduction at this point. For more information see for example the book of Horn and Johnson [HJ94] or the one of Berman and Plemmons [BP94].

Definition 2.28 (M-matrix)**.**
A matrix $A = (a_{i,j})_{i,j} \in \mathbb{R}^{n\times n}, i, j \in I_n$, is called M-matrix if all off-diagonal entries are non-positive: $a_{i,j} \leq 0, i, j \in I_n, i \neq j$, and the real part of all Eigenvalues of A is positive.

M-matrices have some very helpful characterisations which we will use in the following sections. Here we summarise them as

Lemma 2.9.
Let $A = (a_{i,j})_{i,j} \in \mathbb{R}^{n\times n}$ be a matrix with $a_{i,j} \leq 0, i, j \in I_n, i \neq j$. The following statements:

1. *A is an M-matrix.*
2. *The matrix A is regular and inverse positive, that is $A^{-1} \geq 0$.*
3. *The matrix A is monotone, i.e. for $x \in \mathbb{R}^n$ it follows from $Ax \geq 0$ that $x \geq 0$.*
4. *The diagonal entries of A are non-negative and there is a vector $x \in \mathbb{R}^n$ so that $Ax \geq 0$.*
5. *There exists a positive diagonal matrix $D \in \mathbb{R}_+^{n\times n}$ so that the matrix AD is*

strictly diagonal dominant, i. e.

$$d_{i,i}|a_{i,i}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{i,j}|d_{j,j}, \quad i \in I_n.$$

are equivalent [HJ94, BP94].

The proofs of the single equivalences can be found in [HJ94] or [BP94], respectively.

Lemma 2.9 enables us to prove one important property of the discretisation matrix $B^i_{\gamma,\theta}$ in (2.42).

Lemma 2.10.
Let a grid $\mathcal{G}(L, \{t^i\}, \mathcal{S}, \Omega_2)$ satisfying Assumption 2.9 or Assumption 2.23, respectively, be given. Let the spatial discretisation be given by 2.18 and (2.19)*, or* (2.33) *and* (2.36) *or the corresponding generalisations for $n \geq 3$. Furthermore, let $r \geq 0$. Then, for every time step $t^{i-1} \to t^i$ and every control $\gamma \in \Theta_n$ the discretisation matrix $B^i_{\gamma,\theta}$ in* (2.42) *is an M-matrix*

Proof.
From Lemma 2.2, Lemma 2.6, Corollary 2.7, applied for the respective n, we know that A^i_γ has non-positive off-diagonal entries and non-negative diagonal entries. Obviously, the same holds true for $B^i_{\gamma,\theta}$.
By definition of the diagonal entries of A^i_γ in (2.18) the matrix is strictly diagonally dominant for $r > 0$ by the triangle inequality. This property can be maintained for $r \geq 0$ for the matrix $B^i_{\gamma,\theta}$ since we add a positive constant to each scaled diagonal entry.
By Lemma 2.9 with $D = I$ the proof is completed.

□

The result of Lemma 2.10 is very useful for pricing American options, s. the next section, and the determination of the optimal control $\gamma \in \Theta_n$, s. Section 3.1.
From the above proof we know that we always get a discretisation matrix $B^i_{\gamma,\theta}$ with the M-matrix property in (2.42) if a discretisation of the Black-Scholes-Barenblatt equation as described in Section 2.2 can be constructed. Thus the main difficulty is to construct such a discretisation. The monotonicity of the scheme in the sense of Definition 1.23 does not result from this construction generally, s. Chapter 4.

2.4 American options

In the previous section we derived the discretisation for the European option pricing for an option on n assets. In this section we explain how the value of an American option can be calculated numerically. From arbitrage reasons we know that the option value could not fall below its payoff. Either it is equal to the payoff and does not satisfy the pricing equation or it is larger than the payoff and the pricing equation is satisfied. This fact is reflected in the complementary problem of Problem 2.4 on page 26.
We will apply the well known penalty iteration for this purpose. The intrinsic idea is to solve the pricing equation for a European option, but whenever the option value falls below the payoff it is penalised. Thereby, the option value is increased to the payoff at the next step of the iteration. If the calculated value is larger than the payoff, no penalisation is necessary. Numerically, this aim is reached by adding an artificial term, the penalty term, to the pricing equation. This iteration repeats as long as the changes in the option value are above a predefined tolerance level.
In the literature the term

$$\mathcal{P}v := \frac{1}{\varepsilon}(\Psi - v)^+ \tag{2.43}$$

where ε is a small positive figure, is widely used, cf. [FV02, ZWYT09, KLM07]. In the notation used above and the formulation of Problem 2.4 the problem of valuing an American option can be reformulated as

Problem 2.29 (Valuing American options via the Penalty method)**.**
Solve the equation

$$\frac{\partial}{\partial \tau} - \mathcal{L}^n_{BSB} v - \mathcal{P}v = 0 \tag{2.44}$$

on Ω_n*.*

We remind the reader that we previously assumed that the price of an American option in the non-linear model is given by the assigned non-linear complementary problem, cp. Remark 1.9.

The open question is: How accurate is the discrete solution of the penalised pricing equation (2.44)? The answer is the following: the quality of the solution depends on the choice of ε. For the standard Black-Scholes equation it can be shown that the smaller we choose ε the more accurate the complementary problem is solved by the penalised solution, cf. [FV02]. For the one-dimensional Uncertain Volatility model

this property has also been proven by Schaeling and Heider [SH11].

If we solve the penalised pricing equation (2.44) numerically, we handle the penalty term implicitly for all three methods. That is we want to solve the equation

$$B^i_{\gamma,\theta} v^i = C^i_{\gamma,\theta} v^{i-1} + D v^{i-2} + r^i_\theta + h_\tau P_\varepsilon(v^i) \left(p - v^i\right)^+ \tag{2.45}$$

where $P_\varepsilon(v^i) \in \mathbb{R}^{M^i \times M^i}$ is a matrix defined by

$$\left(P_\varepsilon(v^i)\right)_{j,k} := \begin{cases} \frac{1}{\varepsilon}, & v_k < p_k, j = k \\ 0, & \text{else} \end{cases}$$

and p is a vector which contains the payoff at the grid points in lexicographical ordering. The factor h_τ is due to the time discretisation.
Collectively, we obtain a non-linear system of equations to approximately solve the discrete, non-linear complementary problem. We solve the non-linear scheme (2.45) by a generalised Newton iteration.

Notation 2.30.
In the j-th iteration step we denote by $v^{i,(j)}$ the current value at time level t_i.

If we define the generalised Jacobi-matrix $\mathrm{D}P_\varepsilon(v) \in \mathbb{R}^{M^i \times M^i}$ via

$$\left(\mathrm{D}P_\varepsilon(v)\right)_{j,k} = \left(\frac{\partial}{\partial v_k} P_\varepsilon \left(p_j - v_j\right)^+\right)_{j,k} := \left(P_\varepsilon(v)\right)_{j,k}$$

the Newton iteration can be formulated as

$$\begin{aligned} \left(B^i_{\gamma,\theta} + P_\varepsilon(v^{i,(j)})\right) \Delta v^{(j)} &= -B^i_{\gamma,\theta} v^{i,(j)} + C^i_{\gamma,\theta} v^{i-1} + D v^{i-2} \\ &\quad + r^i_\theta + h_\tau P_\varepsilon(v^{i,(j)}) \left(p - v^{i,(j)}\right)^+ \\ v^{i,(j+1)} &= v^{i,(j)} + \Delta v^{(j)} \end{aligned} \tag{2.46}$$

The generalised Newton method is summarised in

Algorithm 2.1: Newton iteration for the Penalty method.

Input: $\theta, h_\tau, B^i_{\gamma,\theta}, C^{i-1}_{\gamma,\theta}, D, r^i_\theta, v^{i-1}, v^{i-2}, p, \varepsilon$
Output: v^i

1 **penalty_iteration**$(\theta, h_\tau, B^i_{\gamma,\theta}, C^{i-1}_{\gamma,\theta}, D, r^i_{\gamma,\theta}, v^{i-1}, v^{i-2}, p, \varepsilon)$
2 $\quad j \leftarrow 1$;
3 $\quad v^{i,(0)} \leftarrow v^{i-1}$;
4 $\quad err \leftarrow 1$;
5 $\quad$ **while** $err \geq 10^{-6}$ **do**
6 $\quad\quad$ *calculate* $P_\varepsilon \leftarrow P_\varepsilon\left(v^{i,(j-1)}\right)$;
7 $\quad\quad$ *solve* $\left(B^i_{\gamma,\theta} - h_\tau P_\varepsilon\right) v^{i,(j)} = C^{i-1}_{\gamma,\theta} v^{i-1} + D v^{i-2} + r^i_\theta + h_\tau P_\varepsilon p$;
8 $\quad\quad err \leftarrow \|v^{i,(j-1)} - v^{i,(j)}\|_\infty$;
9 $\quad\quad j \leftarrow j + 1$;
10 $\quad$ **return** $v^{i,(j-1)}$;

As initial value for the Newton iteration we use the numerical approximation to the option value of the previous time step v^{i-1}. As we will see in the following theorem, the convergence can be guaranteed for any start value.

Remark 2.31.
The system of equations in line 6 of Algorithm 2.1 is not the original Newton iteration (2.46)*, but an equivalent one. In this special form the numerical effort is reduced. We save one matrix-vector-multiplication and two vector-additions per iteration.*

The convergence of Algorithm 2.1 can be guaranteed if the iteration matrix is an M-matrix, compare section 2.3. The following theorem states the convergence. A proof can be found in [FV02].

Theorem 2.11 (Convergence of Algorithm 2.1.)**.**
Let the matrix $B^i_{\gamma,\theta} \in \mathbb{R}^{M^i \times M^i}$ in (2.45) *be an M-matrix. Then the iteration of Algorithm 2.1 converges monotonically in a finite number of iterations to a unique solution of the discrete non-linear pricing equation* (2.45)*.*

The idea of the proof is to use the property of M-matrices to preserve positivity, cp. point 3. of Lemma 2.9.

Additionally to the convergence of the penalty iteration, the error of the discrete complementary problem has to be investigated. In the publication [FV02] Forsyth and Vetzal analysed this question for a Finite Volume method for the standard Black-Scholes model. The result, though, is more general and also applies in this case.

Lemma 2.12.
Let the matrix $B^i_{\gamma,\theta} \in \mathbb{R}^{M^i \times M^i}$ in (2.45) be an M-matrix and $C^i_{\gamma,\theta} \in \mathbb{R}^{M^i \times M^i}_+$ be positive.
Then for every time level i the solution v^i of Algorithm 2.1 solves

$$\begin{aligned} B^i_{\gamma,\theta} v^i - C^i_{\gamma,\theta} v^{i-1} - r^i_\theta &\geq 0, \\ v^i - p &\leq -\varepsilon C, \\ B^i_{\gamma,\theta} v^i - C^i_{\gamma,\theta} v^{i-1} - r^i_{\gamma,\theta} = 0 \quad \vee \quad |v^i - p| &\leq \varepsilon C, \end{aligned}$$

where $C \in \mathbb{R}_+$ is a constant independent of $h_\tau, h_1,$ and h_2.

Accordingly to the above result, the smaller we choose $\varepsilon > 0$ the more accurate the discrete complementary problem will be solved.
The conditions for the discretisation matrices $B^i_{\gamma,\theta}$ being an M-Matrix will be investigated thoroughly in Chapter 4.

3 Optimal controls

In Chapter 2 we have constructed different types of Finite Difference discretisation for the n-dimensional BSB equation. Provided that the set of volatility parameters fulfils some assumptions their properties we need in this chapter are guaranteed for all controls $\gamma \in \Theta_n$.

In the previous chapters we assumed that they optimal controls respectively volatilities and correlations for each spatial grid point were known. In this context optimal means that volatilities and correlations a locally determined to optimise an objective function including the approximated derivatives of the option value. In this chapter a solution for several distinctive tasks is presented. By the given structure of the pricing problem it is obvious that at a time level i the option value and the optimal control for each grid point strongly depend on each other. Thus, we need two different things: first, a method to find the optimal relation between option value and control and second a procedure to solve the optimisation problem at a grid point. The first one has global character since it works on the whole spatial grid while the latter one only works local. To solve the first task we introduce an iteration developed by Forsyth and Labahn [FL07]. With the discretisations constructed in Chapter 2 its convergence can be easily guaranteed.

For the local optimisation problem we resort to standard optimisation results as much as possible. We analyse the n-dimensional optimisation problem for $n = 1, 2, 3$. We have to take into account that also the control and the choice of the difference quotient depend on each other, cp. Section 2.2.1. If only central difference quotients are used we use the Karush-Kuhn-Tucker (KKT) conditions to find optimal points. Otherwise the optimisation becomes much more costly and we have to relax it.

This chapter is divided into two sections. The first one introduces the so called policy iteration for the global aspect of the determination of the optimal controls. The complete proof and a detailed algorithm for the method are given, cp. Section 3.1. The second section focusses on the local optimisation at a single grid point, s. Section 3.2.

3.1 Global determination of the optimal controls

The aim of this section is to determine the option value v^i and the control $\gamma^i \in \Theta_n^{M^i}$ in such a way that they solve the discrete version of the transformed BSB equation (2.2) and the discrete optimisation problem

$$\max_{\substack{\sigma_i \in \mathcal{V}_i \\ \rho_{k,l} \in \mathcal{C}_{k,l}}} \left\{ \sum_{i=1}^{n} \sigma_i^2 \left(\delta_{x_i}^2 v_{\vec{j}}^i - \delta_{x_i}^{\cdot} v_{\vec{j}}^i \right) + \sum_{\substack{k,l=1 \\ k \neq l}}^{n} 2\sigma_k \sigma_l \rho_{k,l} \delta_{x_k x_l}^{\cdot} v_{\vec{j}}^i \right\} \tag{3.1}$$

at the same time. We remind the reader that the solution of 3.1 is included in the discrete schemes (2.42). The optimisation problem is independent of the fact whether we price American or European options.

We use the following notation to properly describe the policy iteration.

Notation 3.1.
In the j-th iteration step we denote by $v^{i,(j)}$ the current value at time level t_i. Similarly, we write $A_\gamma^{i,(j)}$, $b_\gamma^{i,(j)}$, and $\gamma^{i,(j)}$...

Let us for now assume that we know an optimal control $\gamma^{i,(j)}$ for each $v^{i,(j)}$, s. Section 3.2.
The idea of the iteration by Forsyth and Labahn [FL07] is to calculate the optimal control, the so called policy, over the last iterate v^{i-1}. With adequately updated iteration matrix the next iterate is then calculated by solving (2.42). We remind the reader that the controls contained implicitly in the Neumann boundary conditions are not updated, cp. Remark 2.16. The stopping criterion is fulfilled if the difference between the last two iterates in the Euclidean norm is smaller than a predefined tolerance level.
As initial values the option value v^{i-1} and the control γ^{i-1} of the previous time level are chosen. For the very first iteration we use constant controls.

In Algorithm 3.1 a pseudocode of the described iteration is given.

The function *det_control*$(\cdot)$ called in line 8 of Algorithm 3.1 will be explained in the next section, s. Algorithms 3.2 and 3.3.
The policy iteration has successfully been applied to different stochastic control problems like asset allocation, cf. [WF10], or unequal borrowing and lending interest rates [FL07].

Algorithm 3.1: Policy iteration for implicit option value.

Input: $\theta, h_\tau, v^{i-1}, v^{i-2}, \gamma^{i-1}, B^{i-1}_{\gamma,\theta}, C^{i-1}_{\gamma,\theta}, D, r^i_\theta$

Output: $v^{i,(j)}, B^{i-1,(j)}_{\gamma,\theta}, r^{i,(j)}_\theta$

pol_iteration$\left(\theta, h_\tau, v^{i-1}, v^{i-2}, \gamma^{i-1}, B^{i-1}_{\gamma,\theta}, C^{i-1}_{\gamma,\theta}, D, r^{i-1}_\theta\right)$
- $j \leftarrow 0, \quad err \leftarrow 1;$
- $v^{i,(0)} \leftarrow v^{i-1}, \quad \gamma^{i,(0)} \leftarrow \gamma^{i-1};$
- $B^{i,(0)}_{\gamma,\theta} \leftarrow B^{i-1}_{\gamma,\theta}, \quad r^{i,(0)}_\theta \leftarrow r^{i-1}_\theta$;
- **while** $err \geq 10^{-6}$ **do**
  - **for** $(x_{1,k}, x_{2,l}), k \in \mathcal{J}_{M_1}, l \in \mathcal{J}_{M_2}$ **do**
    - *calculate* a, b, c *as derivatives of* $v^{i,(j)}$; //s. (3.6)
    - $\gamma^{i,(j+1)}_{k,l} = det\_control(\cdot);$
  - *update* $B^{i,(j+1)}_{\gamma,\theta} \curvearrowleft B^{i,(j)}_{\gamma,\theta};$
  - *solve* $\quad B^{i,(j+1)}_{\gamma,\theta} v^{i,(j+1)} = C^{i-1}_{\gamma,\theta} v^{i-1} + D v^{i-2} + r^{i,(j+1)}_\theta;$
  - $err \leftarrow \|v^{i,(j+1)} - v^{i,(j)}\|_2;$
  - $j \leftarrow j + 1;$
- **return** $v^{i,(j)}, B^{i,(j)}_{\gamma,\theta}, r^{i,(j)}_\theta;$

To prove the convergence of Algorithm 3.1 we need some properties of the determined control at every step of the algorithm. They are summarised in

Remark 3.2.
The control $\gamma^{i,(j)} \in \Theta_n^{M^i}$ *calculated in the* j*-th step of Algorithm 3.1 satisfies*

$$\gamma^{i,(j)} \in \underset{\gamma \in \Theta_n^{M^i}}{\arg\max} \left\{ -B^{i,(j-1)}_{\gamma,\theta} v^i + r^{i,(j-1)}_\theta \right\}. \tag{3.2}$$

The maximisation in (3.2) *is understood component wise, which means for each grid / row* $\bar{j}$ *point separately. Relation* (3.2) *holds, since the* $\bar{j}$*-th row of the system of equations equals the objective function* (3.1) *at the corresponding grid point with some additional constants independent of the control. Remark the sign of* $B^i_{\gamma,\theta}$ *by comparing (2.4) and (3.6).*
We write "$\in$*" instead of "*$=$*" since a unique optimal control does not necessarily exist.*

The returned value $v^{i,(j)}$ of Algorithm 3.1 is accepted as the option value v^i. For the practical implementation an additional variable $iter_{max}$ is included in order to limit the iteration steps.
Huang, Forsyth, and Labahn prove that the iteration converges from every starting point $v^{i-1}$and that the limit is unique. The theorem is as follows.

Theorem 3.1 (cf. [HFL12, WF10])**.**
Let an initial value v^{i-1} be given. Furthermore let the matrices $B^i_{\gamma,\theta}$, $C^{i-1}_{\gamma,\theta}$, and D of equation (2.42) as well as v^{i-1} and v^{i-2} be bounded in the maximum norm.
If $B^i_{\gamma,\theta}$ is an M-matrix for all $\gamma \in \Theta_n$ then Algorithm 3.1 converges monotonously to a unique value v^*.

Proof (cf. [HFL12, WF10][1])**.**
First we show that the iterates $v^{i,(j)}$ are bounded independently of j. From Algorithm 3.1, line 10 we have

$$\begin{aligned}\left\|v^{i,(j+1)}\right\|_\infty &\leq \left\|\left(B^{i,(j+1)}_{\gamma,\theta}\right)^{-1}\right\|_\infty \left(\left\|C^{i-1}_{\gamma,\theta}v^{i-1}\right\|_\infty + \left\|Dv^{i-2}\right\|_\infty + \left\|r^{i,(j+1)}_\theta\right\|_\infty\right)\\ &\leq C_1\|v^{i-1}\|_\infty + C_2\|v^{i-2}\|_\infty + C_3,\end{aligned}$$

where C_1, C_2, C_3 are a positive constants independent of $v^{i,(j)}$. The constant may be chosen as the maximum of the matrix norm over all feasible controls in Θ_2. Thus, sincce the approximations of the option value v^{i-1} and v^{i-2} for previous time levels are bounded this also holds true for the sequence of the algorithm
The next step is to prove that the $v^{i,(j)}$ form a non-decreasing sequence. From Algorithm 3.1, line 10 we have

$$\begin{aligned}B^{i,(j+1)}_{\gamma,\theta}(v^{i,(j+1)} - v^{i,(j)}) &= C^{i-1}_{\gamma,\theta}v^{i-1} + Dv^{i-2} + r^{i,(j+1)}_\theta - B^{i,(j+1)}_{\gamma,\theta}v^{i,(j)}\\ &\quad - \underbrace{\left(-B^{i,(j)}_{\gamma,\theta}v^{i,(j)} + C^{i-1}_{\gamma,\theta}v^{i-1} + Dv^{i-2} + r^{i,(j)}_\theta\right)}_{=0}\\ &= -B^{i,(j+1)}_{\gamma,\theta}v^{i,(j)} + r^{i,(j+1)}_\theta - \left(-B^{i,(j)}_{\gamma,\theta}v^{i,(j)} + r^{i,(j)}_\theta\right)\\ &\geq 0.\end{aligned} \tag{3.3}$$

The last appraisal follows from Remark 3.2.
With Lemma 2.9 we have $v^{i,(j+1)} \geq v^{i,(j)}$ for $B^{i,(j)}_{\gamma,\theta}$ is an M-matrix.
Since the sequence $v^{i,(j)}$ is non-decreasing and bounded Algorithm 3.1 converges.
Finally, the uniqueness of the limit v^* is shown. Assume that there are two solutions v^* and w^* which satisfy

$$\begin{aligned}B^{i,(j)}_{\gamma_1,\theta}v^* &= C^{i-1}_{\gamma,\theta}v^{i-1} + Dv^{i-2} + r^{i,(j)}_\theta, \quad \gamma_1 \in \arg\max\left\{-B^i_{\gamma,\theta}v^* + r^i_\theta\right\}\\ B^{i,(j)}_{\gamma_2,\theta}w^* &= C^{i-1}_{\gamma,\theta}v^{i-1} + Dv^{i-2} + r^{i,(j)}_\theta, \quad \gamma_2 \in \arg\max\left\{-B^i_{\gamma,\theta}w^* + r^i_\theta\right\}\end{aligned}$$

[1]The proof given in [HFL12] gives a result for a more general case than the one considered here.

for $j \in \mathbb{N}$. By rearranging the two equations, we obtain

$$\begin{aligned} B^{i,(j)}_{\gamma_1,\theta}(v^* - w^*) &= -B^{i,(j)}_{\gamma_1,\theta} w^* + C^{i-1}_{\gamma,\theta} v^{i-1} + D v^{i-2} + r^{i,(j)}_{\theta} \\ &\quad - \left(-B^{i,(j)}_{\gamma_2,\theta} w^* + C^{i-1}_{\gamma,\theta} v^{i-1} + D v^{i-2} + r^{i,(j)}_{\theta}\right) \\ &= -B^{i,(j)}_{\gamma_1,\theta} w^* + r^{i,(j)}_{\theta} - \left(-B^{i,(j)}_{\gamma_2,\theta} w^* + r^{i,(j)}_{\theta}\right) \leq 0 \end{aligned}$$

Since γ_2 maximises $-B^i_{\gamma,\theta} w^* + r^i_\theta$ and $B^{i,(j)}_{\gamma,\theta}$ is an M-matrix we have $v^* \leq w^*$. If we interchange v^* and w^* we obtain the equality of the both solutions.
Collectively, Algorithm 3.1 converges monotonously to a unique solution v^*.

□

With Theorem 3.1 we can guarantee the convergence of Algorithm 3.1 for the BDF1, the BDF2, and the Crank-Nicolson scheme for every time level. For all three schemes the matrix $B^i_{\gamma,\theta}$ is an M-matrix by Lemma 2.10 for all $\gamma \in \Theta_n$ and $i \in I_L$.

Remark 3.3.
Another method that solves the same problem as the iteration by Huang, Forsyth, and Labahn has been developed by Witter and Reisinger [WR11]. The idea here is to penalise all choices of controls that are non-optimal. For their method they consider a discretised domain for the control. In each step of the iteration a matrix-vector multiplication has to be evaluated for every discrete control. Compared to the iteration of Huang, Forsyth, and Labahn the effort doing this would be to high for our three-dimensional control space combined with the size of the iteration matrix.

3.2 Local determination of the optimal controls

A key feature of the controls in Algorithm 3.1 was their maximisation property of the constraint quadratic equation given for any grid point by the discrete BSB equation, s. Remark 3.2.
In this section we describe how the associated optimisation problem can be solved. Speaking in the terms of Algorithm 3.1 the construction of the function *det_ control(·)* is now described.
For only one unknown the task is quite simple. If the BSB equation depends on two or more spatial variables it becomes more costly. Moreover, due to the possible non-convexity of the problem and the constraints which arise from the choice of difference quotients for the first order derivative the problem is more elaborate. Thus, we have to investigate carefully, whether the conditions derived in Section 2.2.2 for the exclusive use of central difference quotients are satisfied.

We start by giving a general suitable representation of the optimisation problem. Then, in the following two subsections we separately focus on those problems where only central difference quotients are applied and those where this is not the case. In the rest of this section we always consider w. l. o. g. the grid point $\overline{x}_{\overline{j}}$. For ease of notation we omit the index $\overline{j}$ where possible.

One asset only. For the one-dimensional Problem 2.2 the optimisation is quite simple. The use of centered difference quotients only can be assured by choosing the step widths small enough, cp. Corollary 2.8. The optimal control $\sigma^*_{1,k} \in \Theta_1$ at the grid point (τ^i, x_k) can be chosen according to the sign of $\delta_x^2 v_k^i - \delta_{\dot{x}} v_k^i$:

$$\sigma^*_{1,k} = \begin{cases} \overline{\sigma_1}, & \operatorname{sign}\left(\delta_x^2 v_k^i - \delta_{\dot{x}} v_k^i\right) \geq 0, \\ \underline{\sigma_1}, & \text{else} \end{cases} \tag{3.4}$$

If the step sizes are not small enough, monotonicity is maintained as described in Section 2.2.4. The interval is then divided into (at most) three subintervals. Their limits are implicitly given by equations (2.13) with the constants in (2.32). For each of the subintervals the maximising volatility is taken as above and then the one of these with maximal objective function is taken as $\sigma^*_{1,k}$. Compare also the following section for an analogue approach.

Two and more assets. The optimisation problem arising in the pricing equation for options on at least two assets is in general more difficult. The two-dimensional problem was formulated in Problem 2.3.
In the discrete setting considered here, we think of the optimal control at time level τ^i as a matrix $\gamma^i \in \Theta_n^{M^i} \subset \mathbb{R}^{M^i \times n}, n \geq 2$, cp. Notation 2.7.
It is important to notice that v^i and γ^i do depend on each other. In the previous section we introduced an iteration that converges to a (unique) v^i for a sequence of controls $\gamma^{i,(l)}, l = 1, \ldots$. A necessary key feature for the controls in every step of the iteration was their property of maximising the discrete equation at every grid point.
Now, we explain how to get the optimal control $\gamma^{i,(j)} = \gamma^{i,*} \in \Theta_n^{M^i}$ at each grid point $\overline{x}_{\overline{j}}, \overline{j} \in I_{M^i}$.
We rewrite the optimisation problem 3.1 at this grid point in the form

$$\max_{\Theta_n} \left(\sum_{\mu=1}^{d} a_\mu \sigma_\mu^2 + \sum_{\substack{\mu,\nu=1 \\ \mu \neq \nu}}^{d} b'_{\mu,\nu} \sigma_\mu \sigma_\nu \rho_{\mu,\nu} \right) \tag{3.5}$$

Here a_μ and $b'_{\mu,\nu}$ represent the discrete approximation of the factors used in the transformed equation (2.4)

$$a_\mu := \delta^2_{x_\mu} v^i_{\bar{j}} - \delta^{\cdot}_{x_\mu} v^i_{\bar{j}}, \quad b'_{\mu,\nu} := \delta^+_{x_\mu,x_\nu} v^i_{\bar{j}} \tag{3.6}$$

For the first order derivative the approximating difference quotient is not yet specified. Their choice implicitly also determines the approximation of the mixed derivative. We use the same difference quotients for the maximisation problem (3.5) that have been used to construct the monotone discretisation. This is absolutely necessary, to guarantee the convergence of Algorithm 3.1, cp. Remark 3.2 and the proof of Theorem 3.1.

The maximisation for the correlation can be done separately. If $b'_{\mu,\nu}$ is negative we choose $\rho^*_{\mu,\nu} = \underline{\rho_{\mu,\nu}}$ and $\rho^*_{\mu,\nu} = \overline{\rho_{\mu,\nu}}$ otherwise. With $b_{\mu,\nu} := b'_{\mu,\nu}\rho^*_{\mu,\nu}$ we obtain an equivalent optimisation problem to (3.5).

Problem 3.4.
Solve:

$$\begin{aligned} -\min_{\mathcal{V}_1\times\ldots\times\mathcal{V}_n} z^n_j(\sigma_1,\ldots\sigma_n) &:= -\left(\sigma_1,\ldots,\sigma_n\right) G^n_j \left(\sigma_1,\ldots,\sigma_n\right)^T, \\ \textit{where} \quad G^n_j = (g^n_{\mu,\nu})_{\mu,\nu} \quad \textit{with} \quad g^n_{\mu,\nu} &:= \begin{cases} a_{\mu,\nu}, & \mu = \nu \\ \dfrac{b_{\mu,\nu}}{2}, & \textit{else} \end{cases} \end{aligned} \tag{3.7}$$

Summing up, we have to solve a general constraint quadratic optimisation problem. This class of problems can be split up. The function z^n_j is not necessarily a convex function, since G^n_j does not have to be positive definite. Non-convex problems are known to be $\mathcal{NP}$-hard[2], cp. [PV91]. Problems of this class cannot be solved in polynomial time. Finding efficient algorithms to solve those problems is still a vast field in research. A survey on existing literature is given in the article of Floudas and Gounaris [FG09].
Meanwhile, for the convex case a solution can be found in polynomial time.
Furthermore, we have not taken into account, the different choices of difference quotients in the constants a_ν and $b_{\nu,\mu}$. This will be done in the two sections to follow.

[2]For a detailed introduction to $\mathcal{NP}$-problems we refer to the books of Garey and Johnson [GJ90] and Cormen et al. [CLRS01].

3.2.1 Centered difference quotients only

In the case where only centered difference quotients are used we have to solve only one optimisation problem per grid point. In this section we show how the optimisation can be done for $n = 2, 3$. These are the dimensions where Finite Difference Methods are practicable. For both cases we assume $\Theta_2 \subset \Theta_{cc}^{\bar{j}}$ and $\Theta_3 \subset \Theta_{ccc}^{\bar{j}}$, respectively. For the optimisation problem this means that there are no restrictions but the bounds for volatility and correlation.
The above condition can easily be tested at every grid point $\overline{x}_{\bar{j}}$ by checking if all vertices of $\mathcal{V}_1 \times \ldots \times \mathcal{V}_n$ satisfy the inequalities necessary for central differencing. For the two-dimensional problem the necessary inequalities are given by (2.13c) and (2.15c).
Another possibility is to choose the spatial step sizes a priori according to the bounds given in Section 2.2.2. So we have two criterions one of local and one of global character. For a uniform grid in a spatial direction these two are equivalent.

Options on two assets. We will use the Karush-Kuhn-Tucker (KKT) conditions to solve the corresponding optimisation problem. For an introduction to optimisation theory we refer to the books of Strang [Str86] and Fletcher [Fle00]. The proof of the following theorem which will be very useful for our task can be found in their books.

Theorem 3.2.
Let $z : \mathbb{R}^n \to \mathbb{R}$ be a smooth quadratic function. Furthermore, let $W \in \mathbb{R}^{m \times n}$ and $w \in \mathbb{R}^m$. Then there exists a vector $\lambda \in \mathbb{R}_+^m$, so that for every solution x^* of the constrained optimisation problem

$$\max z(x), \quad \textit{subject to} \quad Wx \leq w$$

the tuple (x^*, λ) satisfies the KKT conditions

$$\nabla z(x) + \lambda^T W = 0, \tag{3.8a}$$
$$(Wx - w)\lambda = 0, \tag{3.8b}$$
$$Wx \leq w,$$
$$\lambda \geq 0.$$

The converse statement of Theorem 3.2 does not have to be true for our type of

optimisation. For Problem 3.4, $n = 2$, we have

$$W := \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad w := \begin{pmatrix} \overline{\sigma_1} \\ \underline{\sigma_1} \\ \overline{\sigma_2} \\ \underline{\sigma_2} \end{pmatrix}.$$

and thus the KKT conditions are given by

$$(-2a\sigma - b_{1,2}\sigma_2, -2a_2\sigma_2 - b_{1,2}\sigma) + (\lambda_1 - \lambda_2, \lambda_3 - \lambda_4) = 0 \tag{3.9a}$$

$$(\sigma - \overline{\sigma_1})\lambda_1 + (\underline{\sigma_1} - \sigma)\lambda_2 + (\sigma_2 - \overline{\sigma_2})\lambda_3 + (\underline{\sigma_2} - \sigma_2)\lambda_4 = 0 \tag{3.9b}$$

$$W \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} \leq w$$

$$\lambda \geq 0$$

The procedure now is as follows. We determine all points $\xi_1, \ldots$ which solve the KKT-conditions and compare their objective values $z^2_{\bar{j}}(\xi_1), \ldots$ step by step.
There are three possible cases for the optimal solution $\gamma^{i,*}_{\bar{j}} \in \Theta_2$ that can occur:

2DO-a. It is at the vertices:
$\sigma_1^* \in \partial\mathcal{V}_1, \sigma_2^* \in \partial\mathcal{V}_2$

2DO-b. At one coordinate it is at the vertices and the other one on the edges except the vertices:
$\sigma_1^* \in \partial\mathcal{V}_1, \sigma_2^* \in \mathcal{V}_2^\circ \quad \text{or} \quad \sigma_1^* \in \mathcal{V}_1^\circ, \sigma_2^* \in \partial\mathcal{V}_2$

2DO-c. It is in the interior:
$(\sigma_1^*, \sigma_2^*) \in (\mathcal{V}_1 \times \mathcal{V}_2)^\circ$

In the third case 2DO-c all differences in (3.9b) are negative and thus we have $\lambda = 0$. From (3.9a) we get the optimal solution by solving

$$G^2_j \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} = 0$$

If G^2_j is of full rank the optimal solution would be zero which is not reasonable for volatility. Otherwise, if $\det G^2_j = 0 \Leftrightarrow 4a_2 = b^2_{1,2}$ we get the linear dependency $b_{1,2}\sigma_2 = 2a\sigma_1$ and thus $z^2_{\bar{j}}$ simplifies to a quadratic function in one variable. If the corresponding control to the minimum of $z^2_{\bar{j}}$ is feasible, it is taken into account.
In the case 2DO-b $z^2_{\bar{j}}$ only depends on one variable if we consider the other one fixed at a vertex. Via (3.9) we can express the free variable in dependence on the

fixed one. If the corresponding point is feasible the value of the objective function is calculated.
In the first case 2DO-a we simply compare all possible objective values and take those values of the controls where the minimum is attained.
Collectively, we have twelve different cases. Algorithm 3.2 gives a pseudocode to solve Problem 3.4 at every grid point.

Algorithm 3.2: Determination of the optimal control for the 2D-problem.

Input: $a_1, a_2, a_3, b'_{1,2}, \mathcal{V}_1, \mathcal{V}_2, \mathcal{C}_{1,2}$
Output: $\sigma_1, \sigma_2, \rho_{1,2}$

det_control_2D$\big(a_1, a_2, a_3, b'_{1,2}, \mathcal{V}_1, \mathcal{V}_2, \mathcal{C}_{1,2}\big)$
  if $b_{1,2} \leq 0$ **then**
    $b_{1,2} \leftarrow b'_{1,2}\underline{\rho_{1,2}}; \rho_{1,2} = \underline{\rho_{1,2}};$
  else
    $b_{1,2} \leftarrow b'_{1,2}\overline{\rho_{1,2}}; \rho_{1,2} = \overline{\rho_{1,2}};$
  $(\sigma_1, \sigma_2) \leftarrow \arg\min\{g_j^2(\underline{\sigma_1}, \underline{\sigma_2}), g_j^2(\underline{\sigma_1}, \overline{\sigma_2}), g_j^2(\overline{\sigma_1}, \underline{\sigma_2}), g_j^2(\overline{\sigma_1}, \overline{\sigma_2})\};$
  for $(x, y) = \left(\underline{\sigma_1}, \frac{-b_{1,2}}{2a_2}\underline{\sigma_1}\right), \left(\overline{\sigma_1}, \frac{-b_{1,2}}{2a_2}\overline{\sigma_1}\right), \left(\frac{-b_{1,2}}{2a}\underline{\sigma_2}, \underline{\sigma_2}\right), \left(\frac{-b_{1,2}}{2a}\overline{\sigma_2}, \overline{\sigma_2}\right)$ **do**
    if $(x, y) \in \mathcal{V}_1 \times \mathcal{V}_2$ **then**
      $(\sigma_1, \sigma_2) \leftarrow \arg\min\{g_j^2(\sigma_1, \sigma_2), g_j^2(x, y)\};$
  if $4a_2 = b_{1,2}^2$ **then**
    for $(x, y) = \left(\underline{\sigma_1}, \frac{2a}{b_{1,2}}\underline{\sigma_1}\right), \left(\overline{\sigma_1}, \frac{2a}{b_{1,2}}\overline{\sigma_1}\right), \left(\underline{\sigma_2}, \frac{b_{1,2}}{2a}\underline{\sigma_2}\right), \left(\overline{\sigma_2}, \frac{b_{1,2}}{2a}\overline{\sigma_2}\right)$ **do**
      if $(x, y) \in \mathcal{V}_1 \times \mathcal{V}_2$ **then**
        $(\sigma_1, \sigma_2) \leftarrow \arg\min\{g_j^2(\sigma_1, \sigma_2), g_j^2(x, y)\};$
  return $(\sigma_1, \sigma_2, \rho_{1,2})$;

At the grid points on the boundary we use the Neumann boundary conditions (2.20) to calculate $a_\mu, \mu = 1, 2$. Remember that the correlation is assumed to be zero on the boundary

Options on three assets. For the optimisation problem 3.4 for $n = 3$ we proceed analogously as before. The optimal control $\gamma_j^k \in \Theta_3$ is found by analysing four cases stemming from the KKT conditions (3.8). Of course, with out loss of generality we assume the optimal correlations $\rho_{1,2}^*, \rho_{1,3}^*$, and $\rho_{2,3}^*$ to be known. The optimal volatilities are determined according to the following four cases

3DO-a. they lie at the vertices of Θ_3:
$\sigma_1^* \in \partial\mathcal{V}_1$, $\sigma_2^* \in \partial\mathcal{V}_2$, and $\sigma_3^* \in \partial\mathcal{V}_3$,

3DO-b. they lie on the edges of Θ_3:
$\sigma_\mu^* \in \partial\mathcal{V}_\mu$, $\sigma_\nu^* \in \partial\mathcal{V}_\nu$, and $\sigma_\iota^* \in \mathcal{V}_\iota^\circ$, $\iota, \mu, \nu \in \{1,2,3\}$, $\mu, \nu \neq \iota, \mu < \nu$,

3DO-c. they lie on the faces of Θ_3:
$\sigma_\mu^* \in \partial\mathcal{V}_\mu$, $\sigma_\nu^* \in \mathcal{V}_\nu^\circ$, and $\sigma_\iota^* \in \mathcal{V}_\iota^\circ$, $\iota, \mu, \nu \in \{1,2,3\}, \mu < \nu$, $\iota \neq \mu, \nu$,

3DO-d. they lie in the interior of Θ_3 and:

i. $G_{\bar{j}}^3$ is of rank two,

ii. $G_{\bar{j}}^3$ is of rank one.

The first case 3DO-a is handled in the exact same way as before.
To analyse the next instance 3DO-a let σ_ν, σ_μ, and σ_ι be given as described in 3DO-b. Furthermore let the first two be fix. Inserting these assumptions into (3.8) we then obtain the result that σ_ι is given by

$$\sigma_\iota = -\frac{1}{2a_\iota}\left(b_{\iota,\mu}\sigma_\mu + b_{\iota,\nu}\sigma_\nu\right), \quad \iota \in \{1,2,3\}\backslash\{\mu,\nu\}$$

to satisfy the KKT conditions. If $\sigma_{\iota,j}$ is feasible we calculate the value of the objective function.
For the third case 3DO-c we consider the function in two-variables $z_{\bar{j}}^3(\sigma_{\mu,j}, \sigma_{\nu,j}, \overline{\sigma_{\iota,j}})$ and $z_{\bar{j}}^3(\sigma_{\mu,j}, \sigma_{\nu,j}, \underline{\sigma_{\iota,j}})$ over $(\mathcal{V}_\mu \times \mathcal{V}_\nu)^\circ$ for $\iota, \mu, \nu \in \{1,2,3\}, \mu \neq \nu$, $\iota \neq \mu, \nu$. Again from the KKT conditions, σ_μ and σ_ν have to satisfy a system of equations

$$\begin{pmatrix} 2a_\mu & b_{\mu,\nu} \\ b_{\mu,\nu} & 2a_\nu \end{pmatrix} \begin{pmatrix} \sigma_\mu \\ \sigma_\nu \end{pmatrix} = \begin{pmatrix} b_{\mu,\iota}\sigma_\iota \\ b_{\nu,\iota}\sigma_\iota \end{pmatrix}, \quad \sigma_\iota \in \{\underline{\sigma_\iota}, \overline{\sigma_\iota}\}. \tag{3.10}$$

This system can be solved as in the two-dimensional case. If it has a unique and feasible solution, we calculate the corresponding value of the objective function. Otherwise, since also G_j^3 is not of full rank, we actually are in a situation as in case 3DO-d.
The fourth case 3DO-d is clearly more difficult. We omitted the case when G_j^3 is of full rank. The solution $\sigma_1^* = \sigma_2^* = \sigma_3^* = 0$ can be excluded w. l. o. g., since it has no financial interpretation.
Let $\tilde{G}_{\bar{j}}^3 = (\tilde{g}_{k,l})$ be the equivalent upper triangular matrix of $G_{\bar{j}}^3 =$ where we w. l. o. g. that no permutations have been used. If $G_{\bar{j}}^3$ has rank two $z_{\bar{j}}^3$ reduces to a function in one variable: $z_{\bar{j}}^3(\tilde{\sigma}_1, \tilde{\sigma}_2, \sigma_3)$, where

$$\tilde{\sigma}_1 := -\frac{1}{\tilde{g}_{1,1}}\left(\tilde{g}_{1,2}\tilde{\sigma}_2 + \tilde{g}_{1,3}\sigma_3\right), \quad \tilde{\sigma}_2 := -\frac{\tilde{g}_{2,2}}{\tilde{g}_{2,3}}\sigma_3, \tag{3.11}$$

and σ_3 is free. Thus, we have a one-dimensional function which has to be minimised over $\mathcal{V}_3$. This task can be done with standard methods.
If G_j^3 has rank one the objective function reduces to a function with two unknowns: $z_j^3(\tilde{\sigma}_1, \sigma_2, \sigma_3)$, where

$$\tilde{\sigma}_1 := -\frac{1}{\tilde{g}_{1,1}^3} \left(g_{1,2}^3 \sigma_2 + \tilde{g}_{1,3}^3 \sigma_3 \right),$$

and σ_2 and σ_3 are free. Therefore, we have a quadratic optimisation problem in two variables with box constraints given by the set $\mathcal{V}_2 \times \mathcal{V}_3$. With adjusted coefficients, this problem can be solved by Algorithm 3.2.

The above deliberations are summarised in Algorithm 3.3.

At the boundary points we use the information of the Neumann boundary conditions to calculate the optimal controls. The three correlations are set to zero.

Remark 3.5.
By the detailed analysis of the three-dimensional optimisation Problem, we hope to save some effort. Of course, the problem could be treated approximately by suitable methods. But since the problem is in general $\mathcal{NP}$-hard this does not seem like a promising approach. Especially, since the optimisation has to be done many times just for one time step. Proceeding the way described above though, we obtain the optimum at a single grid point in a known number of steps.

3.2.2 Mixed difference quotients

In Section 2.2.1 we learned that for certain sets of controls and / or to big step sizes we have to use different difference quotients for the first order derivative in order to obtain a discretisation matrix with the M-matrix property. This choice again determines the objective function of Problem 3.4. Thus, if we cannot guarantee the exclusive use of centered difference quotients, special care has to be taken when solving Problem 3.4.
In this subsection we focus on the optimisation problem for $n = 2$. The three-dimensional analogon can be treated in the same way with some additional computational effort.

In the following deliberations we consider a fixed grid point with lexicographical index $\bar{j}$ and therefore omit it where possible.

Algorithm 3.3: Determination of the optimal control for the 3D-problem.

Input: $a_1, a_2, a_3, b'_{1,2}, b'_{1,3}, b'_{2,3}, \mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \mathcal{C}_{1,2}, \mathcal{C}_{1,3}, \mathcal{C}_{2,3}$
Output: $\sigma_1, \sigma_2, \sigma_3, \rho_{1,2}, \rho_{1,3}, \rho_{2,3}$

det_control_3D$\big(a_1, a_2, a_3, b'_{1,2}, b'_{1,3}, b'_{2,3}, \mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \mathcal{C}_{1,2}, \mathcal{C}_{1,3}, \mathcal{C}_{2,3}\big)$
  for $(\mu, \nu) \in \{(1,2),(1,3),(2,3)\}$ **do**
    if $b_{\mu,\nu} \leq 0$ **then**
      $b_{\mu,\nu} \leftarrow b'_{\mu,\nu}\underline{\rho_{\mu,\nu}}$; $\rho_{\mu,\nu} = \underline{\rho_{\mu,\nu}}$;
    else
      $b_{\mu,\nu} \leftarrow b'_{\mu,\nu}\overline{\rho_{\mu,\nu}}$; $\rho_{\mu,\nu} = \overline{\rho_{\mu,\nu}}$;
  for *all vertices* (x, y, z) *of* $\mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{V}_3$ **do**
    $(\sigma_1, \sigma_2, \sigma_3) \leftarrow \arg\min\{g_j^3(\sigma_1, \sigma_2, \sigma_3), g_j^3(x, y, z)\}$;
  for $(x, y) \in \Big\{(\underline{\sigma_\mu}, \underline{\sigma_\nu}), (\overline{\sigma_\mu}, \underline{\sigma_\nu}), (\underline{\sigma_\mu}, \overline{\sigma_\nu}), (\overline{\sigma_\mu}, \overline{\sigma_\nu}) | \mu, \nu \in \{1,2,3\}, \mu < \nu\Big\}$ **do**
    $\iota \leftarrow \{1,2,3\}\backslash\{\mu, \nu\}$; $z \leftarrow -(2a_\iota)^{-1}(b_{\iota,\mu}x + b_{\iota,\nu}y)$;
    if $z \in \mathcal{V}_\iota$ **then**
      $(\sigma_1, \sigma_2, \sigma_3) \leftarrow \arg\min\{g_j^3(\sigma_1, \sigma_2, \sigma_3), g_j^3(x, y, z)\}$;
  for $(x, y, z) \in \big\{(\underline{\sigma_\iota}, \sigma_\mu, \sigma_\nu), (\overline{\sigma_\iota}, \sigma_\mu, \sigma_\nu) | \mu, \nu, \iota \in \{1,2,3\} \mu < \nu, \iota \neq \mu, \nu\big\}$ **do**
    if $b_{\mu,\nu} \neq 2a_\mu, 2a_\nu$ **then**
      calculate (y, z) as solution of equation system (3.10);
      if $(y, z) \in \mathcal{V}_\mu \times \mathcal{V}_\nu$ **then**
        $(\sigma_1, \sigma_2, \sigma_3) \leftarrow \arg\min\{g_j^2(\sigma_1, \sigma_2, \sigma_3), g_j^2(x, y, z)\}$;
  if $rank(G_j^3) = 1$ **then**
    $(x, y, z) \leftarrow$ **det_control_2D**$(\cdot)$;
    $(\sigma_1, \sigma_2, \sigma_3) \leftarrow \arg\min\{g_j^2(\sigma_1, \sigma_2, \sigma_3), g_j^2(x, y, z)\}$;
  else if $rg(G_j^3) = 2$ **then**
    $(\sigma_1, \sigma_2, \sigma_3) \leftarrow \arg\min\{g_j^2(\sigma_1, \sigma_2, \sigma_3), g_j^2(x, y, z)\}$ with $y = -(\tilde{g}_{1,1}^3)^{-1}\tilde{g}_{2,3}^3 z$ and $x = -(\tilde{g}_{1,1}^3)^{-1}(\tilde{g}_{1,3}^3 z + \tilde{g}_{1,2}^3 y)$; // cp. (3.11)
  return $(\sigma_1, \sigma_2, \sigma_3, \rho_{1,2}, \rho_{1,3}, \rho_{2,3})$;

Continuous optimisation. For now, let us assume that we know a subset

$$\overline{\Theta} \subset \{\Theta_{ff}, \Theta_{bb}, \Theta_{cc}, \Theta_{fb}, \Theta_{fc}, \Theta_{bf}, \Theta_{bc}, \Theta_{cf}, \Theta_{cb}\},$$

with non-empty elements for every grid point (x_1^k, x_2^l). We have

$$\Theta = \bigcup_{\overline{\theta} \subset \overline{\Theta}} \overline{\theta}.$$

In opposite to the previous sections we cannot determine $\rho_{1,2}$ a priori. The choice of the first order difference quotient might also affect the sign of the mixed derivative and therefore also the correlation ρ_{1_2}. The foregoing is as follows, we split up

Problem 3.4 into p different optimisation problems.

Problem 3.6.
Solve

$$\min_{\substack{(\sigma_1,\sigma_2,\rho_{1,2})\in\overline{\Theta}_{uw}\\ \overline{\Theta}_{uw}\subset\overline{\Theta}}} \sigma_1^2\left(\delta_{x_1}^2 v_{k,l}^i - \delta_{x_1}^u v_{k,l}^i\right) + \sigma_1\sigma_2\rho_{1,2}\delta_{x_1,x_2}^+ v_{k,l}^i + \sigma_2^2\left(\delta_{x_2}^2 v_{k,l}^i - \delta_{x_2}^w v_{k,l}^i\right),$$

for $u, w \in \{f, b, c\}$.

So, we have to solve at most nine quadratic optimisation problems with non-linear constraints for each grid point. Problem 3.6 consists of an objective function and constraints which are non-convex and the additional box constraints. Thus, it is contained in the class of $\mathcal{NP}$-hard problems, see above (v. s.) Problem 3.4.
Compared to the previous section this is not only a strong increase concerning the number of problems but also in the difficulty of solving it.

Let $\gamma_{uw} \in \Theta_{uw} \subset \overline{\Theta}$ be the optimal solution of Problem 3.6. Then, we set

$$(\sigma_1^*, \sigma_2^*, \rho_{1,2}^*) = \underset{\substack{\gamma_{uw}\in\Theta_{uw}\\ \Theta_{uw}\in\overline{\Theta}}}{\arg\min}\, z_{uw}(\gamma_{uw}), \tag{3.12}$$

where we denote by $z_{uw}(\gamma_{uw})$ the objective function minimised in Problem 3.6. The tuple (u, w) for which the optimal value is obtained determines the difference quotients which are used at the actual grid point: in x_1- direction we use $\delta_{x_1}^u$ and in x_2-direction we use $\delta_{x_2}^w$

The remaining question is convergence of Algorithm 3.1. The key property of the determined controls was

$$-B_{\gamma,\theta}^{i,(j+1)} v^{i,(j)} + r_{\gamma,\theta}^{i,(j+1)} - \left(-B_{\gamma,\theta}^{i,(j)} v^{i,(j)} + r_{\gamma,\theta}^{i,(j)}\right) \geq 0,$$

which was guaranteed by relation (3.2), cp. (3.3). The sets $\Theta_{uw} \subset \overline{\Theta}$ are independent of the actual option value $v^{i,(j)}$. Therefore, for each $\Theta_{uw} \subset \overline{\Theta}$ and $\gamma \in \Theta_{uw}$ optimal the above inequality would hold. This is also true for the tuple (u, w) chosen as in (3.12).
The uniqueness of the solution can be proven as in the proof of Theorem 3.1.
So, the following lemma holds.

Lemma 3.3.
For a time step $t^{i-1} \to t^i$ *let the solution* v^{i-1} *and the control* $\gamma^{i-1} \in \Theta_2^{M^i}$ *of the*

previous time level and an iterate $v^{i,(j)}$ *and controls* $\gamma^{i,(j)}$, $\gamma^{i,(j-1)} \in \Theta_2^{M^i}, j > 1$, *calculated via* (3.12) *be given.*
Then Algorithm 3.1 converges monotonously and in a finite number of iterations to a unique solution.

The effort to determine which sets Θ_{uw} are non-empty and to solve the corresponding non-linear optimisation problems for each grid point in each iteration would possibly burst runtime. Therefore we relax the optimisation problems if several difference quotients might be used.

Relaxed optimisation. We will only solve a relaxed version of Problem 3.6, which does not require explicit optimisation any longer. For this purpose we choose a finite set of points

$$\mathcal{D}_2 := \left\{ (\xi_1^i, \xi_2^i, \xi_3) \;\middle|\; \xi_\mu^i = \underline{\sigma_\mu} + i\frac{\overline{\sigma_\mu} - \underline{\sigma_\mu}}{cM_\mu},\ i = 0, \dots, cM_\mu, \mu = 1, 2, \xi_3 \in \{\underline{\rho_{1,2}}, \overline{\rho_{1,2}}\} \right\}, \tag{3.13}$$

where $cM_1, cM_2 \in \mathbb{N}$. We call $\mathcal{D}_2 \subset \Theta_d$ the discrete set of admissible controls. We have to make one further assumption if we only allow $\xi_3 \in \{\underline{\rho_{1,2}}, \overline{\rho_{1,2}}\}$ in (3.13). This is necessary because the choice of the difference quotient and that of ξ_3 might depend on each other. Assume that for a specific selection of difference quotients $\overline{\rho_{1,2}}$ is optimal because of the sign of $\delta^+_{x_1x_2}$. But for $\overline{\rho_{1,2}}$ these difference quotients might not result in a monotone discretisation. These cases are excluded by

Assumption 3.7.
Assume that for all grid points there are no $\rho^+_{1,2}, \rho^-_{1,2} \in \mathcal{C}_{1,2}$ *for which*

$$\operatorname{sign}\left(\frac{2}{h_1^{k+1} + h_1^k} d^{k,l}_{1,\gamma}(\rho^-_{1,2}) + c^{k,l}_{1,\gamma}(\rho^-_{1,2})\right) \neq \operatorname{sign}\left(\frac{2}{h_1^{k+1} + h_1^k} d^{k,l}_{1,\gamma}(\rho^+_{1,2}) + c^{k,l}_{1,\gamma}(\rho^+_{1,2})\right)$$

and

$$\operatorname{sign}\left(\frac{2}{h_1^{k+1} + h_1^k} d^{k,l}_{1,\gamma}(\rho^-_{1,2}) - c^{k,l}_{1,\gamma}(\rho^-_{1,2})\right) \neq \operatorname{sign}\left(\frac{2}{h_1^{k+1} + h_1^k} d^{k,l}_{1,\gamma}(\rho^+_{1,2}) - c^{k,l}_{1,\gamma}(\rho^+_{1,2})\right)$$

It is also assumed that these two conditions hold for the x_2*-direction.*

If Assumption 3.7 is fulfilled the problem described above does not exist any longer.

Example 3.8.
For the following two cases Assumption 3.7 is fulfilled

1. *The grid is uniform in each spatial direction. In this situation $c^{k,l}_{\cdot,\gamma}$ does not depend on correlation.*

2. *Correlation is not uncertain.*

Now, instead of solving the original problem for every element of $\overline{\Theta}$, we look for the point where the minimal value of the objective function restricted to $\mathcal{D}_2$ is attained

$$(\sigma_1^*, \sigma_2^*, \rho_{1,2}^*) = \underset{\substack{q \in \mathcal{D}_2 \cap \Theta_{uw} \\ \Theta_{uw} \subset \overline{\Theta}}}{\arg\min}\ z_{uw}^2(q) \tag{3.15}$$

As in the continuous case inequality (3.3) is satisfied, but only for the discrete set $\mathcal{D}_2$. Thus we have

Corollary 3.4.
For a time step $t^{i-1} \to t^i$ let the solution v^{i-1}, the control $\gamma_{\overline{j}}^{i-1} \in \Theta_{uw} \cap \mathcal{D}_2$ of the previous time level, and an iterate $v^{i,(j)}$ and controls $\gamma_{\overline{j}}^{i,(j)}$, $\gamma_{\overline{j}}^{i,(j-1)} \in \Theta_{uw} \cap \mathcal{D}_2, j > 1$, $u, w \in \{b, c, f\}$, for all $\overline{j} \in J_M^i$, calculated via (3.15) be given. Furthermore, let Assumption 3.7 hold.
Then Algorithm 3.1 converges monotonously and in a finite number of iterations to a unique solution.

Given the set $\mathcal{D}_2$ we can a priori determine for every grid point which difference quotients have to be used to preserve monotonicity of the iteration matrix, provided Assumption 3.7 holds. For this purpose we check conditions (2.13) and (2.15), respectively, for each grid point. During the run of the final algorithm the optimisation reduces to calculation of at most $(cM_1 + 1)(cM_2 + 1)$ evaluations of the objective function and comparisons.

Remark 3.9.
For the relaxed optimisation problems the optimal value of (3.1) is in general not attained. But the approximated one is at least as good as if we would calculate with constant volatilities and correlation. Thus, we would expect the option value calculated with relaxed optimisation to lie between those of the linear BS model and of the non-linear BSB model.

4 Convergence of schemes

In this chapter we analyse the convergence for the Finite Difference schemes constructed in Chapter 2. In Section 1.2 the theoretical framework to prove convergence to the viscosity solution of the BSB equation was provided. For each scheme, consistency, stability and monotonicity have to be verified. We consider BDF1 and CN schemes. For the BDF2 scheme convergence cannot be proven within the theoretical framework at hand, cp. Remark 2.27.
In the first subsections the convergence of schemes for the two-dimensional pricing problem is treated. All the proofs given here are to the knowledge of the author new to the literature and are therefore given in more detail. Then we focus on the schemes for one-dimensional pricing problems. Here convergence results have been established, cf. Heider [Hei10] and Schaeling [Sch10]. In both subsections we only consider the pricing of European options.
In the last subsection we then take a general look at the convergence of schemes for pricing American options.

For the whole chapter we assume that the discretisation is constructed on a grid $\mathcal{G}(L, \{t^i\}, \mathcal{S}, \Omega_2)$, cp. Notation 2.5. The grid in time however might be non-uniform. The set of controls is given by Θ_2, cp. Notation 2.7. In order to properly denote which optimization technique is used we add a superscript to the index of the discretisation. By $D^{c,i}_{\bar{j},\gamma}(v^i, \tau^i, \mathcal{S}^i, \tau^{i-1}, \mathcal{S}^{i-1}, v^{i-1}) = 0$ we denote the discretisation where continuous optimisation is used. Otherwise, "c" is replaced by "d".
We remind the reader that the BDF1 and the CN methods for European option pricing can be denoted in the form

$$B^i_{\gamma,\theta} v^i = C^{i-1}_{\gamma,\theta} v^{i-1} + r^i_\theta, \tag{4.1a}$$

with

$$B^i_{\gamma,\theta} = (I + \theta h_\tau A^i_\gamma), \tag{4.1b}$$

$$C^{i-1}_{\gamma,\theta} = (I - (1-\theta) h_\tau A^{i-1}_\gamma), \quad \text{and} \tag{4.1c}$$

$$r^i_\theta = h_\tau \left(\theta b^i + (1-\theta) b^{i-1}\right), \tag{4.1d}$$

cp. equations (2.40) and (2.41). The BDF1 scheme is obtained for $\theta = 1$ and the CN scheme for $\theta = \frac{1}{2}$. The matrix A^i_γ and the vector b^i, $i \in I_{L-1}$, depend on the dimension of the problem and the method we used for the discretisation, cp. Chapter 2. Additional to consistency, stability and monotonicity we need the equation itself to fulfil a maximum principle. Thus, we make the following

Assumption 4.1.
For all $n \in \mathbb{N}$ for two solutions $u(t,x), v(t,x)$ of the BSB equation we have

$$u \leq v,$$

whenever $u(0,\cdot) \leq v(0,\cdot)$.

4.1 Convergence of two-dimensional schemes

In this section we prove the convergence of the BDF1 and the CN scheme for the two-dimensional pricing problem under certain restrictions regarding the step size $h^i_\tau, i \in I_{L-1}$. For the first method the result is quite clear from the work that has already been done. For the latter one a more thorough analysis has to be conducted. The following theorems are formulated for those problems where only continuous optimisation, cp. Sections 2.2.2 and 3.2.1, can be used. For those cases where the optimisation might be done over $\mathcal{D}_2$ we have to make some additional effort. Not only the convergence of the discretisation in time and space has to be analysed, but also the one for the discrete optimisation, i. e. $cM_1, cM_2 \to \infty$. Having established the necessary framework, we state the convergence results in seperate corollars, v. i.

Convergence for continuous optimisation. To assure that only continuous optimisation techniques, i. e. Algorithm 3.2, are used, the spatial step sizes have to be bounded as stated in (2.25) and (2.27), respectively.

Theorem 4.1 (Convergence of the BDF1 scheme)**.**
Let a grid $\mathcal{G}(L, \{t^i\}, \mathcal{S}, \Omega_2)$ which satisfies Assumption 2.9 be given. The spatial step sizes are assumed to be chosen in such a way that only centered difference quotients may be used. Furthermore, let $r \geq 0$ and let Assumption 3.7 hold.
Then the BDF1 scheme given by

$$D^{c,i}_\gamma\left(v^i, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1}\right) = B^i_{\gamma,1} v^i - C^{i-1}_{\gamma,1} v^{i-1} - r^i_1 = 0$$

for a time step $t^{i-1} \to t^i$ converges to the unique viscosity solution of the Black-Scholes-Barenblatt equation (2.4) for $\max_i\{h^i_\tau\}, H \to 0$.

Proof.

Since all conditions of Lemma 1.5 are fulfilled we know that the BDF1 scheme is monotone. For the spatial discretisation only central difference quotients are used, since (2.25) *and* (2.27) *hold , respectively. Thus, for every time step we find a unique solution* v^i *for the next time level* t^i *by Algorithm 3.1 since* $B^i_{\gamma,1}$ *is an M-matrix on* $\mathcal{G}$*, see Lemma 2.10 and Theorem 3.1 with* $C^{i-1}_{\gamma,1}$ *and* $D = 0$.

The stability of the method is proven as follows. We consider equation (4.1) for $\theta = 1$ *at an inner grid point* (t^i, x_1^k, x_2^l).

Furthermore, we assume $|v^i_{k,l}| = \max_{\mu\in J_{M_1},\nu\in J_{M_2}}\{v^i_{\mu,\nu}\}$. *Using the notation of the difference stencil (2.9), the method is given by*

$$(1 + h^i_\tau a^{k,l}_{k,l,\gamma})v^i_{k,l} = -h^i_\tau \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} a^{\mu,\nu}_{k,l,\gamma} v^i_{\mu,\nu} + v^{i-1}_{k,l} + (r^i_1)_{k,l}$$

Since $a^{k,l}_{k,l,\gamma} \geq 0$ *and* $-a^{\mu,\nu}_{k,l,\gamma} \geq 0, (\mu,\nu) \in \mathcal{N}(k,l)$, *by the choice of the grid, and* $(r^i_1)_{k,l} = 0$ *we have*

$$(1 + h^i_\tau a^{k,l}_{k,l,\gamma})\|v^i\|_\infty \leq h^i_\tau \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} -a^{\mu,\nu}_{k,l,\gamma}\|v^i\|_\infty + \|v^{i-1}\|_\infty$$

By rearranging terms and using the definiton of $a^{k,l}_{k,l,\gamma}$ *in (2.18) we obtain*

$$\left(1 + h^i_\tau \underbrace{\left(a^{k,l}_{k,l,\gamma} + \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} a^{\mu,\nu}_{k,l,\gamma}\right)}_{=r}\right)\|v^i\|_\infty \leq \|v^{i-1}\|_\infty.$$

We now conclude

$$\|v^i\|_\infty \leq \frac{1}{1 + h^i_\tau r}\left\|v^{i-1}\right\|_\infty \leq \|v^{i-1}\|_\infty,$$

since $r \geq 0$.

This argumentation also holds true for the boundary points of the grid. At the boundary points of the grid a similar appraisal con be done, taking into consideration the definition of the boundary conditions given by r^i_1.

Thus, since v^0 *is bounded by construction all iterates* $v^i, i \in J_L$, *are bounded.*

The consistency of the scheme is guaranteed, see for example [SB02].

Thus with Assumption 3.7 and by Theorem 1.4 of Barles and Souganidis the BDF1 scheme converges to the unique viscosity solution.

□

To prove the convergence of the CN scheme to the viscosity solution we will proceed in several steps. First, we prove a technical result which will be needed for monotonicity and stability results. Subsequently, we prove the actual convergence result.

The following lemma gives sufficient conditions to guarantee $1 - \frac{h_\tau^i}{2} a_{k,l,\gamma}^{k,l} \geq 0$ to hold true. This property will be needed to appraise the explicit terms of the CN scheme.

Lemma 4.2.
Let a grid $\mathcal{G}(L, \{t^i\}, \mathcal{S}, \Omega_2)$ which satisfies Assumption 2.9 be given. Let the constants $a_{k,l,\gamma}^{\cdot}$ of the stencil (2.9) be given as defined in (2.18). Then we have

$$1 - \frac{h_\tau^i}{2} a_{k,l,\gamma}^{k,l} \geq 0, \quad \text{for } k \in J_{M_1}, l \in J_{M_2}, \tag{4.2}$$

if

$$\max\{c_{ff}, c_{fc}, c_{fb}, c_{cf}, c_{cc}, c_{cb}, c_{bf}, c_{bc}, c_{bb}\} + H^2 r + \overline{\rho}_{1,2}\overline{\sigma}_1\overline{\sigma}_2 \leq \frac{2\tilde{h}_x^2}{h_\tau}, \tag{4.3}$$

where $0 \leq \tilde{h}_x \leq h_1^k, h_2^l \leq H, k \in I_{M_1}, l \in I_{M_2}$ and

$$\begin{aligned}
c_{ff} &:= -f_{x_1} - f_{x_2} & c_{fc} &:= -f_{x_1} - c_{x_2} & c_{fb} &:= -f_{x_1} - b_{x_2} \\
c_{cf} &:= -c_{x_1} - f_{x_2} & c_{cc} &:= -c_{x_1} - c_{x_2} & c_{cb} &:= -c_{x_1} - b_{x_2} \\
c_{bf} &:= -b_{x_1} - f_{x_2} & c_{bc} &:= -b_{x_1} - c_{x_2} & c_{bb} &:= -b_{x_1} - b_{x_2}
\end{aligned}$$

are constants originating from the combined use of forward (f), centered (c), and backward (b) difference quotients in the two space dimensions. They only depend on the bounds of volatilities and correlation.

The exact definitions of the above constants are given in the following proof.

Proof.
We will not consider all nine possible combinations of forward, centered, and backward difference quotients. Instead the three choices of the same difference quotients in each space direction are analysed and combined to one part of appraisal (4.3) each.
Before we start to derive the appraisals we recollect that for the grid in consideration the inequalities

$$\Xi_{1,2}^d \leq \frac{h_1^\mu}{h_2^\nu} \leq \Xi_{1,2}^u \quad \text{and} \quad \frac{1}{\Xi_{1,2}^u} \leq \frac{h_2^\nu}{h_1^\mu} \leq \frac{1}{\Xi_{1,2}^d} \tag{4.4}$$

hold for all $\mu \in J_{M_1}, \nu \in J_{M_2}$ by Assumption 2.9.

For the rest of the proof let (k,l) be a fixed index of an inner grid point. First we consider the usage of backward differentiation $\delta^b v^i_{k,l}$ for the first derivative. By (2.18) and definition of $\tilde{h}_x$ it is

$$\begin{aligned} a^{k,l}_{k,l} &= r - \frac{2d^{k,l}_{1,\bar{j}}}{h^k_1 h^{k+1}_1} - \frac{2d^{k,l}_{2,\bar{j}}}{h^l_2 h^{l+1}_2} + \frac{c^{k,l}_{1,\bar{j}} h^{k+1}_1}{h^k_1 h^{k+1}_1} + \frac{c^{k,l}_{2,\bar{j}} h^{l+1}_2}{h^l_2 h^{l+1}_2} - a^\gamma_1 - a^\gamma_2 \\ &\leq r - \frac{1}{\tilde{h}^2_x}\left((2d^{k,l}_{1,\bar{j}} - c^{k,l}_{1,\bar{j}} h^{k+1}_1) + (2d^{k,l}_{2,\bar{j}} - c^{k,l}_{2,\bar{j}} h^{l+1}_2) - \rho_{1,2,\bar{j}} \sigma_{1,\bar{j}} \sigma_{2,\bar{j}} \right) \end{aligned} \tag{4.5}$$

Now we separately appraise each of the four summands from above. We only use backward differentiation if (2.13b) and (2.15b) hold true. Thus all four summands are assumed to be non-positive. We have

$$\begin{aligned} 0 &\geq 2d^{k,l}_{1,\bar{j}} - c^{k,l}_{1,\bar{j}} h^{k+1}_1 \\ &= -\sigma^2_{1,\bar{j}} - h^{k+1}_1 \left(\frac{1}{2}\sigma^2_{1,\bar{j}} - r + \delta_1 \right) + \frac{1}{2}\rho_{1,2,\bar{j}} \sigma_{1,\bar{j}} \sigma_{2,\bar{j}} \frac{h^k_1 + h^{k+1}_1}{h^l_2} \\ &\overset{(4.4)}{\geq} -\sigma^2_{1,\bar{j}} \left(1 + \frac{h^{k+1}_1}{2} \right) + h^{k+1}_1 (r - \delta_1)^- + \Xi^d_{1,2} \rho_{1,2,\bar{j}} \sigma_{1,\bar{j}} \sigma_{2,\bar{j}} \\ &\geq -\overline{\sigma}^2_1 \left(1 + \frac{H}{2} \right) + H(r - \delta_1)^- + \underline{\rho}_{1,2} \overline{\rho}_{1,2} \underline{\sigma}_1 \overline{\sigma}_1 =: b_{x_1} \end{aligned} \tag{4.6}$$

and

$$\begin{aligned} 0 &\geq 2d^{k,l}_{2,\bar{j}} - c^{k,l}_{2,\bar{j}} h^{l+1}_2 \\ &= -\sigma^2_{2,\bar{j}} - h^{l+1}_2 \left(\frac{1}{2}\sigma^2_{2,\bar{j}} - r + \delta_2 \right) + \frac{1}{2}\rho_{1,2,\bar{j}} \sigma_{1,\bar{j}} \sigma_{2,\bar{j}} \frac{h^l_2 + h^{l+1}_2}{h^k_1} \\ &\overset{(4.4)}{\geq} -\sigma^2_{2,\bar{j}} \left(1 + \frac{h^{l+1}_2}{2} \right) + h^{l+1}_2 (r - \delta_2)^- + \frac{1}{\Xi^u_{1,2}} \rho_{1,2,\bar{j}} \sigma_{1,\bar{j}} \sigma_{2,\bar{j}} \\ &\geq -\overline{\sigma}^2_2 \left(1 + \frac{H}{2} \right) + H(r - \delta_2)^- + \underline{\rho}_{1,2} \overline{\rho}_{1,2} \underline{\sigma}_2 \overline{\sigma}_2 =: b_{x_2} \end{aligned} \tag{4.7}$$

We now insert (4.6) *and* (4.7) *into (4.5) to obtain*

$$a^{k,l}_{k,l,\gamma} \leq r - \frac{1}{\tilde{h}^2_x} \left(b_{x_1} + b_{x_2} - \overline{\rho}_{1,2} \overline{\sigma}_1 \overline{\sigma}_2 \right)$$

where we used $h^{k+1}_1, h^{l+1}_2 \leq H$. By claiming condition (4.3), constraint (4.2) is ensured to hold true since

$$\frac{1}{2} \left(c_{bb} + H^2 r + \overline{\rho}_{1,2} \overline{\sigma}_1 \overline{\sigma}_2 \right) h^i_\tau \leq \tilde{h}^2_x \;\Rightarrow\; \frac{h^i_\tau}{2} a^{k,l}_{k,l,\gamma} \leq 1 \tag{4.8}$$

can easily be verified with $\tilde{h}_x \leq 1$.

The argumentation for the use of forward differentiation applies a similar argumentation as in the case presented above. We only state the main steps. For the element $a_{k,l}^{k,l,\gamma}$ *the appraisal*

$$a_{k,l,\gamma}^{k,l} \leq r - \frac{1}{\tilde{h}_x^2}\left((2d_{1,\bar{j}}^{k,l} + c_{1,\bar{j}}^{k,l}h_1^{k+1}) + (2d_{2,\bar{j}}^{k,l} + c_{2,\bar{j}}^{k,l}h_2^{l+1}) - \rho_{1,2,\bar{j}}\sigma_{1,\bar{j}}\sigma_{2,\bar{j}}\right) \tag{4.9}$$

is valid. All summands are non-positive since we can assume (2.13a) and (2.15a) to hold true. The first two can be gauged by

$$0 \geq 2d_{1,\bar{j}}^{k,l} + c_{1,\bar{j}}^{k,l}h_1^k \geq -\overline{\sigma}_1^2 - H(r-\delta_1)^+ + \underline{\rho}_{1,2}\overline{\rho}_{1,2}\underline{\sigma}_1\overline{\sigma}_1 =: f_{x_1} \tag{4.10}$$

and

$$0 \geq 2d_{2,\bar{j}}^{k,l} + c_{2,\bar{j}}^{k,l}h_2^l \geq -\overline{\sigma}_2^2 - H(r-\delta_2)^+ + \underline{\rho}_{1,2}\overline{\rho}_{1,2}\underline{\sigma}_2\overline{\sigma}_2 =: f_{x_2}. \tag{4.11}$$

Inserting these two appraisals into (4.9) we obtain

$$a_{k,l,\gamma}^{k,l} \leq r - \frac{1}{\tilde{h}_x^2}\left(f_{x_1} + f_{x_2} - \overline{\rho}_{1,2}\overline{\sigma}_1\overline{\sigma}_2\right).$$

If condition (4.3) holds true, with $\tilde{h}_x \leq 1$*, and by the implication*

$$\frac{1}{2}\left(c_{ff} + H^2 r + \overline{\rho}_{1,2}\overline{\sigma}_1\overline{\sigma}_2\right)h_\tau^i \leq \tilde{h}_x^2 \;\Rightarrow\; \frac{h_\tau^i}{2}a_{k,l,\gamma}^{k,l} \leq 1, \tag{4.12}$$

constraint (4.2) is satisfied.

For the use of central differentiation we proceed in the same way as before. Only the terms become a little longer. For the diagonal entry the inequality

$$\begin{aligned} a_{k,l,\gamma}^{k,l} \leq r - \frac{1}{\tilde{h}_x^2}\Big(&\left(2d_{1,\bar{j}}^{k,l} + c_{1,\bar{j}}^{k,l}\left(h_1^k - h_1^{k+1}\right)\right) + \left(2d_{2,\bar{j}}^{k,l} + c_{2,\bar{j}}^{k,l}\left(h_2^l - h_2^{l+1}\right)\right) \\ &- \rho_{1,2,\bar{j}}\sigma_{1,\bar{j}}\sigma_{2,\bar{j}}\Big) \end{aligned} \tag{4.13}$$

is valid. The first summand can be appraised by

$$\begin{aligned} 0 &\geq 2d_{1,\bar{j}}^{k,l} + c_{1,\bar{j}}^{k,l}\left(h_1^k - h_1^{k+1}\right) \\ &\geq -\sigma_{1,\bar{j}}^2 - \left(h_1^k - h_1^{k+1}\right)\left(\frac{1}{2}\sigma_{1,\bar{j}}^2 - r + \delta_1\right) \\ &\qquad + \frac{1}{2}\rho_{1,2,\bar{j}}\sigma_{1,\bar{j}}\sigma_{2,\bar{j}}\left(\frac{h_1^k - h_1^{k+1}}{h_2^l} - \frac{h_1^k - h_1^{k+1}}{h_2^{l+1}}\right) \end{aligned}$$

$$\geq -\sigma_{1,\bar{j}}^2\left(1+\frac{h_1^k}{2}\right) - H|r-\delta_1| + \left(\Xi_{1,2}^d - \Xi_{1,2}^u\right)\rho_{1,2,\bar{j}}\sigma_{1,\bar{j}}\sigma_{2,\bar{j}}$$
$$\geq -\overline{\sigma}_1^2\left(1+\frac{H}{2}\right) - H|r-\delta_1| + \left(\underline{\rho}_{1,2}\overline{\rho}_{1,2}-1\right)\underline{\sigma}_1\overline{\sigma}_1 =: c_{x_1} \tag{4.14}$$

and in an analogue manner we obtain

$$0 \geq 2d_{2,\bar{j}}^{k,l} + c_{2,\bar{j}}^{k,l}\left(h_2^l - h_2^{l+1}\right)$$
$$\geq -\overline{\sigma}_2^2\left(1+\frac{H}{2}\right) - H|r-\delta_2| + \left(\underline{\rho}_{1,2}\overline{\rho}_{1,2}-1\right)\underline{\sigma}_2\overline{\sigma}_2 =: c_{x_2}. \tag{4.15}$$

Combining these two inequalities and (4.13) with each other, we obtain the appraisal

$$a_{k,l,\gamma}^{k,l} \leq r - \frac{1}{\tilde{h}_x^2}\left(c_{x_1} + c_{x_2} - \overline{\rho}_{1,2}\overline{\sigma}_1\overline{\sigma}_2\right).$$

As in the two cases above, we now obtain

$$\frac{1}{2}\left(c_{cc} + H^2 r + \overline{\rho}_{1,2}\overline{\sigma}_1\overline{\sigma}_2\right)h_\tau^i \leq \tilde{h}_x^2 \;\Rightarrow\; \frac{h_\tau^i}{2}a_{k,l,\gamma}^{k,l} \leq 1 \tag{4.16}$$

which in combination with (4.2) leads us to (4.3).
For the remaining six combinations of the three differencing methods a similar combination of the appraisals (4.6), (4.7), (4.10), (4.11), (4.14), *and* (4.15) *can be derived. The resulting constants are defined as given in the formulation of the lemma.*
For the boundary points of the grid we have $\mathcal{C}_{1,2} = \{0\}$. *A close look to* (4.6), (4.7), (4.10), (4.11), (4.14), *and* (4.15) *reveals that these appraisals also hold for all boundary points.*
Combining all these appraisals for each single case we conclude that (4.3) *needs to be fulfilled to guarantee* (4.2).

□

With Lemma 4.2 we can now prove the convergence of the CN scheme to the unique viscosity solution of the Black-Scholes-Barenblatt equation (2.4) under certain assumptions.

Theorem 4.3 (Convergence of the CN scheme)**.**
Let a grid $\mathcal{G}\left(L, \{t^i\}, \mathcal{S}, \Omega_2\right)$ *which satisfies Assumption 2.9 be given. The spatial step sizes are assumed to be chosen in such a way that only centered difference quotients are used. Furthermore, let the following conditions hold true*

1. $r \geq 0$
2. $\tilde{h}_x \leq 1$

3. the terms of Lemma 4.2 hold.

Then the CN scheme given by

$$D_\gamma^{c,i}\left(v^i, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1}\right) = B^i_{\gamma,\frac{1}{2}} v^i - C^{i-1}_{\gamma,\frac{1}{2}} v^{i-1} - r^i_{\frac{1}{2}} = 0$$

for a timestep $t^{i-1} \to t^i$ converges to the unique viscosity solution of the Black-Scholes-Barenblatt equation (2.4) for $\max_i\{h^i_\tau\}, H \to 0$.

Proof.

We will prove Theorem 4.3 in three steps by verifying monotonicity, stability, and consistency of the scheme. The result then follows with Assumption 3.7 by Theorem 1.4 of Barles and Souganidis [BS90].

To show monotonicity we directly verify the two conditions of Definiton 1.23.

Let $\varepsilon^i, \varepsilon^{i-1} \in \mathbb{R}_+^{M^i}$ be two vectors.

W. l. o. g. we consider an inner grid point with index (k,l). For ease of notation we denote by $\overline{a}_{k,l}$ the entries of the matrix A^{i-1}_γ, compare the definition of $C^{i-1}_{\gamma,\frac{1}{2}}$ in (4.1c).

Then the following inequality holds true:

$$\begin{aligned}
&D^{c,i}_{\bar{j},\gamma}\left(\left[v^i + \varepsilon^i\right]_{\bar{j}}, v^i_{\bar{j}}, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1} + \varepsilon^{i-1}\right) \\
&= \left(1 + \frac{h^i_\tau}{2} a^{k,l}_{k,l,\gamma}\right) v^i_{k,l} + \frac{h^i_\tau}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} a^{\mu,\nu}_{k,l,\gamma}(v^i_{\mu,\nu} + \varepsilon^i_{\mu,\nu}) \\
&\qquad + \frac{h^i_\tau}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} \overline{a}^{\mu,\nu}_{k,l,\gamma}(v^{i-1}_{\mu,\nu} + \varepsilon^{i-1}_{\mu,\nu}) - \left(1 - \frac{h^i_\tau}{2}\overline{a}^{k,l}_{k,l,\gamma}\right)(v^{i-1}_{k,l} + \varepsilon^{i-1}_{k,l}) \\
&\leq \left(1 + \frac{h^i_\tau}{2} a^{k,l}_{k,l,\gamma}\right) v^i_{k,l} + \frac{h^i_\tau}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} a^{\mu,\nu}_{k,l,\gamma} v^i_{\mu,\nu} \\
&\qquad + \frac{h^i_\tau}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} \overline{a}^{k,l}_{k,l,\gamma} v^{i-1}_{k,l} - \left(1 - \frac{h^i_\tau}{2}\overline{a}^{\mu,\nu}_{k,l,\gamma}\right) v^{i-1}_{\mu,\nu} \\
&= D^{c,i}_{\bar{j},\gamma}\left(\left[v^i\right]_{\bar{j}}, v^i_{\bar{j}}, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1}\right),
\end{aligned}$$

by Lemma 4.2 and since all off-diagonal entries $a^{\mu,\nu}_{k,l,\gamma}, (\mu,\nu) \in \mathcal{N}(k,l)$, are non-positive.

Let now $\varepsilon^i \in \mathbb{R}_+$. Since the diagonal entries of $B^i_{\gamma,\frac{1}{2}}$ are non-negative, we have

$$\begin{aligned}
&D^{c,i}_{\bar{j},\gamma}\left(\left[v^i\right]_{\bar{j}}, v^i_{\bar{j}} + \varepsilon^i, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1}\right) \\
&= \left(1 + \frac{h^i_\tau}{2} a^{k,l}_{k,l,\gamma}\right)\left(v^i_{k,l} + \varepsilon^i\right) + \frac{h^i_\tau}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} a^{\mu,\nu}_{k,l,\gamma} v^i_{\mu,\nu}
\end{aligned}$$

$$
\begin{aligned}
&\quad + \frac{h_\tau^i}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} \overline{a}_{k,l,\gamma}^{\mu,\nu} v_{\mu,\nu}^{i-1} - \left(1 - \frac{h_\tau^i}{2}\overline{a}_{k,l,\gamma}^{k,l}\right) v_{k,l}^{i-1} \\
&\geq \left(1 + \frac{h_\tau^i}{2} a_{k,l,\gamma}^{k,l}\right) v_{k,l}^i + \frac{h_\tau^i}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} a_{k,l,\gamma}^{\mu,\nu} v_{\mu,\nu}^i \\
&\quad + \frac{h_\tau^i}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} \overline{a}_{k,l,\gamma}^{\mu,\nu} v_{\mu,\nu}^{i-1} - \left(1 - \frac{h_\tau^i}{2}\overline{a}_{k,l,\gamma}^{k,l}\right) v_{k,l}^{i-1} \\
&= D_{\bar{j},\gamma}^{\mathrm{c},i}\left(\left[v^i\right]_{\bar{j}}, v_j, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1}\right),
\end{aligned}
$$

So, the monotonicity of the scheme is proven for an inner grid point. At the boundary points similar appraisals also hold true due to the definition of the boundary conditions given by $r_{\frac{1}{2}}^i$.
Thus, the Crank-Nicolson scheme is monotone if the conditions of the theorem are satisfied.

To prove the stability of the scheme we show that each iterate $v^i, i \in I_L$, *is bounded from above by a constant. This is true by construction for* v^0. *For the inner grid point with index* (k,l) *we now also assume w. l. o. g.*

$$
v_{k,l}^i = \max_{\mu\in J_{M_1},\nu\in J_{M_2}} \{v_{\mu,\nu}^i\}.
$$

For the Crank-Nicolson scheme at this grid point we then have

$$
\begin{aligned}
\left(1 + \frac{h_\tau^i}{2} a_{k,l,\gamma}^{k,l}\right)\|v^i\|_\infty &\leq \frac{h_\tau^i}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} -a_{k,l,\gamma}^{\mu,\nu}|v_{\mu,\nu}^i| + \left(1 - \frac{h_\tau^i}{2}\overline{a}_{k,l,\gamma}^{k,l}\right)|v_{k,l}^{i-1}| \\
&\quad + \frac{h_\tau^i}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} -\overline{a}_{k,l,\gamma}^{\mu,\nu}|v_{\mu,\nu}^{i-1}| \\
&\leq \frac{h_\tau^i}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} -a_{k,l,\gamma}^{\mu,\nu}\|v^i\|_\infty + \left(1 - \frac{h_\tau^i}{2}\overline{a}_{k,l,\gamma}^{k,l}\right)\|v^{i-1}\|_\infty \\
&\quad + \frac{h_\tau^i}{2} \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} -\overline{a}_{k,l,\gamma}^{\mu,\nu}\|v^{i-1}\|_\infty,
\end{aligned}
$$

by the non-positivity of the off-diagonal entries $a_{k,l,\gamma}^{\mu,\nu}, (\mu,\nu) \in \mathcal{N}(k,l)$, *and Lemma 4.2.*
If we now rearrange terms and exploit the definition of $a_{k,l,\gamma}^{k,l}$ *in (2.18), we obtain*

$$
\underbrace{\left(1 + \frac{h_\tau^i}{2} r\right)}_{\geq 1}\|v^i\|_\infty \leq \underbrace{\left(1 - \frac{h_\tau^i}{2} r\right)}_{\leq 1}\|v^{i-1}\|
$$

$$\Leftrightarrow \qquad \|v^i\|_\infty \leq \frac{1-\frac{h_\tau^i}{2}r}{1+\frac{h_\tau^i}{2}r}\|v^{i-1}\|_\infty \leq \|v^{i-1}\|_\infty,$$

where we used

$$a_{k,l,\gamma}^{k,l} + \sum_{(\mu,\nu)\in\mathcal{N}(k,l)} a_{k,l,\gamma}^{\mu,\nu} = r$$

which holds for all inner grid points. For the points on the boundary we can replace "=" with "≥" because some negative summands of the sum are missing.
Since $b_\gamma^i, i \in I_L$, and $\|v^0\|$ are bounded independent of h_τ^i and H we conclude by induction that all iterates $v^i, i \in I_L$ of the scheme are also bounded independently of the stepwidth in time and space.

The CN scheme is consistent, see for example [SB02].

□

Remark 4.2.
For the EE scheme, cp. Example 1.21, the right hand side would contain the matrix $I - h_\tau^i A_\gamma^i$. Thus, the conditions to guarantee stability and monotonicity will be even more restrictive as for the CN scheme. Meanwhile the method is only first order and we have no possibility to determine optimal controls due to the explicit character.

Convergence for discrete optimisation. Now we allow that at some grid points backward or forward difference quotients are used. In this situation we pass over to a relaxed, discrete optimisation problem in order to keep the overall effort easy to handle, cp. Section 3.2.2. To guarantee the convergence of the BDF1 and the CN scheme towards the viscosity solution we, have to make sure that the discretisation error of the discrete optimisation decreases for $cM_1, cM_2 \to \infty$. For all $u, w \in \{b, c, f\}$ we want

$$\left| \min_{\gamma_{uw}\in\Theta_{uw}\cap\mathcal{D}_2} z_{uw}^2(\gamma_{uw}) - \min_{\gamma_{uw}\in\Theta_{uw}} z_{uw}^2(\gamma_{uw}) \right| \to 0, \quad \text{for } cM_1, cM_2 \to \infty. \tag{4.17}$$

Condition (4.17) holds if the discrete derivatives are bounded and z_{uw}^2 is a continuous function of the controls for all $u, w \in \{b, c, f\}$, cp. [WF08]. Therefore, (4.17) holds for the problems which are considered in this thesis.

The following lemma states the convergence result for the discretisations with dis-

crete optimisation.

Lemma 4.4.
Let a consistent, stable, and monotone discretisation $D_\gamma^{c,i}(v^i, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1}) = 0$ with continuous optimisation be given. And let it converge to the unique viscosity solution of the BSB equation (BSB_n) for $n = 2$. Furthermore, let 4.17 hold.
Then the solution of the discretisation $D_\gamma^{d,i}(v^i, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1}) = 0$ converges to the unique viscosity solution of the BSB equation (BSB_n), $n = 2$, for $H, h_\tau \to 0$ and $cM_1, cM_2 \to \infty$.

Proof.
Stability and monotonicity of the scheme are not affected whether we use continuous or discrete optimisation. They were directly derived from the properties of the discretisation matrix $B^i_{\gamma,\theta}$. Their properties are indpendent of the controls at each single grid point as long as they are in Θ_2.
For the consistency we have

$$\begin{aligned}
&\left|v_\tau - \mathcal{L}^2_{BSB} v - D_\gamma^{d,i}(v^i, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1})\right| \\
&\leq \underbrace{\left|v_\tau - \mathcal{L}^2_{BSB} v - D_i^{c}(v^i, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1})\right|}_{\to 0} \\
&\quad + \underbrace{\left|D_\gamma^{c,i}(v^i, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1}) - D_\gamma^{d,i}(v^i, \tau^i, \mathcal{S}, \tau^{i-1}, \mathcal{S}, v^{i-1})\right|}_{\overset{(4.17)}{\to} 0} \to\ 0
\end{aligned}$$

Thus, also the solution of the discretisation with discrete controls converges to the unique viscosity solution for an increasing number of points in $\mathcal{D}_2$ and decreasing step sizes H and h_τ.

□

A similar result to the lemma above has been proven by Wang and Forsyth [WF08].

4.2 Convergence of one-dimensional schemes

The convergence results for the one-dimensional non-linear schemes can be proven analogue to the proofs of section 4.1. We will obtain similar conditions for the convergence of each scheme.
The pricing of options on one asset in the BSB model is a known problem in the literature. Pooley, Forsyth and Vetzal [PFV03a] apply a Finite Volume (FV) method to the BSB equation BSB_1. Heider [Hei10] and Schaeling [Sch10] consider convergence

of Finite Difference (FD) methods. In the latter two papers the pricing equation is transformed to log-prices and then additionally scaled by the underlying's value. In the publication mentioned first the equation is solved in the original variables. In all papers the non-linearity is handled by Newton iteration in each time step. Its convergence is guaranteed if the iteration matrix is an M-matrix [PFV03a]. In [Sch10] it is proven that for an appropriate choice of the stepwidth in space their iteration matrix is an M-matrix. This result also applies to the discretisation in [Hei10].
The results we present are of the same quality as those in [Hei10] and [Sch10]. Thus, we keep this section short. They can be understood as a special case of the results presented in Section 4.1. For the pricing of American options, see the results in Section 4.3.

For the determination of the optimal control at a time level i Algorithm 3.1 is used. Its convergence is guaranteed by Theorem 3.1 since we know from Lemma 2.10 that the iteration matrix is an M-matrix.
Thus, the remaining task is the proof of convergence of the BDF1 and the CN scheme to the unique viscosity solution of equation (2.3). As before, we assume that the spatial grid is constant over time. For the BDF1 scheme we have

Lemma 4.5.
Let a grid $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}\}, \Omega_1)$ be given and let $r \geq 0$. Furthermore, let the spatial step sizes be bounded so that only centered difference quotients are used.
Then the BDF1

$$D_\gamma^i \left(v^i, \tau^i, \mathcal{S}^i, \tau^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right) = B_{\gamma,1}^i v^i - C_{\gamma,1}^{i-1} v^{i-1} - r_1^i = 0$$

converges to the unique viscosity solution of the BSB equation (2.3) for $h_\tau, H \to 0$ and $cM_1 \to \infty$.

The proof is completely analogue to the one of Theorem 4.1.

For the CN scheme we need an additonal condition to assure its convergence.

Lemma 4.6.
Let a grid $\mathcal{G}\left(L, \{t^i\}, \{\mathcal{S}^i\}, \Omega_1\right)$ be given and let $r \geq 0$. Furthermore, let

$$\max\{c_f,\, c_b,\, c_c\} h_\tau \leq \tilde{h}_x^2, \tag{4.18}$$

where

$$\tilde{h}_x \leq h_1^k, k \in I_{M^i}, \qquad c_f = \frac{1}{2}\left(\overline{\sigma}_1^2\left(1+\frac{H}{2}\right)+H^2 r\right),$$
$$c_b = \frac{1}{2}\left(\overline{\sigma}_1^2 + r\left(H+H^2\right)\right), \qquad c_c = \frac{1}{2}\left(\overline{\sigma}_1^2\left(1+\frac{H}{2}\right)+\left(H+H^2\right)r\right).$$

Then the CN scheme

$$D_\gamma^i\left(v^i,\tau^i,\mathcal{S}^i,\tau^{i-1},\mathcal{S}^{i-1},v^{i-1}\right) = B^i_{\gamma,\frac{1}{2}}v^i - C^{i-1}_{\gamma,\frac{1}{2}}v^{i-1} - r^i_{\gamma,\frac{1}{2}} = 0$$

converges to the unique viscosity solution of the BSB equation (BSB_1) (2.3) *for* $h_\tau, h_1 \to 0$.

The proof can be performed analogously to the one of Lemma 4.2.

Proof.
We only give the appraisal of each diagonal entry of the matrix A^i_γ *defined in* (2.33). *The stability and the monotonicity of the Crank-Nicolson scheme can then be deduced by the same arguments used in the proof of Lemma 4.2.*
W. l. o. g. we consider a grid point with index $(\cdot,k)$. *By the construction of the discretisation we know that the minuends in the following inequalities are non-positive. If we use central differencing, we have*

$$\begin{aligned} a^k_{k,\gamma} &\leq r - \frac{1}{\tilde{h}_x^2}\left(2d^{k,l}_{1,\bar{j}} + c^{k,l}_{1,\bar{j}}\left(h_1^k - h_1^{k+1}\right)\right) \\ &\leq \frac{1}{\tilde{h}_x^2}\underbrace{\left(\overline{\sigma}_1^2\left(1+\frac{H}{2}\right)+H^2 r - H|r-\delta_1|\right)}_{=2c_c}. \end{aligned}$$

For backward differencing we obtain the appraisal:

$$a^k_{k,\gamma} \leq r - \frac{1}{\tilde{h}_x^2}\left(2d^{k,l}_{1,\bar{j}} + c_{x_1}\gamma h_1^k\right) \leq \frac{1}{\tilde{h}_x^2}\underbrace{\left(\overline{\sigma}_1^2 + H^2 r + H(r-\delta_1)^-\right)}_{=2c_b}$$

Finally, for the use of forward differencing we have

$$a^k_{k,\gamma} \leq r - \frac{1}{\tilde{h}_x^2}\left(2d^{k,l}_{1,\bar{j}} - c_{x_1}\gamma h_1^{k+1}\right) \leq \frac{1}{\tilde{h}_x^2}\underbrace{\left(\overline{\sigma}_1^2\left(1+\frac{H}{2}\right)+H^2 r - H(r-\delta_1)^+\right)}_{=2c_f}.$$

Defining $c := \max\{c_f, c_c, c_b\}$*, we have:*

$$ch_\tau^i \leq \tilde{h}_x^2 \Rightarrow \frac{h_\tau^i}{2\tilde{h}_x^2} c \leq 1 \Rightarrow \frac{1}{2} h_\tau^i a_k^k \leq 1$$

Now, if inequality (4.18) *is satisfied, we could proceed proving stability and monotonicity as in the proof of Theorem 4.3.*
If discrete optimisation has to be used, the discretisation error converges to zero for $cM_1 \to \infty$*, cp. Lemma 4.4 .*

□

4.3 Convergence of schemes for American option pricing

In Section 2.4 we learned that the value of an American option can be approximated by solving a penalized version of the pricing equation for European options instead of solving a complementary problem directly. For the linear BS model we saw that Algorithm 2.1 converges to the solution of the discretised pricing equation, compare Theorem 2.11. In this section we extend this result and those of Section 4.1 to the discrete schemes for Problem 2.29 for pricing American options. This problem is a valid formulation for the one-dimensional as well as the multi-dimensional pricing problems.
On the one hand a convergence result for the combination of the Penalty iteration (Algorithm 2.1) and the Policy iteration (Algorithm 3.1) has to be established. On the other hand we have to prove that the discrete schemes represented by (2.45) also converge to the solution of Problem 2.29. Of course, as mentioned in Remark 1.9, we only can assume that there exists a unique viscosity solution to this problem.

We assume that the spatial grid is constant over time, i. e. $\mathcal{S}^0 = \ldots = \mathcal{S}^L$. We recall that the θ-scheme for a time step $t^{i-1} \to t^i$ is given by

$$B_{\gamma,\theta}^i v^i = C_{\gamma,\theta}^i v^{i-1} + r_\theta^i + h_\tau^i P_\varepsilon\left(v^i\right)\left(p - v^i\right)^+ .$$

For the definitions of the matrices and vectors see Section 2.4.

We now proof a lemma which guarantees the convergence of scheme (2.45) for the penalized equation whenever the corresponding scheme (2.40) for the non-penalized equation is convergent.

Lemma 4.7.
Let a grid $\mathcal{G}(L, \{t^i\}, \{\mathcal{S}\}, \Omega_n)$ be given. Assume that a time step $t^{i-1} \to t^i$ of the scheme for solving the pricing equation for European options on n assets

$$D^i_\gamma \left(v^i; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right) = B^i_{\gamma,\theta} v^i - C^i_{\gamma,\theta} v^{i-1} - r^i_\theta$$

is monotone, stable, and consistent.
Then, the corresponding scheme for solving the penalized pricing equation (2.45)

$$\begin{aligned} &D^{i,pen}_\gamma \left(v^i; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right) \\ &\qquad = B^i_{\gamma,\theta} v^i - C^i_{\gamma,\theta} v^{i-1} - r^i_\theta - h^i_\tau P_\varepsilon \left(v^i\right) \left(p - v^i\right)^+ \end{aligned} \tag{4.19}$$

also is monotone, consistent and stable. Thus, it converges to the unique viscosity solution of the penalized equation.

Proof.
We first verify the monotonicty of (4.19) *by verifying Definition 1.23. Let $\varepsilon^i, \varepsilon^{i-1} \in \mathbb{R}^{M^i}_+$. Without loss of generality we consider the grid point with index (i,k,l) with lexicographical index $\overline{j}$ for spatial grid points. We have*

$$\begin{aligned} &D^{i,pen}_{\overline{j},\gamma} \left(\left[v^i + \varepsilon^i\right]_{\overline{j}}, v^i_{\overline{j}}; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1} + \varepsilon^{i-1}\right) \\ &\leq D^i_{\overline{j},\gamma} \left(\left[v^i\right]_{\overline{j}}, v^i_{\overline{j}}; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right) - h^i_\tau P_\varepsilon \left(v^i + \left[\varepsilon^i\right]_{\overline{j}}\right)_{\overline{j}} \left(p - v^i - \left[\varepsilon^i\right]_{\overline{j}}\right)^+ \\ &= D^{i,pen}_{\overline{j},\gamma} \left(\left[v^i\right]_{\overline{j}}, v^i_{\overline{j}}; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right), \end{aligned}$$

since $P_\varepsilon(v)$ is a diagonal matrix and the disturbance ε^i does not effect the diagonal element with index $\overline{j}$.
Let $\xi^i \in \mathbb{R}_+$. Then the following inequality holds true

$$\begin{aligned} &D^{i,pen}_{\overline{j},\gamma} \left(\left[v^i\right]_{\overline{j}}, v^i_j + \xi^i; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right) \\ &\geq D^i_{\overline{j},\gamma} \left(\left[v^i\right]_{\overline{j}}, v^i_j; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right) - h^i_\tau P_\varepsilon \left(v^i + e_{\overline{j}} \xi^i\right)_{\overline{j}} \left(p - v^i - e_{\overline{j}} \xi^i\right)^+ \\ &\geq D^{i,pen}_{\overline{j},\gamma} \left(\left[v^i\right]_{\overline{j}}, v^i_{\overline{j}}; t^i, \mathcal{S}^i, t^{i-1}, \mathcal{S}^{i-1}, v^{i-1}\right), \end{aligned}$$

where $e_{\overline{j}}$ is the $\overline{j}$-th canonical unit vector in $\mathbb{R}^M$.
Thus, also the penalized discretisation is monotone.
The stability of the penalized scheme can be shown analogously to the one of the non-penalized one, since the penalty term is bounded from above.
We already know that the non-penalized discretisation is consistent. By adding the

penalty term the consistency of the scheme for the penalized equation is not affected. The scheme remains consistent.
The determination of the optimal controls is for both continuous and discrete optimisation the same as for European options. All results regarding convergence also hold. □

To use the penalized discretisation for American option pricing in the BSB model we need an algorithm to solve the discrete non-linear equation (2.45). The following one realizes the Penalty iteration and the Policy iteration at the same time. It is a combination of Algorithm 3.1 and Algorithm 2.1. That is, we execute the Newton iteration for the penalty method with an iteration matrix which is updated in every step by the optimal control.
Let $p \in \mathbb{R}^{M^i}$ denote a vector in lexicographical ordering containing the payoff on the grid in space.

Algorithm 4.1: Combined Penalty and Policy iteration.

Input: $\theta, h_\tau, v^{i-1}, \gamma^{i-1}, B^i_{\gamma,\theta}, C^{i-1}_{\gamma,\theta}, r^{i-1}_\theta, p$
Output: $v^{i,(j)}, B^{i,(j)}_{\gamma,\theta}, r^{i,(j)}_\theta$

pen_pol_iteration$\left(\theta, h_\tau, v^{i-1}, \gamma^{i-1}, B^{i-1}_{\gamma,\theta}, C^{i-1}_{\gamma,\theta}, r^i_\theta, p\right)$
  $j \leftarrow 0, \quad err \leftarrow 1;$
  $v^{i,(0)} \leftarrow v^{i-1}, \quad \gamma^{i,(0)} \leftarrow \gamma^{i-1};$
  $B^{i,(0)}_{\gamma,\theta} \leftarrow B^{i-1}_{\gamma,\theta}, \quad r^{i,(0)}_{\gamma,\theta} \leftarrow r^{i-1}_\theta;$
  while $err \geq 10^{-6}$ **do**
    for $(x_{1,k}, x_{2,l}), k \in \mathcal{J}_{M_1}, l \in \mathcal{J}_{M_2}$ **do**
      calculate a_1, a_2, a_3 *as derivatives of* $v^{i,(j)}$; //s. (3.6)
      $\gamma^{i,(j)}_{k,l} = det\_control(\cdot);$
    update $B^{i,(j+1)}_{\gamma,\theta} \curvearrowleft B^{i,(j)}_{\gamma,\theta};$
    calculate $P^{i,(j)}_\varepsilon \leftarrow P_\varepsilon\left(v^{i,(j)}\right);$
    solve $\left(B^{i,(j+1)}_{\gamma,\theta} + h_\tau P^{i,(j)}_\varepsilon\right) v^{i,(j+1)} = C^{i-1}_{\gamma,\theta} v^{i-1} + r^{i,(j+1)}_\theta + h_\tau P^{i,(j)}_\varepsilon p;$
    $err \leftarrow \|v^{i,(j+1)} - v^{i,(j)}\|_2;$
    $j \leftarrow j + 1;$
  return $v^{i,(j)}, B^{i,(j)}_{\gamma,\theta}, r^{i,(j)}_\theta;$

The following lemma states a convergence result for Algorithm 4.1.

Lemma 4.8.
Let a start value v^{i-1} be given. Furthermore, let the matrices $B^i_{\gamma,\theta}$ and $C^{i-1}_{\gamma,\theta}$ in equation (4.19) *be bounded in the maximum norm.*

If $B^i_{\gamma,\theta}$ is an M-matrix for all $\gamma \in \Theta^d_\cdot$ and $\theta \in \{0.5, 1\}$, then Algorithm 4.1 converges monotonously to a unique solution v^i.

The arguments in the proof are analogue to those used to prove Theorem 3.1. With the iteration matrix being an M-matrix and the boundedness of the iterates the mentioned results can be deduced directly.

Proof.
We consider one step $j \to j+1$ of the iteration.
Since $B^i_{\gamma,\theta}$ is bounded by assumption this is also true vor $B^i_{\gamma,\theta} + h^i_\tau P_\varepsilon(v^{i,(j)})$ the boundedness of the iterates can be affiliated from equation (4.19) *as in the proof of Theorem 3.1.*
The monotone convergence and the uniqueness of the solution in the proof of Theorem 3.1 were an instantaneous result of the M-matrix property of the iteration matrix. The matrix $B^i_{\gamma,\theta} + h^i_\tau P_\varepsilon(v^{i,(j)})$ maintains this property. We only add a non-negative figure to every diagonal entry and thus the off-diagonal entries remain non-positive. Applying the triangle inequality and Lemma 2.9 the M-Matrix property of $B^i_{\gamma,\theta} + h^i_\tau P_\varepsilon(v^{i,(j)})$ is proven. Therefore the monotone convergence and the uniqueness of the solution of the above algorithm follow directly.

□

5 Numerical Results

In this chapter an analysis of the numerical results for the methods introduced in the Chapters 2 and 3 is given. We will investigate accuracy, convergence, and runtime. The analysis is presented for options on one and on two assets. For both we consider European as well as American options for several sets of parameters. Different schemes with different approaches to the optimization problem included are compared among each other. In addition we analyse the behaviour of the methods if the theoretical conditions for convergence, cp. Chapter 4, are not met.
For one two-dimensional option we show the evolution of the controls over time in order to clarify the nature of the optimal control problem for options on financial markets.
In the first section of this chapter we briefly address certain aspects of the implementation. Hereafter, we focus on methods for options on two assets and conclude with those for single-asset options.

5.1 Aspects of implementation

Before we present the results for the methods introduced in the previous chapters let us make some preliminary remarks.
All methods which will be used are implemented in `C++`. They were executed on a MacBook Air with two 1.7 GHz processors and an internal memory of 4GB. The timekeeping was done via the "clock()" - function of the "time.h" package provided by `C++`.
For the solution of the systems of equations with more than six diagonals the package "lsolver" is applied. It is provided by C. Badura [Bad98], v. i.. Here several iterative solvers are implemented. The implementation is based on routines of the `LAPACK` [LAP] and `Blas` [BLA] packages.
For those systems which are in tridiagonal form we use the Thomas algorithm, cf. [SB02].

To keep the algorithm efficient the discretisation matrix is initialised as a sparse matrix with seven or three diagonals, respectively. Each diagonal is stored in a

separate array, where an element of the matrix can be accessed by the lexicographical index of its row and column index. Furthermore, all calculations regarding the choice of difference quotients and the dependent approach to the optimization method are done a priori.
The algorithm could easily be made more efficient by including parallelisation. There are two main aspects to apply it. For every spatial grid point the optimization has to be done independently of all others. This is an obvious starting point to benefit of the use of several cores. Additionally, the update of the discretisation matrix for each step of the policy iteration also is done for each grid point separately. Applying parallelisation in this context could be as beneficial as for the optimization.

5.2 Detailed numerical results

In this section we present a numerical analysis of the different approaches which were introduced in the previous chapters for solving Problem 2.3 and Problem 2.29 for $n = 2$. Especially Sections 2.2.1, 2.3, 2.4, and 4.1 and Chapter 3 serve as basis for the numerical experiments. For $n = 3$ we will not perform numerical experiments. For the one-dimensional pricing problems a detailed analysis can be found in the publications of Heider [Hei10] and Schaeling [Sch10].

For every time step of the considered schemes we have to solve a system of equations several times. This system is of sparse structure and can be solved efficiently by iterative methods. Direct methods could cause a fill-in in the matrix $B^i_{\gamma,\theta}$ of the linear system and might therefore require much memory and become slow. Since it is not symmetric or positive definite a quite general iterative solver is chosen. In the following we will use the Biconjugate Gradient Stabilized (BiCGSTAB) method developed by van der Vorst [vdV92]. It is an extension of the Conjugate Gradient method for non-symmetric matrices. A detailed introduction to this method among others can be found in the book of Saad [Saa03]. Since the BiCGSTAB works iteratively an error tolerance has to be provided for the algorithm to quit. We set $\varepsilon_{IS} := 10^{-12}$. As initial value we use the option value of the previous time level. This choice should reduce the overall iterations perceptibly.

We start with the three schemes (BDF1, BDF2, CN) constructed on uniform grids where continuous optimisation is possible. A reasonable criterion to validate their performance is the calculation of the convergence rates. Moreover, we look at the number of iterations needed by the policy iteration and the BiCGSTAB, respectively. Hereafter, we compare the schemes constructed on a uniform grid using continuous optimisation with those constructed on a uniform grid with discrete optimisation

and those constructed on a non-uniform grid with discrete optimisation. The focus of this comparison lies on run time and accuracy of the solution.

For the numerical experiments we choose $\Omega_2^D = [0,T] \times [-3,4] \times [-3,4]$. For the BDF2 and the CN scheme Rannacher (cp. [Ran84]) time stepping was used for the first five time steps. By Δ_C^S, Δ_C^t, I_{PI}, and I_{IS} we denote the convergence rates, the number of iterations executed by the policy iteration and those of the iterative solver, respectively. In Table 5.1 the parameters for the first analysis are given.

Tab. 5.1: *Parameter set #1 for options on two assets.*

K	$S_{1,0}$	$S_{2,0}$	r	T	δ_1	δ_2	$\mathcal{V}_1$	$\mathcal{V}_2$	$\mathcal{C}_{1,2}$
1	1	1	0.05	1	0.01	0.01	[0.3,0.4]	[0.2,0.35]	[0.2,0.3]

Let us furthermore assume that the spatial step sizes of the uniform grid are small enough so that only centered difference quotients can be used, cp. Lemma 2.3. The first example is a put on the minimum of two assets with the parameters given in Table 5.1 which is at the money.

Tab. 5.2: *Results for an increasing number of spatial grid points of the BDF1 (top), BDF2 (middle), and CN (bottom) scheme for a European minimum put with parameters given in Table 5.1, and $N = 500$ time steps.*

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.196481	-	2.0	1.74	0.7
100	0.198950	-	2.06	2.05	3.8
200	0.199551	4.11	2.76	2.49	23.3
400	0.199700	4.02	3.27	4.15	144.7
800	0.199737	4.01	3.96	9.02	1182.8

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.196547	-	2.0	1.69	0.7
100	0.199013	-	2.01	1.95	3.7
200	0.199613	4.11	2.57	1.95	20.7
400	0.199763	4.02	3.02	3.74	129.3
800	0.199800	4.01	3.86	7.27	1005.0

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.196547	-	2.0	2.0	1
100	0.199013	-	2.02	2.10	5.1
200	0.199613	4.11	2.73	2.81	30.6
400	0.199763	4.02	3.18	4.64	174.4
800	0.199800	4.01	3.89	9.65	1311.4

We consider the convergence rates in time and in space separately. For the results in Table 5.2 the number of grid points in the two spatial directions was doubled

for every run. The number of iterations increases with the doubling of M_1 and M_2. Which is reasonable since with the number of grid points also the influence of the controls grows. Thus the difference between the iterates of the policy iteration decays more slowly. Of course for a minimum put we would not expect a strong variation of the controls neither in time nor in space due to the structure of the option value.
For all three schemes the policy iteration almost needs the same number of iterations. In each run the number of iterations is nearly constant.
A conspicuous conduct is the increasing time for each run. We would expect the run time to quadruple with each doubling of M_1 and M_2. However, due to the increasing average iterations I_{PI} and I_{IS} it grows with a factor of five to seven stronger than expected.

Tab. 5.3: *Results for an increasing number of grid points in time of the BDF1 (top), BDF2 (middle), and CN (bottom) scheme for a European minimum put with parameters given in Table 5.1, $M_1 = M_2 = 400$.*

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.198981	-	3.54	10.43	4.0
100	0.199297	-	3.23	6.61	10.0
200	0.199455	2.0	3.03	4.21	20.1
400	0.199535	2.0	2.82	3.08	37.3
800	0.199574	2.0	2.6	2.13	65.4

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.199557	-	3.24	8.96	3.5
100	0.199600	-	3.11	5.33	9.0
200	0.199610	4.0	2.87	3.33	18.0
400	0.199613	3.99	2.7	2.48	33.5
800	0.199614	4.02	2.31	2.05	58.3

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.199554	-	3.52	9.89	4.3
100	0.199599	-	3.31	6.45	11.5
200	0.199610	3.97	3.0	4.24	23.3
400	0.199613	3.98	2.8	3.1	43.4
800	0.199614	3.99	2.43	2.66	76.9

In Table 5.3 the convergence results are given for a doubled number of time steps for each run. The convergence rates are nearly identical to the theoretical ones. We obtain a rate of two for the BDF1 scheme and of four for the BDF2 and the CN scheme. For all three the average number of policy iterations is similar and decreases slowly with an increasing number of time steps. For the BiCGSTAB solver

the average effort decreases while the step size h_τ decreases. This behaviour can be explained regarding the condition of the discretisation matrix. For $h_\tau \to 0$ its structure becomes more and more similar to the identity matrix and so the condition reduces. Thus, the convergence speed increases.

For an American minimum put the results are of similar quality as those presented above. They can be found in Appendix B.1 in Tables B.1 and B.2. The convergence rates are close to four and two for doubling the number of spatial grid points and those in time, respectively. The average number of policy iterations is slightly increased. This circumstance is explained by the additional penalty term included in the iteration. The convergence speed of BiCGSTAB iteration instead is noticeably smaller. With the penalty term added to the diagonal of the discretisation matrix its condition deteriorates and thus the the error decays more slowly.

The convergence rates given in Tables 5.2 and 5.3 above were quite close to the theoretical ones. In the numerical examples presented hereafter we will see that this conformity is not what could be expected in general.

The next example is a capped minimum put, cp. Figure 1.3. The parameters of the underlying model are given in Table 5.4. Also in this case all conditions are satisfied so that only centered difference quotients are used and the optimisation problems are solved analytically.

Tab. 5.4: *Parameter set #2 for options on two assets.*

K	$S_{1,0}$	$S_{2,0}$	r	T	δ_1	δ_2	$\mathcal{V}_1$	$\mathcal{V}_2$	$\mathcal{C}_{1,2}$
1	1	1	0.05	1	0.0	0.0	[0.4,0.5]	[0.2,0.35]	[0.25,0.25]

For the parameter set #2 the condition of Lemma 2.3 is satisfied and only centered difference quotients are used.
For this type of option the Deltas as well as the Gammas change their sign for change values of x_1 and x_2. In Table 5.5 the data for the analysis of the convergence behaviour in space are given. As for the European minimum put all three schemes converge to the same option value despite the fact that the BDF2 scheme is not monotone at all and the CN scheme does not possess this property for the last run. The average number of iterations of the policy iteration is nearly constant. For each run the first few time steps require some additional iterations. But for the majority only two iterations are needed for convergence. The computational costs for the iterative solver are of the same quality as in the previous example. With a discretisation getting finer also the average effort increases.

Tab. 5.5: *Results for an increasing number of spatial grid points of the BDF1 (top), BDF2 (middle), and CN (bottom) scheme for a European capped minimum put with parameters given in Table 5.1, $N = 500$.*

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.126311	-	2.02	1.73	0.6
100	0.128963	-	2.02	2.02	3.3
200	0.129200	11.81	2.06	2.21	14.9
400	0.129330	1.81	2.12	3.5	82.9
800	0.129350	6.42	2.21	7.57	702.4

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.126343	-	2.01	1.68	0.6
100	0.128988	-	2.02	1.93	3.2
200	0.129225	11.19	2.05	2.11	14.64
400	0.129355	1.81	2.09	3.02	73.15
800	0.129376	6.40	2.17	6.46	625.15

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.126343	-	2.01	2.0	0.8
100	0.128988	-	2.02	2.13	4.6
200	0.129225	11.19	2.05	2.38	21.0
400	0.129355	1.81	2.08	3.81	102.5
800	0.129376	6.4	2.15	8.16	709.9

The convergence rates are of special interest. For all three schemes these are essentially the same. Still, for all three schemes the option value seems to converge monotonically to the same value but not with the theoretical convergence rates. With the smallest ratio of 1.81 we have at least first order convergence. The other two ratios in opposite clearly indicate higher convergence speed as could theoretically be expected. In Table 5.6 the results with doubled N are displayed. For the BDF1 scheme we observe the expected convergence rates of two. For the BDF2 and the CN scheme convergence is slower. They still converge super linearly but with rates of approximately 3.4 we do not have second order convergence. For the average iterations used by the policy iteration and the BiCGSTAB we observe the same behaviour as in Table 5.3. The decreasing step size h_τ accelerates the convergence of the iterations.

Tab. 5.6: *Results for an increasing number of grid points in time of the BDF1 (top), BDF2 (middle), and CN (bottom) scheme for a capped minimum put with parameters given in Table 5.1,* $M_1 = M_2 = 200$.

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.128972	-	2.68	10.22	2.9
100	0.129098	-	2.35	5.89	6.8
200	0.129162	2.0	2.2	3.42	12.8
400	0.129193	2.0	2.1	2.34	23.0
800	0.129209	1.99	2.03	1.84	41.3

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.129201	-	2.66	8.40	2.6
100	0.129218	-	2.36	4.94	6.3
200	0.129223	3.44	2.19	2.98	11.9
400	0.129224	3.2	2.08	2.22	21.8
800	0.129225	3.42	2.02	1.77	39.8

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.129198	-	2.64	9.60	3.2
100	0.129217	-	2.34	5.83	8.0
200	0.129222	3.43	2.65	3.53	15.9
400	0.129224	3.25	2.07	2.57	29.7
800	0.129225	3.4	2.02	2.15	56.0

The last example for schemes working on a completely uniform grid with continuous optimisation is the butterfly spread, cp. Figure 1.4. The data for the model in which we priced the options are given in Table 5.7.

Tab. 5.7: *Parameter set #3 for options on two assets.*

K	α	$S_{1,0}$	$S_{2,0}$	r	T	δ_1	δ_2	$\mathcal{V}_1$	$\mathcal{V}_2$	$\mathcal{C}_{1,2}$
1	0.2	1	1	0.05	1	0.02	0.01	[0.2,0.4]	[0.35,0.5]	[0.1,0.25]

Given the structure of the payoff there should be one change of the sign both for the first and second derivative. Thus, the controls will vary across the spatial domain. In Figure 5.1 the evolution of σ_1 is displayed for different time levels and different spatial discretisations.

In each graph the interesting area for $S_1, S_2 \in [0, 2.5]$ is plotted. Those on the left hand side contain 141 points in each direction and those on the right hand side 562. The value of σ_1, which is plotted at a grid point, is the one for which the objective function attained its minimal respectively maximal value.

At first, we notice that the results for σ_2 show a similar structure. For the correlation the results directly depend on the sign of the mixed derivative, v. s.

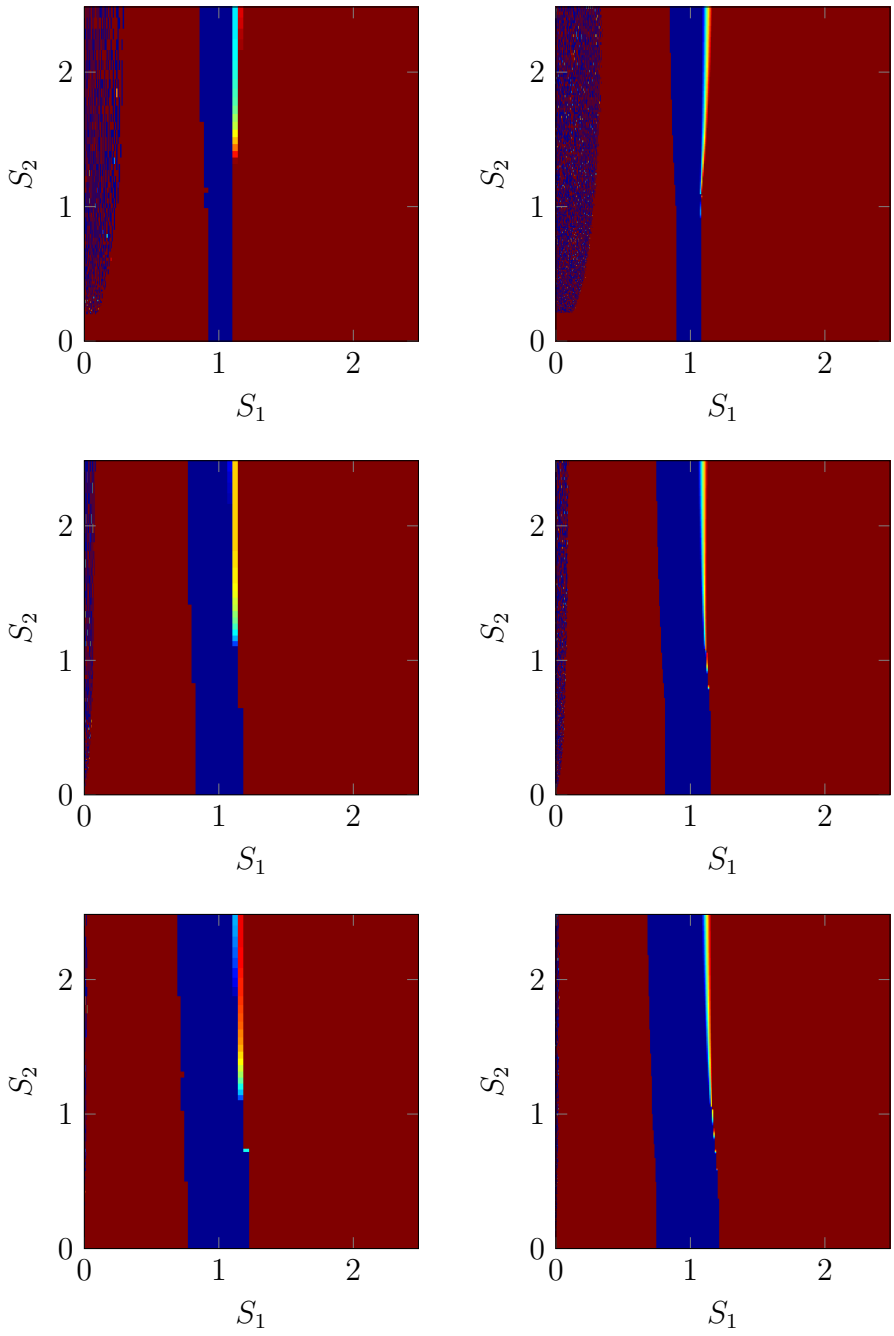

Fig. 5.1: *Evolution of σ_1 over time for a butterfly spread for two different spatial discretisations. The data are for $M_1 = M_2 = 200$ on the left hand side and for $M_1 = M_2 = 800$ on the right hand side. The time levels $\tau = 0.1$ (top), $\tau = 0.5$ (middle), and $\tau = 1$ (bottom) are displayed. The colours display the volatility starting with $\underline{\sigma_1}$ (red) and ending with $\overline{\sigma_1}$ (blue).*

The next observation is that mostly the minimal and maximal values of σ_1, respectively, occur. This fact will be of special interest, v. i.

For all three time levels σ_1 shows the same structure for $M_1 = M_2 = 200$ and

$M_1 = M_2 = 800$. With the spatial discretisation getting finer the different domains of the controls refine and the structure does not change. Thus, the evolution of the control over time seems to be "consistent" and independent of the spatial discretisation.
Similar results have been observed for other types of options. The results are displayed for the butterfly spread since the controls are most volatile for this option.
In the Tables 5.8 and 5.9 the convergence results for the butterfly spread are summarised. The convergence rates in Table 5.8 vary around four for all three schemes.

Tab. 5.8: *Results for an increasing number of spatial grid points of the BDF1 (top), BDF2 (middle), and CN (bottom) scheme for a European butterfly spread with parameters given in Table 5.7, $N = 500$.*

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.086776	-	2.09	1.8	0.6
100	0.078525	-	2.13	2.01	3.7
200	0.076995	5.39	2.31	2.51	19.1
400	0.076652	4.47	2.52	4.82	112.9
800	0.076551	3.39	2.66	10.63	935.3

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.086756	-	2.06	1.74	0.7
100	0.078518	-	2.10	1.91	3.6
200	0.076994	5.41	2.25	2.31	18.1
400	0.076654	4.47	2.44	4.19	102.7
800	0.076553	3.38	2.59	8.89	768.9

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.086756	-	2.04	2.0	0.8
100	0.078518	-	2.09	2.19	4.9
200	0.076994	5.41	2.25	3.08	24.6
400	0.076653	4.47	2.4	5.35	133.2
800	0.076552	3.38	2.59	11.04	967.1

But they are not as uniform as they can be observed for Finite Differences applied to multi-dimensional Black-Scholes models. Which would suggest the assumption that the varying controls influence the convergence of the schemes.
For the convergence in time smaller convergence rates can be observed for a smaller number of time steps. But with increasing N convergence improves, cp. Table 5.9.

In the next segment we price the butterfly spread again with the parameters as given in Table 5.4 and 5.7, respectively. But this time we use the discrete control set $\mathcal{D}_2$ instead of Θ_2. Consequently, we do not use the KKT conditions but compare the objective values for each discrete control. To analyse the influence of how fine

Tab. 5.9: *Results for an increasing number of grid points in time of the BDF1 (top), BDF2 (middle), and CN (bottom) scheme for a European butterfly spread with parameters given in Table 5.7, $N = 400$.*

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.076617	-	3.30	13.28	4.18
100	0.076612	-	3.09	7.56	9.94
200	0.076634	0.07	2.88	4.45	18.61
400	0.076648	1.13	2.44	2.91	31.59
800	0.076658	1.58	2.15	2.01	52.69

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.076461	-	3.22	11.02	3.72
100	0.076560	-	3.07	6.35	9.15
200	0.076619	1.69	2.79	3.9	17.31
400	0.076648	2.0	2.35	2.60	29.71
800	0.076661	2.29	2.13	1.91	50.37

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.076457	-	3.18	12.31	4.31
100	0.076557	-	3.06	7.47	10.96
200	0.076617	1.66	2.75	4.87	21.46
400	0.076647	1.99	2.31	3.36	37.94
800	0.076661	2.29	2.11	2.43	66.66

Θ_2 is discretised we vary $cM_1, cM_2 \in \{3, 5, 7, 9, 11\}$. We omit the results for the other options considered so far, since we have the biggest variations of control for the butterfly spread. Therefore, the results for those do not offer any new findings. In our experiments we observed that noticeable differences only occurred in a region where the controls changed significantly across the spatial grid. One example is displayed in Table 5.10. Here the butterfly spread was priced for $S_{1,0} = 1.15$, $S_{2,0} = 1.1$ which is a point where the volatilities vary strongly, cp. Figure 5.1. The difference between the option values calculated with continuous and discrete optimisation are of order 10^{-6} for $cM_1 = cM_2 = 3$. With a growing number of discrete controls the error reduces. For $cM_1 = cM_2 = 11$ the error is of order 10^{-8}. The constants cM_1 and cM_2 have to be chosen carefully. To solve the KKT-conditions for a two-dimensional problem we need at most twelve evaluations of the objective function and some additional calculations. If the discretised optimisation problem is solved $(cM_1+1)(cM_2+1)$ evaluations of z_{uw}^2 have to be done for a single grid point. For the example given in the table below the algorithm needed 91.9 sec., 110.8 sec. and 141.1 sec. for $cM_1 = cM_2 \in \{3, 7, 11\}$, respectively. With continuous optimisation the algorithm took 91.5 sec.

Before we analyse the utility of non-uniform grid we note the following. The exper-

Tab. 5.10: *Comparison of option values for continuous and discrete optimisation with $cM_1, cM_2 \in \{3, 5, 7, 9, 11\}$ for a European butterfly spread with parameters given in Table 5.7 and $S_{1,0} = 1.15$, $S_{2,0} = 1.1$. ($M_1 = M_2 = 400$, $N = 500$).*

	cont. controls	discrete controls	
		$cM. = 3$	$cM. = 5$
BDF1	0.060419707	0.604188735	0.060419416
BDF2	0.060437554	0.060436725	0.060437264
CN	0.060437080	0.06436251	0.060436790
	discrete controls		
	$cM. = 7$	$cM. = 9$	$cM. = 11$
BDF1	0.060419560	0.060419618	0.060419646
BDF2	0.060437406	0.060437465	0.060437493
CN	0.060436932	0.060436991	0.060437019

iments shown above are representative for the convergence behaviour we observed during our test runs. The schemes converge monotonically to the viscosity solution except for cases in which $(S_{1,0}, S_{2,0})$ lies in regions with volatile controls. But generally, the schemes cannot be expected to converge with the rates implied by theory. For linear and almost linear problems these rates are roughly attained. For highly non-linear problems this does not hold true, v. s..

In the last segment of this section we examine the efficiency increase which can be gained by non-uniform spatial grids. From Section 2.2.2 we know that for most non-uniform grids the discrete optimisation problem has to be solved at each grid point. Therefore, for the solution of the pricing problem on a non-uniform grid with discretised controls we have several limitations:

1. the scheme has to stay monotone and thus the step sizes are bounded as described in Section 2.2.1,

2. the number of discrete controls can significantly increase the overall run time, and

3. the number of discrete controls influence the accuracy of the final result.

Of course, the third point is the least restrictive one.

In the following example we apply a non-uniform grid as described in Section 2.2.2. The upper bounds for the spatial step sizes are directly implied by the bounds for volatility and correlation. As parameters for the non-uniform grid we used $x^+ = 0.5$ and $d = 5$. Thus, the points $K - \alpha, K, K + \alpha$ are all placed in a region where the minimal step size is applied.

In the following deliberations we consider two of the examples from above, namely the minimum put and the butterfly spread with the parameters given in Tables 5.1 and 5.7, respectively. We do not analyse the convergence behaviour of the schemes for increasing N, since it is similar to the results presented above.

In Table 5.11 the option values and the run time for the three different schemes and with a different number of discrete controls are given.

Tab. 5.11: *Option values for a European butterfly spread priced on a non-uniform grid with discrete control, parameters given in Table 5.1, and $N = 500$.*

	M_1,M_2	$cM. = 5$	$cM. = 9$
		$v(0,\cdot)$	$v(0,\cdot)$
BDF1	29/17	0.09485323	0.09485324
BDF2		0.09482174	0.09482174
CN		0.09482168	0.09482168
BDF1	58/33	0.08011376	0.08011395
BDF2		0.08010682	0.08010701
CN		0.08010643	0.08010662
BDF1	116/66	0.07765864	0.07765875
BDF2		0.07765869	0.07765879
CN		0.07765799	0.07765809
BDF1	230/134	0.07677106	0.07677117
BDF2		0.07677217	0.07677229
CN		0.07677152	0.07677164

In the table above we see again the influence of the discrete controls as described above. All three schemes converge to the same solution independent of the non-uniform grid and the choice of cM_1, cM_2. In the following Figure 5.2 run time and error are plotted against each other. For this purpose we used the option value calculated on a uniform grid with $M_1 = M_2 = 3200$ grid points and continuous optimisation as benchmark. We focus on the BDF1 and the CN scheme. Additionally to the results for the nonuniform grids, the butterfly spread was priced on grids with $(M_1, M_2) \in \{(20,20),(40,40), \ldots\}$ and $(M_1, M_2) \in \{(30,30),(60,60), \ldots\}$ grid points, respectively, and continuous optimisation. For the results obtained with discrete optimisation we applied $cM_1 = cM_2 = 5$.

In Figure 5.2 we see that the schemes solved on non-uniform grid and with discretized controls are slightly more efficient than those schemes executed on a uniform grid with continuous controls. But the efficiency increase is quite small which is explained by bounded maximal step sizes due to the volatility and correlation ranges. For the experiment with the lowest number of spatial grid points $h^1_{max} = 0.4$ was used. This example illustrates the limited success adaptive Finite Difference methods might

have for option pricing in the Uncertain Volatility model. Of course, for special cases, for example very small or zero correlation, the possible range for the spatial step sizes is larger.

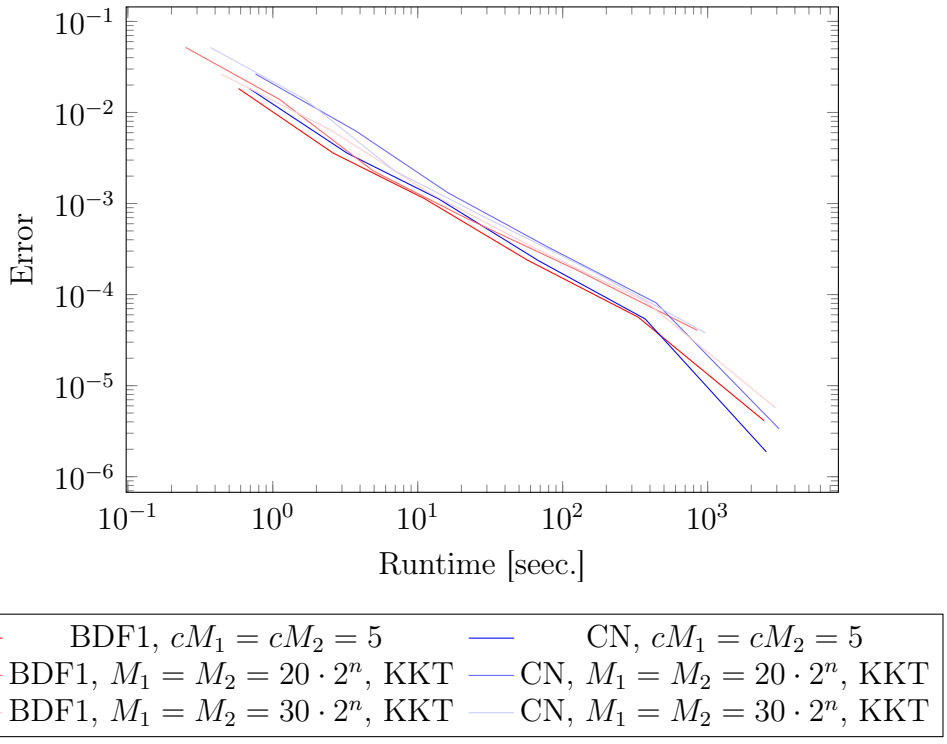

Fig. 5.2: *Plot of error versus runtime for option values calculated on uniform and non-uniform spatial grids. The BDF1 scheme on a uniform grid with continuous controls marked light red was calculated with $M_1 = M_2 = 2^n \cdot 30$ spatial grid points. While the other BDF1 scheme calculated on a uniform grid and continuous controls was calculated with $M_1 = M_2 = 2^n \cdot 20, n = 0, \ldots, 3$ spatial grid points. The same holds for the uc-CN schemes. For the schemes calculated on a non-uniform grid with discrete controls we applied $N_3 - N_2 = 2^n \cdot 5$ for each spatial dimension and $cM_1 = cM_2 = 5$.*

To conclude this chapter, we give a short comparison of the considered methods. We note that the BDF1 scheme is convergent for monotone discretisations and solves the pricing problem faster than the Crank-Nicolson method. But it theoretically and practically has the smaller rate of convergence. Assuring the convergence of the Crank Nicolson scheme is more restrictive than for the BDF1 scheme due to the step size condition. The rates of convergence of the CN scheme are higher than those of the BDF1 scheme and comparable to those of the BDF2 scheme. The BDF2 scheme is approximately as fast as the BDF1 scheme, but it converges with the rates of the CN scheme. The scheme however is not monotone and its convergence to the viscosity solution could not be proven.

For all three schemes non-uniform grids could be applied. Here the local optimisation problem should be approximated by considering discrete controls. The results obtained are slightly improved compared to their counterparts on uniform grids.

6 Conclusion

This chapter summarizes the analytical and practical results, which were elaborated in the present thesis. The theoretical framework regarding the Uncertain Volatility model and the concept of viscosity solution in which the introduced schemes are included are given in the first chapter.
In the second chapter we performed the Finite Difference discretisation of the Black-Scholes-Barenblatt equation for European options. For American options it was assumed that the option value is given by a complementary problem which was then approached by the well known penalty mehtod.
For the two-dimenisional option pricing problem the fundamental requirement is the flexible choice of the difference quotients, which are used to approximate the first derivative of the option value. But also with this additional flexibility of the discretisation there might be no monotone discretisation based on Finite Differences. Applying this technique, it was shown that under certain sufficient conditions the matrix representation of the non-linear pde is an M-matrix. These sufficient conditions were transferred to the discretisation matrix of higher dimensional option pricing problems.
The local choice of the difference quotient for the first order derivative has material influence on the objective function of the discrete optimisation problem. The other way around, the choice of the difference quotients is affected by the local volatilities and correlations which are the solution of the optimisation problem itself. Thus, the interconnection of the monotone discretisation and the discrete optimisation cause the non-linearity of the discrete option pricing problem. To resolve this dependence, a detailed analysis of the application of different quotients has been done. Furthermore, for the usage of mixed difference quotients a decomposition for the space of admissible volatilities has been developed.
The second working field of solving the discrete option pricing problem concerns the determination of the optimal volatilities and correlations and the global solution of the non-linear problem at every time step. For the global determination of an approximation of the option value for a given time level we applied the policy iteration described by Huang, Forsyth, and Labahn [HFL12] among others. The local determination of optimal volatilities and correlations the optimisation problems

belong to the class of $\mathcal{NP}$-hard problems. The optimisation problem was analysed for options on two and three assets in combination with situations where only centered difference quotients can be used. For discrete pricing problems where mixed difference quotients have to be used several optimisation problems with additional non-linear constraints arise. Therefore the original problem was replaced by one where the control was discretised.
To assure the convergence of the schemes to the viscosity solution it is sufficient to proove consistency, stability, and monotonicity of the scheme. The first two properties for the presented schemes are directly guaranteed as in the linear case. For the BDF1 and the CN scheme we gave a detailed analysis under which conditions monotonicity can be assured. These are similar to those for the schemes applied to linear pdes. The BDF1 scheme is monotone if the discretisation is monotone. For the Crank-Nicolson scheme its monotonicity can be guaranteed, if the step sizes in space and time fulfil a condition of the form $Ch_\tau = h_x^2, C \in \mathbb{R}_+$. The properties which were proven for the European option pricing problem also are valid for American option pricing problems.
The numerical results showed that the rates of convergence regarding time correspond to those expectable from theory. In opposite, the rates which display the speed of convergence if the spatial grid is refined did not match the theoretical expectations but even varied strongly for some examples. But nevertheless the option value is approximated monotonically. Regarding the convergence in time, the CN scheme and the BDF2 scheme -although not consistent- converge with roughly the expected rates. In each example they converged faster than the BDF1 scheme. The evolution of the control for an increasing number of time steps and spatial grid points seems to be consistent, no deteriorationis were oberserved. There is no big difference between option values calculated if the discrete optimisation problem is solved with the KKT conditions or by considering discretised controls, respectively. In fact, the error due to the discretisation of the pde seems bigger than the one caused by using discrete controls.

For pricing problems of options on more than two assets Finite Difference schemes can be applied. But one has to be aware of the fact that they might not always result in a monotone and thus convergent scheme. Moreover, in order to assure convergence a possible reduction of the rates of convergence has to be accepted. Due to the restricitions of the step sizes it is limitedly efficient or difficult to apply on non-uniform grids or even adaptive methods to increase the efficency of the scheme. Adaptive might be an interesting possibility to increase the performance of schemes for one-dimensional option pricing. But the monotonicity of those schemes has to

be analysed carefully.
A possible way to improve the performance of the schemes is parallelization. Especially, for the local optimisation problems for each grid point, which are independent of each other, an increase in efficiency could easily be accomplished. The update of the iteration matrix in the policy iteration offers a similar opportunity.

A Calculation of Deltas

In this chapter we shortly explain how the deltas of options on more than two assets can be calculated. They are used for the Neumann boundary conditions of the discrete non-linear pricing problem described in Chapter 2.
In the first section we give the algorithms and the necessary data to calculate the n-variate cumulative standard normal distribution function for $n = 1, 2$. The resulting algorithms are used for the calculation of the deltas. In the following section we focus on the maximum and the minimum on several assets for which analytic formulas are known. The third section gives an analytic formula for the delta of a geometric average option.

A.1 Evaluating the (multi-variate) standard normal distribution function

For the bivariate normal distribution function we apply the algorithm developed by West [Gen04]. A detailed code is given in the book of Haug [Hau06]. The algorithm needs two arrays v and w of constants which are given in Table A.1 and Table A.2.

Tab. A.1: *Parameter v for the approximation of the bivariate normal distribution function $\mathcal{N}_2$*

i	$v_{i,1}$	$v_{i,2}$	$v_{i,3}$
1	-0.932469514203152	-0.981560634246719	-0.993128599185095
2	-0.661209386466265	-0.904117256370475	-0.963971927277914
3	-0.238619186083197	-0.769902674194305	-0.91223448251326
4	-	-0.587317954286617	-0.839116971822219
5	-	-0.36783149899818	-0.746331906460151
6	-	-0.125233408511469	-0.636053680726515
7	-	-	-0.510867001950827
8	-	-	-0.37370608871542
9	-	-	-0.227785851141645
10	-	-	-0.0765265211334973

Algorithm A.1: Approximation of the bivariate normal distribution function $\mathcal{N}_2$

Input: x, y, ρ
Output: v `//approximation to` $\mathcal{N}_2(x, y, \rho)$

BiVar_CumNormal(x, y, ρ)
  initialise arrays w and v as in Table A.1 and A.2.;
  $NG \leftarrow 3;\ LG \leftarrow 10;$
  if $|\rho| < 0.3$ **then**
    $NG \leftarrow 1;\ LG \leftarrow 3;$
  else if $|\rho| < 0.75$ **then**
    $NG \leftarrow 2;\ LG \leftarrow 6;$
  $h \leftarrow -x;\ \ k \leftarrow -y;\ hk \leftarrow h \cdot k;\ \ z \rightarrow 0;$
  if $|\rho| < 0.925$ **then**
    if $\rho \neq 0$ **then**
      $hs \leftarrow 0.5(h^2 + k^2);\ asr = \arcsin \rho;$
      for $i = 1$ **to** LG **do**
        for $ii = -1$, *step 2,* **to** 1 **do**
          $sn \leftarrow \sin(asr \cdot 0.5(ii * v(i, NG) + 1));$
          $z \leftarrow z + w(i, NG) \cdot \exp((sn \cdot hk - hs) \cdot (1 - sn^2)^{-1});$
      $z \leftarrow z \cdot asr \cdot (4\pi)^{-1};$
    $z \leftarrow z + \mathcal{N}_1(-h) \cdot \mathcal{N}_1(-k);$
  else
    if $\rho < 0$ **then**
      $k \leftarrow -k;\ \ hk \leftarrow -hk;$
    if $|\rho| < 1$ **then**
      $ass \leftarrow (1 - \rho) \cdot (1 + \rho);\ bs \leftarrow (h - k)^2;\ c \leftarrow 0.125 \cdot (4 - hk);$
      $d \leftarrow 0.0625 \cdot (12 - hk);\ asr \leftarrow -0.5 \cdot (bs \cdot ass^{-1} + hk);$
      if $asr > 100$ **then**
        $z \leftarrow \sqrt{ass} \cdot \exp(asr) \cdot (1 - \frac{1}{3}c \cdot (bs - ass)(1 - 0.2.d \cdot bs) + 0.2c \cdot d \cdot ass^2);$
      if $-hk < 100$ **then**
        $z \leftarrow z - \exp(-\frac{1}{2}hk)\sqrt{2\pi} \cdot \mathcal{N}_1\left(-\sqrt{\frac{b}{ass}}\right) \cdot \sqrt{bs} \cdot (1 - \frac{1}{3}c \cdot bs \cdot (1 - 0.2d \cdot bs));$
      for $i = 1$ **to** LG **do**
        for $ii = -1$, *step 2,* **to** 1 **do**
          $xs \leftarrow (0.5\sqrt{ass} \cdot (ii \cdot v(i, NG) + 1))^2;\ \ rs \leftarrow \sqrt{1 - xs};$
          $asr \leftarrow -0.5 \cdot (bs \cdot xs^{-1} + hk);$
          if $asr > -100$ **then**
            $z \leftarrow z + 0.5\sqrt{ass} \cdot w(i, NG) \cdot \exp(asr) \cdot (\exp(-hk \cdot (1 - rs)(2 \cdot (1 + rs))^{-1}) \cdot rs^{-1} - (1 + c \cdot xs \cdot (1 + d \cdot xs)))$
    if $\rho > 0$ **then**
      $z \leftarrow -2z \cdot \pi^{-1} + \mathcal{N}_1(-\max\{h, k\});$
    else
      $z \leftarrow 2z \cdot \pi^{-1} + \max\{0, \mathcal{N}_1(-h) - \mathcal{N}_1(-k)\};$
  return z;

Tab. A.2: *Parameter w for the approximation of the bivariate normal distribution function $\mathcal{N}_2$*

i	$w_{i,1}$	$w_{i,2}$	$w_{i,3}$
1	0.17132449237917	0.0471753363865118	0.0176140071391521
2	0.360761573048138	0.106939325995318	0.0406014298003869
3	0.46791393457269	0.160078328543346	0.0626720483341091
4	-	0.203167426723066	0.0832767415767048
5	-	0.233492536538355	0.10193011981724
6	-	0.249147045813403	0.118194531961518
7	-	-	0.131688638449177
8	-	-	0.142096109318382
9	-	-	0.149172986472604
10	-	-	0.152753387130726

For the evaluation of the uni-variate standard normal distribution we use the approximation of the distribution function of Hastings et al. [HHW55]. It is given in

Algorithm A.2: Approximation of the uni-variate normal distribution function $\mathcal{N}_1$

Input: x
Output: y `//approximation to` $\mathcal{N}_1(x)$

UniVar_CumNormal(x)
- $b_1 \leftarrow 0.319381530$; $b_2 \leftarrow -0.356563782$;
- $b_3 \leftarrow 1.781477937$; $b_4 \leftarrow -1.821255978$;
- $b_5 \leftarrow 1.330274429$;
- $p \leftarrow 0.2316419$; $c \leftarrow 0.9189385332024672$;
- $a \leftarrow |x|$;
- $t \leftarrow (1 + a \cdot p)^{-1}$;
- $s \leftarrow ((((b_5 t + b_4)t + b_3)t + b_2)t + b_1)t$;
- $y \leftarrow s \cdot \exp(-0.5x^2 - c)$;
- **if** $x > 0$ **then**
  - $y \leftarrow 1 - y$;
- **return** y;

The above algorithm has an accuracy error less than $7.5 \cdot 10^{-8}$ according to Hastings et al.

For the evaluation of the tri-variate normal distribution function a code is given in the book of Haug [Hau06].

A.2 Minimum and Maximum options

The multi-variate normal distribution function of p multi-variate normal distributed random variables $X_1, \ldots, X_p$ with mean equal to zero, variances equal to one, and correlations $\rho_{i,j}, i, j \in I_n$ is defined as:

$$\mathcal{N}_p(x_1, \ldots, x_p) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} n_p(u_1, \ldots, u_p) \mathrm{d}u_p \cdots \mathrm{d}u_1 \tag{A.1}$$

with the associated density function

$$n_p(x_1, \ldots, x_p) := \frac{1}{(2\pi)^{\frac{p}{2}} \left(\det\Sigma\right)^{\frac{1}{2}}} \exp\left(\frac{1}{2}(x_i)_i^T \Sigma^{-1}(x_i)_i\right),$$

where Σ is the non-singular covariance matrix of the p random variables.

Let $\phi, \eta \in \{-1, 1\}$. Starting from the analytic formulas of Stulz [Stu82], Reiß and Wystup [RW01] derived closed form expression for the delta of Maximum and Minimum options on two assets. They are given below.
Let

$$\begin{aligned}
\hat{\sigma}^2 &= \sigma_1^2 + \sigma_2^2 - 2\rho_{1,2}\sigma_1\sigma_2, \\
d_j &= \frac{\log\left(\frac{S_j}{K}\right) + \left(r - \delta_j + \frac{1}{2}\sigma_j^2\right)(T-t)}{\sigma_j\sqrt{T-t}}, \quad j = 1, 2, \\
d_3 &= \frac{\log\left(\frac{S_2}{S_1}\right) + \left(\delta_1 - \delta_2 - \frac{1}{2}\hat{\sigma}^2\right)(T-t)}{\hat{\sigma}\sqrt{T-t}}, \\
d_4 &= \frac{\log\left(\frac{S_1}{S_2}\right) + \left(\delta_2 - \delta_1 - \frac{1}{2}\hat{\sigma}^2\right)(T-t)}{\hat{\sigma}\sqrt{T-t}}, \\
d_5 &= \frac{\rho_{1,2}\sigma_2 - \sigma_1}{\hat{\sigma}}, \\
d_6 &= \frac{\rho_{1,2}\sigma_1 - \sigma_2}{\hat{\sigma}}.
\end{aligned}$$

Then the deltas are

$$\begin{aligned}
\frac{\partial V(t, S_1, S_2)}{\partial S_1} &= \phi \exp(-\delta_1(T-t))\mathcal{N}_2(\phi d_1, \eta d_3, \phi\eta d_5), \\
\frac{\partial V(t, S_1, S_2)}{\partial S_2} &= \phi \exp(-\delta_2(T-t))\mathcal{N}_2(\phi d_2, \eta d_4, \phi\eta d_6),
\end{aligned}$$

where $\phi = 1$ for a call, $\phi = -1$ for a put, $\eta = 1$ for a minimum option, and $\eta = -1$ for maximum option.

A.3 Geometric Average options

In Section 2.2.1 we have seen that a Geometric Average option can be modelled as a Vanilla call / put option in the standard Black-Scholes model with modified volatility and dividend rate

$$\tilde{\sigma}^2 := \frac{1}{4}\sum_{i,j=1}^{n} \sigma_i \sigma_j \rho_{i,j}, \qquad \tilde{\delta} := \frac{1}{2}\sum_{i=1}^{n}\left(\delta_i + \frac{1}{2}\sigma_i^2\right) - \frac{1}{2}\tilde{\sigma}^2,$$

and

$$y = \prod_{i=1}^{n} \exp\left(\frac{S_i}{n}\right).$$

The delta of a Vanilla put / call is given by

$$\frac{\partial V(t,y)}{\partial y} = \phi \exp(-\tilde{\delta}(T-t))\mathcal{N}_1(\phi d_1)$$

where $\mathcal{N}_1(x)$ is the standard normal distribution function, c.f. (A.1),

$$d_1 = \frac{\log\left(\frac{y}{K}\right) + \left(r - \tilde{\delta} + \frac{1}{2}\tilde{\sigma}^2\right)(T-t)}{\tilde{\sigma}\sqrt{T-t}}$$

and $\phi = 1$ for a call, and $\phi = -1$ for a put.

B Additional numerical results

In this chapter all additional results to those given in Chapter 5 are summarised.

B.1 Numerical results for an American minimum put

In the following two tables the convergence results for an American minimum put are given. Table B.1 contains the results for the doubling of M_1 and M_2 and Table B.2 those for the doubling of N.

Tab. B.1: *Convergence for increasing number of spatial grid points of the BDF1 (top), BDF2 (middle), and CN (bottom) scheme for an American minimum put with parameters given in Table 5.1,* $N = 500$.

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.200462	-	2.13	3.41	1.1
100	0.203083	-	3.0	4.19	7.9
200	0.203742	3.98	3.02	6.27	39.9
400	0.203908	3.96	3.56	11.14	261.5
800	0.203949	3.99	5.0	22.67	2919.2

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.200540	-	2.1	3.3	1.1
100	0.203164	-	3.0	1.95	7.69
200	0.203825	3.97	3.01	5.41	38.0
400	0.203992	3.95	3.47	9.34	235.5
800	0.204034	3.98	4.53	18.06	2255.0

M_1,M_2	$v(0,\cdot)$	Δ_C^S	I_{PI}	I_{IS}	time [sec.]
50	0.200540	-	2.13	3,83	1.3
100	0.203164	-	3.0	4.62	8.7
200	0.203825	3.97	3.01	6.71	47.7
400	0.203992	3.95	4.17	11.72	345.7
800	0.204034	3.98	4.55	26.73	3144.1

Tab. B.2: *Convergence for increasing number of grid points in $[0,T]$ of the BDF1 (top), BDF2 (middle), and CN (bottom) scheme for an American minimum put with parameters given in Table 5.1, $M_1 = M_2 = 200$.*

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.202980	-	4.08	25.58	25.63
100	0.203394	-	3.54	15.78	57.49
200	0.203609	1.93	3.19	10.1	100.77
400	0.203719	1.94	3.03	6.78	170.41
800	0.203776	1.96	3.0	5.12	292.14

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.203620	-	4.04	21.09	22.71
100	0.203752	-	3.51	12.97	51.45
200	0.203802	2.64	3.13	8.57	92.2
400	0.203822	2.58	3.02	5.91	157.37
800	0.203829	2.64	2.88	4.43	269.33

N	$v(0,\cdot)$	Δ_C^t	I_{PI}	I_{IS}	time [sec.]
50	0.203617	-	4.18	25.67	27.36
100	0.203751	-	3.63	16.0	61.35
200	0.203802	2.65	3.2	10.55	107.72
400	0.203822	2.57	3.02	7.51	181.46
800	0.203829	2.68	3.0	5.79	318.63

Bibliography

[ALP95] **Avellaneda**, M., **Levy**, A., and **Paras**, A.: *Pricing and hedging derivative securities in markets with uncertain volatilities.* Applied Mathematical Finance, 2(2):73–88, 1995.

[Bad98] **Badura**, C.: *C++ solvers for sparse systems of linear equations*, 1998. `http://aam.mathematik.uni-freiburg.de/IAM/Research/projectskr/lin_solver/`.

[Bar96] **Barenblatt**, G. I.: *Scaling, Self-similarity, and Intermediate Asymptotics*, volume 14 of *Cambridge Texts in Applied Mathematics*. Cambridge University Press, New York, 1996.

[BD97] **Bardi**, M. and **Dolcetta**, I. C.: *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations.* Modern Birkhäuser Classics. Springer, 1997.

[BDR95] **Barles**, G., **Daher**, C. H., and **Romano**, M.: *Convergence of numerical schemes for degenerate parabolic equations arising in finance theory.* Mathematical Models and Methods in Applied Sciences, 5(1):125–143, 1995.

[BLA] *Basic Linear Algebra Subprogramms.* `http://www.netlib.org/blas/`.

[BP94] **Berman**, A. and **Plemmons**, R. J.: *Non-negative Matrices in the Mathematical Sciences*, volume 9 of *Classics in Applied Mathematics*. Society for Industrial Mathematics, 1994.

[BS73] **Black**, F. and **Scholes**, M.: *The pricing of options and corporate liabilities.* Journal of Political Economy, 81(3):637–654, 1973.

[BS77] **Brennan**, M. J. and **Schwartz**, E. S.: *The valuation of American put options.* The Journal of Finance, 32(2):449–462, 1977.

[BS90] **Barles**, G. and **Souganidis**, P. E.: *Convergence of approximation schemes for fully non-linear second order equations.* In *29 th IEEE Conference on Decision and Control, Honolulu*, pages 2347–2349, 1990.

[CF02] **Cont**, R. and **Fonseca**, J.: *Dynamics of implied volatility surfaces.* Quantitative Finance, 2(1):45–60, 2002.

[CIL92] **Crandall**, M. G., **Ishii**, H., and **Lions**, P.-L.: *User's guide to viscosity solutions of second order partial differential equations.* Bulletin of the American Mathematical Society, 27(1):1–67, 1992.

[CL83] **Crandall**, M. G. and **Lions**, P.-L.: *Viscosity solutions of Hamilton-Jacobi equations.* Transactions of the American Mathematical Society, 277(1):1–42, 1983.

[CLRS01] **Cormen**, T. H., **Leiserson**, C. E., **Rivest**, R. L., and **Stein**, C.: *Introduction to Algorithms.* MIT Press, 3rd edition, 2001.

[Deu09] **Deutsch**, H. P.: *Derivatives and Internal Models.* Finance and Capital Markets Series. Palgrave Macmillan, 4th edition, 2009.

[DL97] **Dolcetta**, I. C. and **Lions**, P. L. (editors): *Viscosity Solutions and Applications.* Lecture Notes in Mathematics. Springer, 1997.

[Dup94] **Dupire**, B.: *Pricing with a smile.* Risk, 7(1):18–20, 1994.

[Eva10] **Evans**, L. C.: *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics.* Oxford University Press, 2nd edition, 2010.

[Fel08] **Feller**, W.: *An Introduction to Probability Theory and its Applications*, volume 2 of *Wiley Publication in Mathematical Statistics.* Wiley, 2008.

[FG09] **Floudas**, C. A. and **Gounaris**, C. E.: *A review of recent advances in global optimization.* Journal of Global Optimization, 45(1):3–38, 2009.

[FL07] **Forsyth**, P. A. and **Labahn**, G.: *Numerical methods for controlled Hamilton-Jacobi-Bellman pdes in finance.* Journal of Computational Finance, 11(2):1–43, 2007.

[Fle00] **Fletcher**, R.: *Practical Methods of Optimization.* Wiley, 2nd edition, 2000.

[FS06] **Fleming**, W. and **Soner**, H.: *Controlled Markov Processes and Viscosity Solutions*, volume 25 of *Stochastic Modelling and Applied Probability.* Springer, New York, 2nd edition, 2006.

[FV02] **Forsyth**, P. A. and **Vetzal**, K. R.: *Quadratic convergence for valuing American options using a penalty method.* SIAM Journal on Scientific Computing, 23(6):2095–2122, 2002.

[GC06] **Giles**, M. and **Carter**, R.: *Convergence analysis of Crank-Nicolson and Rannacher time-marching.* Journal of Compuational Finance, 9(4):89–112, 2006.

[Gen04] **Genz**, A.: *Numerical computation of rectangular bivariate and trivariate normal and t probabilities.* Statistics and Computing, 14(3):251–260, 2004.

[GJ90] **Garey**, M. R. and **Johnson**, D. S.: *Computers and Intractability: A Guide to the Theory of $\mathcal{NP}$ Completeness.* W. H. Freeman & Co., 1990.

[GvL96] **Golub**, G. H. and **Loan**, C. F. van: *Matrix Computations.* Johns Hopkins University Press, 3rd edition, 1996.

[Hau06] **Haug**, E. G.: *The Complete Guide to Option Pricing Formulas.* McGraw-Hill, 2006.

[Hei09] **Heider**, P.: *A local least-squares method for solving nonlinear partial differential equations of second order.* Numerische Mathematik, 111(3):351–375, 2009.

[Hei10] **Heider**, P.: *Numerical methods for non-linear Black-Scholes equations.* Applied Mathematical Finance, 17(1):59–81, 2010.

[Hes93] **Heston**, S. L.: *A closed-form solution for options with stochastic volatility with applications to bond and currency options.* Review of Financial Studies, 6(2):327–343, 1993.

[HFL12] **Huang**, Y., **Forsyth**, P. A., and **Labahn**, G.: *Combined fixed point and policy iteration for HJB equations in finance.* SIAM Journal on Numerical Analysis, 50(4):1861–1882, 2012.

[HHW55] **Hastings**, C., **Hayward**, J. T., and **Wong**, J. P.: *Approximations for digital computers.* A Research Study by the Rand Corporation, Princeton University Press, 1, 1955.

[HJ94] **Horn**, R. A. and **Johnson**, C. R.: *Topics in Matrix Analysis.* Cambridge University Press, 1994.

[IT07] **Ikonen**, S. and **Toivanen**, J.: *Componentwise splitting methods for pricing American options under stochastic volatility.* International Journal of Theoretical and Applied Finance, 10(2):331–361, 2007.

[JLL90] **Jaillet**, P., **Lamberton**, D., and **Lapeyre**, B.: *Variational inequalities and the pricing of American options.* Acta Applicandae Mathematicae, 21(3):263–289, 1990.

[Joh87] **Johnson**, H.: *Options on the maximum or the minimum of several assets.* Journal of Financial and Quantitative Analysis, 22(3):277–283, 1987.

[KD01] **Kushner**, H. J. and **Dupuis**, P.: *Numerical Methods for Stochastic Control Problems in Continuous Time*, volume 24 of *Stochastic Modelling and Applied Probability.* Springer, 2nd edition, 2001.

[KLM07] **Kovalov**, P., **Linetsky**, V., and **Marcozzi**, M.: *Pricing multi-asset American options: A Finite Element method-of-lines with smooth penalty.* Journal of Scientific Computing, 33(3):209–237, 2007.

[Kwo08] **Kwok**, Y.-K.: *Mathematical Models of Financial Derivatives.* Springer Finance. Springer, 2nd edition, 2008.

[LAP] *Linear Algebra Package.* `http://www.netlib.org/lapack/`.

[Lev04] **Levy**, G.: *Computational Finance: Numerical Methods for Pricing Financial Instruments.* Butterworth-Heinemann, 2004.

[Lio83] **Lions**, P. L.: *Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations part 2: viscosity solutions and uniqueness.* Communications in Partial Differential Equations, 8(11):1229–1276, 1983.

[LL08] **Lamberton**, D. and **Lapeyre**, B.: *Introduction to Stochastic Calculus Applied to Finance.* CRC Financial Mathematics Series. Chapman & Hall, check edition, 2008.

[LO08] **Leentvaar**, C. C. W. and **Oosterlee**, C. W.: *Multi-asset option pricing using a parallel Fourier-based technique.* Journal of Computational Finance, 12(1):1–26, 2008.

[LS95] **Longin**, F. and **Solnik**, B.: *Is the correlation in international equity returns constant: 1960 - 1990?* Journal of International Money and Finance, 14(1):3–26, 1995.

[Lyo95] **Lyons**, T. J.: *Uncertain volatility and the risk-free synthesis of derivatives.* Applied Mathematical Finance, 2(2):117–133, 1995.

[Mer73] **Merton**, R. C.: *Theory of rational option pricing.* The Bell Journal of Economics and Managment Science, 4(1):141–183, 1973.

[PFV03a] **Pooley**, D. M., **Forsyth**, P. A., and **Vetzal**, K. R.: *Numerical convergence properties of option pricing pdes with uncertain volatility.* IMA Journal of Numerical Analysis, 23(2):241–267, 2003.

[PFV03b] **Pooley**, D. M., **Forsyth**, P. A., and **Vetzal**, K. R.: *Two factor option pricing with uncertain volatility.* Lecture Notes on Computer Science, pages 158–167, 2003.

[Poo03] **Pooley**, D. M.: *Numerical methods for nonlinear equations in option pricing.* PhD thesis, University of Waterloo, Canada, 2003.

[PV91] **Pardalos**, P. M. and **Vavasis**, S. A.: *Quadratic programming with one negative eigenvalue is $\mathcal{NP}$-hard.* Journal of Global Optimization, 1(1):15–22, 1991.

[PVF03] **Pooley**, D. M., **Vetzal**, K. R., and **Forsyth**, P. A.: *Convergence remedies for non-smooth payoffs in option pricing.* Journal of Computational Finance, 6(4):25–40, 2003.

[Ran84] **Rannacher**, R.: *Finite Element solution of diffusion problems with irregular data.* Numerische Mathematik, 43(2):309–327, 1984.

[RT00] **Randall**, C. and **Tavella**, D.: *Pricing Financial Instruments - The Finite Difference Method.* Wiley, 2000.

[RV00] **Romagnoli**, S. and **Vargiolu**, T.: *Robustness of the Black-Scholes approach in the case of options on several assets.* Finance and Stochastics, 4(3):325–341, 2000.

[RW01] **Reiß**, O. and **Wystup**, U.: *Computing option price sensitivities using homogeneity and other tricks.* The Journal of Derivatives, 9(2):41–53, 2001.

[Saa03] **Saad**, Y.: *Iterative Methods for Sparse Linear Systems.* Society for Industrial and Applied Mathematics, 2nd edition, 2003.

[SB02] **Stoer**, J. and **Bulirsch**, R.: *Introduction to Numerical Analysis*, volume 12 of *Texts in Applied Mathematics.* Springer, 3rd edition, 2002.

[Sch10] **Schaeling**, D.: *Numerik nichtlinearer Black-Scholes-Modelle für amerikanische Optionen.* Diploma thesis, University of Cologne, 2010.

[Sey12] **Seydel**, R. U.: *Tools for Computational Finance*. Universitext. Springer, 5th edition, 2012.

[SH11] **Schaeling**, D. and **Heider**, P.: *Numerical methods for American options in non-linear Black-Scholes models*. submitted, 2011.

[Shi93] **Shimko**, D.: *Bounds of probability*. Risk, 6(4):33–37, 1993.

[Shr04] **Shreve**, S. E.: *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Finance. Springer, 2004.

[Str86] **Strang**, G.: *Introduction to Applied Mathematics*. Wellesley-Cambridge Press, 2nd edition, 1986.

[Stu82] **Stulz**, R. M.: *Options on the minimum or the maximum of two risky assets: Analysis and applications*. Journal of Financial Economics, 10(2):161–185, 1982.

[Var01] **Vargiolu**, T.: *Existence, uniqueness and smoothness for the Black-Scholes-Barenblatt equation*. Technical Report 5, Universita di Padova, Department of Pure and Applied Mathematics, www.math.unipd.it/∼vargiolu/BSB.pdf, 2001.

[vdV92] **Vorst**, H. A. van der: *Bi-cgstab: A fast and smoothly converging variant of bi-cg for the solution of nonsymmetric linear systems*. SIAM Journal on scientific and Statistical Computing, 13(2):631–644, 1992.

[WDH00] **Wilmott**, P., **Dewynne**, J., and **Howison**, S.: *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press, 2000.

[WF08] **Wang**, J. and **Forsyth**, P. A.: *Maximal use of central differencing for Hamilton-Jacobi-Bellman pdes in finance*. SIAM Journal on Numerical Analysis, 46(3):1580–1601, 2008.

[WF10] **Wang**, J. and **Forsyth**, P. A.: *Numerical solution of the Hamilton-Jacobi-Bellman formulation for continuous time mean variance asset allocation*. Journal of Economic Dynamics and Control, 34(2):207 – 230, 2010.

[WR11] **Witte**, J. H. and **Reisinger**, C.: *A penalty method for the numerical solution of Hamilton-Jacobi-Bellman (HJB) equations in finance*. SIAM Journal on Numerical Analysis, 49(1):213–231, 2011.

[ZFV01] **Zvan**, R., **Forsyth**, P. A., and **Vetzal**, K. R.: *A Finite Volume approach for contingent claims valuation.* IMA Journal of Numerical Analysis, 21(3):703–731, 2001.

[ZFV03] **Zvan**, R., **Forsyth**, P. A., and **Vetzal**, K. R.: *Negative coefficients in two-factor option pricing models.* Journal of Computational Finance, 7(1):37–74, 2003.

[ZWYT09] **Zhang**, K., **Wang**, S., **Yang**, X. Q., and **Teo**, K. L.: *A power penalty approach to numerical solutions of two-asset American options.* Numerical Mathematics, A Journal of Chinese Universities (English Series), 2(2):202–223, 2009.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbstständig angefertigt und die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit -einschließlich Tabellen, Karten und Abbildungen-, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie -abgesehen von den unten angegebenen Teilpublikationen- noch nicht veröffentlicht worden ist sowie, dass sich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmung der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Rüdiger Seydel betreut worden.

München, den 15. November 2014

Albrecht Budke