B-Spline Based Methods: From Monotone Multigrid Schemes for American Options to Uncertain Volatility Models

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Sandra Boschert

aus Leverkusen

Köln, 2020

Berichterstatter: Prof. Dr. Angela Kunoth
Prof. Dr. Helmut Harbrecht
Prof. Dr. Ulrich Langer

Tag der mündlichen Prüfung: 26.10.2020

Abstract

In the first part of this thesis, we consider B-spline based methods for pricing American options in the Black-Scholes and Heston model. The difference between these two models is the assumption on the volatility of the underlying asset. While in the Black-Scholes model the volatility is assumed to be constant, the Heston model includes a stochastic volatility variable. The underlying problems are formulated as parabolic variational inequalities. Recall that, in finance, to determine optimal risk strategies, one is not only interested in the solution of the variational inequality, i.e., the option price, but also in its partial derivatives up to order two, the so-called Greeks. A special feature for these option price problems is that initial conditions are typically given as piecewise linear continuous functions. Consequently, we have derived a spatial discretization based on cubic B-splines with coinciding knots at the points where the initial condition is not differentiable. Together with an implicit time stepping scheme, this enables us to achieve an accurate pointwise approximation of the partial derivatives up to order two. For the efficient numerical solution of the discrete variational inequality, we propose a monotone multigrid method for (tensor product) B-splines with possible internal coinciding knots. Corresponding numerical results show that the monotone multigrid method is robust with respect to the refinement level and mesh size.
In the second part of this thesis, we consider the pricing of a European option in the uncertain volatility model. In this model the volatility of the underlying asset is a priori unknown and is assumed to lie within a range of extreme values. Mathematically, this problem can be formulated as a one dimensional parabolic Hamilton-Jacobi-Bellman equation and is also called Black-Scholes-Barenblatt equation. In the resulting non-linear equation, the diffusion coefficient is given by a volatility function which depends pointwise on the second derivative. This kind of non-linear partial differential equation does not admit a weak H^1-formulation. This is due to the fact that the non-linearity depends pointwise on the second derivative of the solution and, thus, no integration by parts is possible to pass the partial derivative onto a test function. But in the discrete setting this pointwise second derivative can be approximated in H^1 by L^1-normalized B-splines. It turns out that the approximation of the volatility function leads to discontinuities in the partial derivatives. In order to improve the approximation of the solution and its partial derivatives for cubic B-splines, we develop a Newton like algorithm within a knot insertion step. Corresponding numerical results show that the convergence of the solution and its partial derivatives are nearly optimal in the L^2-norm, when the location of volatility change is approximated with desired accuracy.

Kurzzusammenfassung

Der erste Teil dieser Arbeit beschäftigt sich mit B-spline basierten Verfahren zur Bewertung Amerikanischer Optionen anhand des Black-Scholes und Heston-Modells. Die beiden Modelle unterscheiden sich in der Annahme an die Volatilität des Basiswertes. Während im Black-Scholes-Modell die Volatiliät als konstant vorausgesetzt wird, wird im Heston-Modell eine stochastische Volatilität angenommen. Die resultierenden Probleme werden als parabolische Variationsungleichungen formuliert. Um optimale Risikostrategien zu entwickeln, sind in der Finanzwelt neben der Berechnung des Optionspreises auch deren partiellen Ableitungen bis zur Ordnung zwei, die sogenannten Griechen, von besonderem Interesse. Eine Besonderheit der Optionspreisprobleme ist, dass die Anfangsbedingung üblicherweise als stückweise lineare stetige Funktion gegeben ist. Aufgrund dessen werden die Probleme hinsichtlich des Ortes mit kubischen B-Splines und zusammenfallenden Knoten an solchen Punkten, wo die Anfangsbedingung nur stetig ist, diskretisiert. Dieser Ansatz ermöglicht zusammen mit einem impliziten Zeitschrittverfahren die punktweise genaue Approximation der Griechen. Zur effizienten Lösung der diskreten Variationsungleichung haben wir ein monotones Mehrgitterverfahren für (Tensorprodukt-)B-Splines mit zusammenfallenden Knoten im Inneren des Gebietes entwickelt. Zugehörige numerische Resultate zeigen, dass das monotone Mehrgitterverfahren robust bezüglich der Verfeinerungslevel und Gitterweiten ist.
In dem zweiten Teil dieser Arbeit werden mit dem sogenannten Uncertain-Volatility-Modell (Modell mit unsicherer Volatilität) Europäische Optionen bewertet. In diesem Modell ist die Volatilität a priori nicht bekannt und es wird angenommen, dass sie in einem Intervall von Extremwerten liegt. Mathematisch kann dieses Problem als eine eindimensionale parabolische Hamilton-Jacobi-Bellman-Gleichung formuliert werden und wird auch Black-Scholes-Barenblatt-Gleichung genannt. In der resultierenden nichtlinearen Gleichung ist der Diffusionskoeffizient durch eine Volatilitätsfunktion gegeben, die punktweise von der zweiten partiellen Ableitung der Lösung abhängt. Diese Art von nichtlinearen partiellen Differentialgleichungen haben keine schwache H^1-Formulierung. Das liegt daran, dass der nichtlineare Term von der punktweisen Auswertung der zweiten Ableitung der Lösung abhängt und keine partielle Integration möglich ist, um die partielle Ableitung auf die Testfunktion zu übertragen. Aber im diskreten Fall kann diese zweite partielle Ableitung in H^1 mit L^1-normalisierten B-Splines approximiert werden. Es stellt sich heraus, dass die Approximation der Volatilitätsfunktion zu Unstetigkeiten in den partiellen Ableitungen führt. Um die Approximation der Lösung und dessen partielle Ableitungen für kubische B-splines zu verbessern, wird das Newton Verfahren um einen Schritt erweitert, in dem Knoten eingefügt werden. Zugehörige numerische Resultate zeigen, dass die Konvergenzraten für die Lösung und dessen partielle Ableitungen fast optimal sind, wenn die Stelle, wo sich die Volatilität verändert, genau genug approximiert wird.

1. Introduction

Motivation and overview

The pricing of options plays an important role on financial markets. Options can be seen as some kind of insurance against future price fluctuations. More precisely, an option is a financial contract, which allows but not obligates the holder to buy (Call) or to sell (Put) an underlying asset S at a specific time T (European option) or within a time period $(0, T]$ (American option) at a specific strike price K. In return the holder receives a non-negative payoff $\mathcal{H}(S)$. In financial mathematics option pricing has been consistently gaining attention since the 1970's, when Black and Scholes introduced the Black-Scholes equation.
The central goal in this research area is to find a fair price of an option. Nevertheless, to determine optimal risk strategies, it is not only important to focus on the option value but also on its sensitivities to a change in the underlying parameter (or model parameter) on which the value of a portfolio or financial instrument is dependent. Mathematically, these sensitivities or so called *Greeks* are described by the partial derivatives of the option value with respect to the model parameters up to order two.
To answer all these questions regarding financial markets many different models have been developed. The most famous one is the *Black-Scholes model* developed by Black and Scholes in 1973 [BS73] in which the option value $V(t, S)$ as a function of the underlying asset $S \in \mathbb{R}^+$ and time $t \in [0, T]$ is described by a parabolic partial differential equation. Moreover, this equation depends on a strike price K, a risk free interest rate r, dividend yields D_0 and on a *constant volatility* σ. One of the main limitations of the model is the assumption of constant volatility and can expose the user to an unexpected risk. To study the fluctuation of the underlying asset researchers started to determine the implied volatility, which can be determined in the Black-Scholes model from real observed option prices. It was discovered that the implied volatilities for different strikes and maturities vary which gives a so called *volatility smile* or skewness effect.
A lot of models were developed to explain this empirical fact. In the most obvious approach, suggested by Merton in 1973 [Mer73], the volatilities are imposed to be a deterministic function of time. This model explains the implied volatilities for different maturities but it still does not consider the effect on the implied volatilities for varying strike prices. To overcome this difficulty Heston developed the *Heston stochastic volatility model* in 1993 [Hes93], in which not only the underlying asset but also the volatility is assumed to follow a stochastic process. Here the option price $V(t, S, v)$ as a function of the stock price $S \in \mathbb{R}^+$, time $t \in [0, T]$ and variance $v \in \mathbb{R}^+$ (root of the volatility) is determined by a parabolic two dimensional partial differential equation. As pointed out by [Hes93], this model allows an arbitrary correlation between the underlying asset

and the volatility and is able to clarify the skewness effect in the Black-Scholes model. Another approach, where the volatility is a priori unknown and is only assumed to lie within a range of extreme values, the so called *Uncertain Volatility model* was developed independently by Avellaneda, Levy and Paras [ALP95] and Lyons [Lyo95] in 1995. As pointed out by [ALP95] the bounds can be inferred from high-low peaks of option-implied volatilities and be considered as defining a confidence interval for future volatilities. With these extreme values of the volatility the option value can be computed in a worst and best case scenario. Assuming the worst case the holder can hedge his position and the value of the portfolio is increasing minimally, regardless of the actual volatility movement. In the resulting non-linear pricing equation, also known as *Black-Scholes-Barenblatt equation*, the volatility depends on the time t, the underlying asset S and on the second derivative of the option value. For simple convex payoff functions the model reduces to the Black-Scholes Problem with one of the maximal or minimal value of the volatility. Therefore, most authors in the literature study non-convex payoff functions such as butterfly-spread or barrier options [Bud13, FP03, Hei10, Sen15, ZW09].

In the first part of this thesis we consider the pricing of *American options with Black-Scholes and Heston's model*. A fundamental difficulty arises from the fact that American options can be exercised at any time; thus leading to a free boundary problem. For American put options in Heston's model with payoff $\mathcal{H}_{\mathrm{P}}(S) := \max(0, K - S)$ the free boundary problem reads as follows: Find $V(t, S, v)$ with $(t, S, v) \in [0, T) \times \Omega$ and $\Omega \subset \mathbb{R}^2$ a domain such that

$$\begin{aligned} V &= \mathcal{H}_{\mathrm{P}}(S) \quad && \text{for } S \leq S_f(t, \cdot), \\ \frac{\partial V}{\partial t} + \mathcal{L}V &= 0 \quad && \text{for } S > S_f(t, \cdot) \end{aligned} \tag{1.1}$$

with end condition $V(0, S, v) = \mathcal{H}_B(S)$ and appropriate boundary conditions. In Heston's model the differential operator is a two dimensional parabolic equation; thus leading to a free boundary $S_f(t, v) : [0, T] \times \mathbb{R}^+ \to \mathbb{R}^+$ depending on time t and variance v. In the Black-Scholes model the partial differential operator $\mathcal{L}V(t, S)$ is a one-dimensional parabolic equation and the free boundary $S_f(t) : [0, T] \to \mathbb{R}^+$ is a function depending on time t. In option pricing problems the free boundary can be interpreted as the optimal underlying price to which exercising the option is optimal.
In order to obtain a formulation where the a priori unknown free boundary does not explicitly appears, the problem is reformulated as a *linear complementarity problem*. A classical way to approximately solve such problems is the finite difference method. In a finite difference method, one defines a finite number of grid points and approximates the partial derivatives with difference quotients. This class of discretizations in the context of American options has been analyzed by many authors [CP99, IT09, Oos03, Sey12, ZFV98]. One well-known drawback of finite difference methods is that the discrete solution is only given at the grid points and that strong regularity assumptions on the solution are required to obtain error estimates.
Another approach is to reformulate the linear complementarity problem as a *parabolic variational inequality*. This formulation is derived by multiplying test functions from a closed convex set in $H^1(\Omega)$ and integrating over the spatial domain Ω. Finally, the vari-

ational inequality (weak formulation) is derived by integration by parts. Well-posedness for (parabolic) variational inequalities is often derived by using a method called *penalization* (cf. [BL82, KS00, KTK17, Mem12]). The penalty approach consists of substituting the variational inequality by a family of non-linear equations and proving some results for the penalized problem. Finally, it can be shown that their solutions converge to the solution of the variational inequality and the solution is unique. This method is also used in this thesis to establish well-posedness for the Black-Scholes and Heston variational inequality. Note that the well-posedness of the Black-Scholes variational inequality for put options and call options with dividend payments were already established in [Ach05, DL19b].

A parabolic variational inequality is typically solved approximately by a time stepping method and by replacing the convex set by a linear finite element space. Thus, in each time step a discrete linear complementarity system is solved. Essential to the success of a spatial discretization with linear basis functions is the nodal basis property that enables an appropriate approximation of the constraint set by comparing the function values. For the numerical computation to solutions of elliptic variational inequalities on closed convex sets employing a discretization with linear basis functions, error estimates, adaptive methods and monotone multigrid solvers have been investigated over the past decades [BHR77, Fal74, GK09, Kor94, Kor96, Man84]. This methods for a discretization with linear basis functions are already applied with success to the semi-discrete variational inequality arising in the valuation of American options [DL19a, HRSW13, KSW12, KTK17, Mau06, Mem12, Sey12, Zha07].

However, of particular importance in the valuation of options are the partial derivatives of the option price up to order two (*Greeks*) to determine optimal risk strategies. A special feature for these option price problems is that initial conditions are typically given as piecewise linear continuous functions. Consequently, in this work, we have derived a spatial discretization based on cubic B-splines with coinciding knots at the points where the initial condition is not differentiable. Together with an implicit Euler scheme, this enables us to achieve an accurate *pointwise approximation of the partial derivatives up to order two*. This is confirmed by numerical computations for an American put option as well as for a call option with dividend payments in the Black-Scholes model and a put option in Heston's model. In former approaches the authors applied a cubic B-spline discretization without internal coinciding knots to the Black-Scholes ([Hol04]) and Heston variational inequality [Bos15, Wei14]. In particular in [Bos15, Wei14] one could observe that an approximation of the initial condition without internal coinciding knots results in oscillations in both the first and second partial derivatives of the numerical solution.

In the context of variational inequalities it should be noted that, due to the free boundary, classical solutions of variational inequalities do not exist. In particular, for elliptic variational inequalities the second derivative of the solution is discontinuous along the free boundary. The threshold of smoothness for elliptic variational inequalities has been established by [Bré71]. It can be shown that for a sufficiently smooth boundary, right side and obstacle function in $C^3(\overline{\Omega})$, the solution lies in the Sobolev-Slobodeckij space $W^{s,p}(\Omega) \cap W^{2,\infty}(\Omega)$ for all $s < 2 + 1/p$ and $1 < p < \infty$. Based on this realistic smoothness assumption, the authors in [BHR77] have established optimal $\mathcal{O}(h^{3/2-\varepsilon})$ a priori

estimates for quadratic basis functions in the $H^1(\Omega)$-norm. Since the obstacle function for pricing American options are given as a piecewise linear function, it is still unknown whether the solution to the semi-discrete Black-Scholes or Heston variational inequality satisfies this regularity result. But numerical experiments for the Black-Scholes and Heston variational inequality with quadratic ($k = 3$) and cubic ($k = 4$) (tensor product) B-splines confirm the optimal convergence rate of $\mathcal{O}(h^{3/2-\varepsilon})$ in the $H^1(\Omega)$-norm for the semi-discrete solution, when the initial condition is approximated with $k - 1$ coinciding knots at those points where the initial condition is not differentiable. This convergence rates for American option pricing problem seems to be never discovered in the previous literature. At this point it should be mentioned that the constraint for an American put option in the Black-Scholes or Heston variational inequality are binding for $S \leq S_f < K$, where the obstacle function is smooth. A similar consideration applies to the Black-Scholes variational inequality with dividend yields for an American call option. Thus, it seems that the irregularity of the obstacle function at $S = K$ has no influence on the smoothness of the solution. We also provide numerical results with linear basis functions, where optimal $\mathcal{O}(h)$ convergence in the $H^1(\Omega)$-norm can be observed. A priori estimates and corresponding numerical results for the Black-Scholes variational inequality with linear basis functions and an implicit time stepping method in Bochner spaces can also be found in [DL19a, DL].
For an efficient computation of numerical solutions we propose a *monotone multigrid method* (MMG) for (tensor product) B-splines with possible coinciding knots. In order to maintain the robustness and monotonicity of the scheme we construct a quasi-optimal monotone coarse grid approximation which is based on the B-spline expansion coefficients. The constructed method profits heavily from the positivity of B-splines. A MMG method for higher order B-splines without coinciding knots was first established by [Hol04, HK07]. It is shown here that the method is globally convergent and reduces asymptotically to a subspace correction method when the contact set is identified. Finally, the method is applied to the Black-Scholes and Heston variational inequality for linear, quadratic and cubic (tensor product) B-splines. In particular, it is observed that the MMG method is robust with respect to the refinement level and mesh size, but the convergence rate and number of iterations increase with the order of the B-spline basis functions.

In the second part of this thesis we consider the valuation of a European butterfly-spread option in the *uncertain volatility model*. The resulting non-linear *Black-Scholes-Barenblatt equation* is a parabolic partial differential equation of Hamilton-Jacobi-Bellman type. Hamilton-Jacobi-Bellman (HJB) equations are fully non-linear second order equations where the differential operator is included in an infimum or supremum. The equation has been introduced in the 1950's by Rowan Hamilton, Carl Gustav Jacobi and Richard Bellman. This type of equation arises from models for optimal control with stochastic processes. A butterfly-spread option is a trading strategy where positions with three different strike prices can be taken. It can be constructed by buying a call option at a strike price K_1, buying a call option at a higher strike price $K_2 > K_1$

and selling two call options at strike price $K := (K_1 + K_2)/2$. In the *worst case scenario* the problem is given as follows: Find $V(t,S)$ with $(t,S) \in [0,T) \times \mathbb{R}^+$ such that

$$F(V) := V_t + \inf_{\underline{\sigma} \in \Sigma} (\mathcal{L}^{\underline{\sigma}} V) := V_t + \inf_{\underline{\sigma} \in \Sigma} \left(\frac{1}{2} S^2 \underline{\sigma}^2 V_{SS} \right) + (r - D_0) S V_S - rV = 0 \text{ in } [0,T) \times \mathbb{R}^+ \tag{1.2}$$

with $\Sigma := [\sigma_{\min}, \sigma_{\max}]$, end condition $V(T,S) = \mathcal{H}_{\mathrm{BS}}(S)$ and appropriate boundary conditions. In this setting $V(t,S)$ corresponds to the option price with dynamic hedging with underlying asset under the worst case volatility path and solves (1.2) with

$$\underline{\sigma}^- [V_{SS}] := \begin{cases} \sigma_{\min}^2 & \text{if } S^2 V_{SS} \geq 0, \\ \sigma_{\max}^2 & \text{if } S^2 V_{SS} < 0. \end{cases} \tag{1.3}$$

The corresponding problem in the best case scenario is obtained by replacing the infimum by a supremum.

For the well-posedness for classical linear parabolic equations one usually introduces a weak formulation. Assuming that the bilinear form is bounded and coercive one derives the well-posedness in the Bochner space $L^2(0,T,H_0^1(\Omega)) \cap H^1(0,T,H^{-1}(\Omega))$ by the well known Lax-Milgram like theorem for parabolic equations [Eva98]. This approach is not applicable for the Black-Scholes-Barenblatt equation or HJB equations in general, since it does not admit a weak formulation. This is due to the fact that the non-linearity depends pointwise on the spatial second derivative of the solution, and thus no integration by parts is applicable to pass the partial derivative onto a test function.

A novel theory for parabolic HJB equations that satisfy the *Cordes condition* has been established in [Sme15] or in the later published article [SS16]. The Cordes condition is an algebraic assumption on the coefficients in the differential operator and has originally been introduced for elliptic equations in non-divergence form. For partial differential equations in non-divergence form it is usually not possible to introduce a weak formulation when the diffusion coefficients are not sufficiently regular. Assuming the Cordes condition and a convex bounded domain Ω enables us to prove existence and uniqueness of the solution in the space $H^2(\Omega) \cap H_0^1(\Omega)$. It turns out that this analysis for the well-posedness fits also in the context of HJB-equations. In this thesis we apply the theory from [Sme15, SS16] to prove the well-posedness for the Black-Scholes-Barenblatt equation. Since this theory was developed for non-degenerate diffusion coefficients we introduce a log-transformation $x := \log(S/K)$ to avoid a degenerate coefficient in S. In particular it is shown that the Black-Scholes-Barenblatt equation is well-posed in the Bochner space $L^2(0,T;\mathcal{W}) \cap H^1(0,T;L^2(\tilde{I}))$ with $\mathcal{W} := H^2(\tilde{I}) \cap H_0^1(\tilde{I})$ and $\tilde{I} := (S_{\min}, S_{\max})$, $0 < S_{\min} < S_{\max}$.

Since no exact solution formula is available in the simple case of European option with a non-convex payoff function, many authors have studied finite difference or finite volume methods to approximately solve the Black-Scholes-Barenblatt equation. Essential to their success was that Barles and Sougandis in [BS91] provided a convergence analysis which is applicable to low order discretizations. In particular, it was shown in [Var01], that the Black-Scholes-Barenblatt equation has a unique viscosity solution. Therefore, the Barles and Souganidis theorem provides the convergence to the viscosity solution if the method is consistent, stable and monotone. This was proven for the

Black-Scholes-Barenblatt equation with a finite difference method in [FP03, Hei10] and with a finite volume method in [ZW09]. Moreover, the author in [Bud13] has discussed finite difference methods for options with one or two underlying assets.
Considerably less literature in the context of HJB equations is available for finite element methods or discontinuous finite element methods. However, the authors in [JS13] provide a finite element method for possible degenerate parabolic HJB equations. In particular, they introduce an approximation of the non-linear second order term in $H^1(\Omega)$ with L^1-normalized linear finite element functions. Furthermore, it was shown that the finite element solution converges to the viscosity solution in the L^∞-norm, as well as in the H^1-norm, if there is a non-degenerate subset of the discrete spatial operator. A higher order discretization with a discontinuous finite element method for elliptic and parabolic HJB equations is established in [Sme15, SS16]. Considering the elliptic case the authors introduce a $H^2(\Omega)$-formulation on a bounded domain Ω and provide well-posedness in $H^2(\Omega) \cap H_0^1(\Omega)$, assuming the Cordes condition. Particularly, the introduction of the Cordes condition enables the development of a discontinuous finite element method with a complete analysis in terms of consistency, stability and error bounds.
However, as mentioned before, in financial mathematics a pointwise accurate approximation of the Greeks is highly relevant to determine optimal risk strategies. Thus, a higher order discretization is preferable for a pointwise approximation of the Greeks up to second order. In this thesis we provide an approximation with B-spline basis functions to approximately solve the Black-Scholes-Barenblatt equation with high precision when the location of the volatility change is approximated with desired accuracy. An essential advantage of this method is its simplicity. In the linear case this method is equivalent to the monotone finite element method from [JS13], where a discretization with L^1-normalized linear finite element functions is analyzed.
One particular difficulty arises from the fact that no weak formulation for the Black-Scholes-Barenblatt equation is available in order to use a classical Galerkin approach for the discretization. For this reason, the second order term in the non-linear partial differential operator is approximated in $H^1(I)$ with $L^1(I)$-normalized B-splines of order k. The corresponding problem is solved by a semismooth Newton algorithm. It turns out that the weak approximation of the volatility function in (1.3) leads to discontinuities in the first derivatives and to small oscillations in the second derivative of the approximated solution. This is due to the fact that the locations of volatility change given by the zeros of the second derivative are only an approximation and the second derivatives are not exactly zero at those points. Since this approximation of the volatility function needs only a stable approximation of the solution in $H^1(I)$ an unstable approximation of the second derivative has no direct influence on the approximation of the volatility function. But it is clear that this approach leads to low order convergence rates for cubic B-splines. To obtain a stable approximation of the second derivative with B-splines of order k it is necessary to repeat those knots $k-1$ times where the volatility jumps from $\sigma_{\min}^2$ to $\sigma_{\max}^2$ or vice versa. In order to improve the approximation of the solution we compute the approximation of the volatility function by a pointwise evaluation of the second derivative with cubic B-splines in the semismooth Newton step and insert knots at the zeros of the second derivative. Corresponding numerical results

show that the convergence of the solution and its partial derivatives are nearly optimal in the L^2-norm, when the location of volatility change is approximated with desired accuracy.

The *central goal of this thesis* is the highly accurate numerical computation of the American option price and its partial derivatives up to order two based on a cubic (tensor product) B-spline discretization with coinciding knots. In particular we consider the American option price in the Heston or Black-Scholes variational inequality and the European option price with the Black-Scholes-Barenblatt equation. While the formulation of the Black-Scholes variational inequality has been already formulated in [Ach05, DL19b], we establish a variational inequality for Heston's problem. In particular, the derivation of the weak form in this thesis differs from [Bos15, Bur16, KSW12, Wei14] due to an error in the integration by parts at the boundaries where a Neumann boundary condition holds. Moreover, we prove the existence and uniqueness of a solution for the Heston variational inequality derived in this thesis.
The special feature of this thesis is the discretization of the mentioned problems based on a (tensor product) B-spline approximation with coinciding knots at the locations where the initial condition (or numerical solution) is not differentiable. Former approaches [Bos15, Hol04, Sen15, Wei14] are based on cubic (tensor product) B-splines without coinciding knots which results in oscillations in the partial derivatives. For the highly accurate and fast numerical computation of the American option price we have developed a monotone multigrid method based on coinciding knots. In the scope of this thesis a Matlab package for the numerical realization of the B-spline discretization for the above mentioned problems with linear, quadratic and cubic (tensor product) B-splines and corresponding numerical solution algorithms was developed. Most of the function works for arbitrary B-spline order, but due to the low regularity of the solution to variational inequalities a linear, quadratic and in particular cubic B-spline discretization is recommended. Several numerical results confirm that a cubic B-spline discretization with coinciding knots facilitates a pointwise highly accurate approximation of the partial derivatives up to order two. These numerical results for the considered option pricing problems with a quadratic or cubic B-spline discretization seems to be never discovered in the previous literature.

Outline

Chapter 2: In Section 2.1, we briefly introduce some financial option pricing instruments which are used throughout the thesis. Then the mathematical descriptions of the Black-Scholes model, Heston stochastic volatility model and the uncertain volatility model are presented. An overview of the mathematical definitions of the Greeks and their relevance on a financial market is discussed. In Section 2.2 we concentrate on the valuation of American options. We present a weak formulation within the Black-Scholes equation as a parabolic variational inequality in a weighted Sobolev space on a bounded interval for put and call options. Note that this formulation has already been established in [Ach05, AP05, DL19a, DL19b]. Furthermore, the valuation of American put options with Heston's model is presented. Therefore, we introduce a log-transformation $x = \log(S/K)$ and the corresponding problem is formulated as a parabolic variational inequality in two dimensions on a bounded domain. In this regard it turns out that one has to proceed somewhat differently in the integration by parts as in [Bos15, BHSW15, KSW12, Wei14].
Chapter 3: Section 3.1 is devoted to study some regularity results for variational inequalities. These results are important for two reasons, namely for the error analysis for further numerical computations and to clarify the maximal smoothness of the Greeks. First, we introduce existence and uniqueness results of weak solutions to elliptic and parabolic variational inequalities, which can be found in standard textbooks [KS80, BL82]. In order to prove higher regularity one main difficulty for American option pricing problems arises from the fact that the obstacle function given by a piecewise linear continuous function is only once weak differentiable. Under some appropriate assumption on the obstacle and the bilinear form we present a regularity result for elliptic and parabolic variational inequalities in Sobolev or Bochner spaces, which fits into the context of many option pricing problems. The results are based on [Mem12] and [KTK17], where the log-transformed Black-Scholes and the partial integro-differential variational inequality are considered.
Moreover, provided that the obstacle function and right side is smooth enough, the maximal smoothness to elliptic variational inequalities from [Bré71] is discussed.
Fundamental properties on B-splines are presented in Section 3.2. Relevant properties on splines for the implementation and their approximation properties are presented.
Of special interest in this thesis is the spatial discretization of the Black-Scholes and Heston variational inequality after a semi-discretization. Thus, in Section 3.3 we present a priori estimates for elliptic variational inequalities for a conforming discretization with linear and quadratic B-spline basis functions based on the results from [BHR77]. In particular, under realistic regularity assumptions for elliptic variational inequalities optimal $\mathcal{O}(h)$ and $\mathcal{O}(h^{3/2-\epsilon})$ error bounds in the $H^1(\Omega)$-norm are obtained for linear and quadratic B-splines. Due to the limited smoothness of the solution to an elliptic variational inequality the global error bound for a uniform grid size cannot be improved when the B-spline order is increased for $k > 3$. Note that the required smoothness result for the semi-discrete Black-Scholes or Heston variational inequality in the case of quadratic basis function is still outstanding.
Chapter 4: In Section 4.1 we discuss the well-posedness of the Black-Scholes vari-

ational inequality in a weighted Sobolev space. Particularly, we employ the results presented in Section 3.1 to the Black-Scholes variational inequality for a put and call option. Note that these results are already established in [AP05, Ach05, DL19a, DL19b] but the proof in this thesis is based on the techniques from [Mem12] and [KTK17].
Next, we introduce a semi-discretization of the Black-Scholes variational inequality and discuss different time stepping methods in Section 4.2.
Furthermore, we derive a spatial B-spline discretization of the Black-Scholes variational inequality. Since initials condition for option pricing problems are typically given as piecewise linear continuous functions we derive a spatial discretization based on linear ($k = 2$), quadratic ($k = 3$) or cubic ($k = 4$) B-splines with $k - 1$ coinciding knots at the points, where the initial condition is continuous.
Chapter 5: In Section 5.1 we discuss the well-posedness of the Heston variational inequality for a put option. Assuming a strictly positive variance and appropriate assumptions on the data we prove that the bilinear form, derived in this thesis, is bounded and satisfies a Gårding inequality. Therefore, the result presented in Section 3.1 implies that there exists a unique weak solution to the parabolic variational inequality. Since the diffusion part of the derived bilinear form is not symmetric the results presented in Section 3.1 to prove higher regularity in Bochner spaces are not applicable. Thus, we prove $H^2(\Omega)$-regularity for the semi-discrete problem.
After semi-discretization a spatial tensor product B-spline discretization with coinciding knots of the Heston variational inequality is presented in Section 5.2. Due to the tensor product structure, the corresponding discretization matrices can be expressed as sums of Kronecker products of matrices with respect to one coordinate. This enables an efficient implementation.
Chapter 6: Next, we discuss the approximation of the Greeks. Since the derivatives of splines can be expressed as a sum of B-splines of lower order, the approximation of the partial derivatives can be computed efficiently by using a Neville-like scheme.
Chapter 7: In order to accelerate basic iterative schemes we propose a monotone multigrid method for the fast and efficient numerical solution of the semi-discrete variational inequality with higher order B-splines arising in the valuation of American options. The novel approach facilitates the numerical solution of discrete variational inequalities arising from a uniform (tensor product) B-spline discretization with possible internal coinciding knots. Former approaches in [Hol04, HK07, Bos15] are based on uniform B-splines without repeating knots. In order to ensure the robustness, or monotonicity, of the multigrid scheme monotone coarse grid approximations with possible coinciding knots are proposed. Finally, the global convergence of the MMG method is proven.
Chapter 8: In order to verify the B-spline discretization and the corresponding numerical solution algorithm to solve elliptic variational inequalities a one dimensional test problem is constructed in Section 8.1. In Section 8.2, several numerical results are proposed to study the discretization error of the solution and its partial derivatives for the Black-Scholes and Heston variational inequality. Finally, several numerical experiments show that the proposed monotone multigrid method is robust with respect to the refinement level and mesh size.
Chapter 9: In this chapter we consider the valuation of a European butterfly-spread

option in the Uncertain Volatility model. In Section 9.2 we prove the well-posedness of the Black-Scholes-Barenblatt equation by assuming the Cordes condition. This theory in the context of HJB equations was first established in [Sme15, SS16].
The discretization with a semi-discretization and a spatial B-spline discretization is discussed in Section 9.3. Since the initial condition given as the butterfly-spread option payoff is also a piecewise linear continuous function, a B-spline discretization with coinciding knots at the critical points is proposed. Since no weak formulation of the Black-Scholes-Barenblatt equation is available in the continuous setting, we present an approximation of the second derivative in a weak form with L^1-normalized B-splines. It turns out that the approximation of the non-linearity or volatility function in $H^1(I)$ leads to discontinuities in the partial derivatives at those knots where the volatility changes.
In order to construct a semi-smooth Newton method to solve the discrete Black-Scholes-Barenblatt equation, one has to find an analog formulation of the Jacobian since the discrete operator of the Black-Scholes-Barenblatt equation is not differentiable in the classical sense. Thus, we first discuss slant-differentiability in Section 9.4. This concept was also used in [Sme12, BMZ09] to solve discrete HJB equations. In order to stabilize the approximation of the partial derivatives knots are inserted at the location of volatility changes. Corresponding numerical results for cubic B-splines in Section 9.5 show that the convergence rates of the solution and its partial derivatives are nearly optimal in the L^2-norm, as for linear variational equations, when the locations of volatility changes are approximated with desired accuracy.

Danksagung

Mit der größte Dank gilt Frau Prof. Dr. Angela Kunoth für ihre ausgezeichnete fachliche und persönliche Betreuung. Ohne ihre motivierenden Worte gegen Ende meiner Masterarbeit und ihr entgegen gebrachtes Vertrauen wäre die Arbeit sicherlich so nicht zustande gekommen. Darüber hinaus möchte ich mich für die vielen Workshops bedanken, an denen ich dank ihr teilnehmen konnte. Zu guter letzt danke ich ihr auch für das Korrekturlesen meiner Dissertation.
Ein weiterer großer Dank gebührt dem Zweitgutachter Prof. Dr. Helmut Harbrecht und dem Drittgutachter Prof. Dr. Ulrich Langer für all ihre aufgewendete Zeit im Zusammenhang mit meiner Dissertation.
Außerdem möchte ich mich auch bei den weiteren Mitgliedern der Prüfungskommission Dr. Roman Wienands und Prof. Dr. Kathrin Bringmann für ihren Einsatz bedanken. Leider werden wir wohl aufgrund der Corona Pandemie auf ein gemeinsames Beisammensein nach der Verteidigung verzichten müssen.
I want to thank Prof. Dr. Charalambos Makridakis for the discussion about the Rannacher timestepping method at the workshop in Warwick. Moreover, I want to thank Dr. Iain Smears for his detailed reply to my questions about HJB equations.
Im Laufe meiner Promotion haben mich viele Kollegen begleitet. An erster Stelle möchte ich mich bei Frau Basmaji für ihren Einsatz bedanken. Sie hat in nur wenigen Tagen kurz vor ihrem Urlaub den schriftlichen Teil meiner Dissertation überprüft.
Meinem Kollegen Samuel Leweke möchte ich für viele mathematische sehr hilfreiche Diskussionen danken. Meinem langjährigen Bürokollegen Dr. Boqiang Huang möchte ich für die schöne gemeinsame Zeit und für die verschiedenen fachlichen Gespräche danken. Gerne erinnere ich mich auch an den Besuch in Bestwig oder den Ski-Ausflug mit den ehemaligen AG-Mitgliedern Christian Mollet, Stephan Gerster und Enis Sen. Ein besonderer Dank gilt Stephan Gerster für das Korrekturlesen meiner Dissertation. Auch den ehemaligen Mitgliedern der AG Gassner möchte ich für die vielen für mich sehr interessanten Fußballgespräche beim Mittagessen danken. Bei Marvin Bohm möchte ich mich an dieser Stelle für das Korrekturlesen meiner Einleitung und die beruhigenden Worte bedanken. In diesem Zusammenhang möchte ich mich auch bei seiner Frau Juliana Bohm für ihre langjährige Freundschaft und die warmherzigen, stärkenden Worte während meiner Promotion bedanken. Auf privater Ebene möchte ich mich auch bei meiner langjährigen Freundin Julia Jäschke für all ihre motivierenden Worte und die schönen Zeiten bedanken.
Mit der größte Dank gilt meiner Mutter, die sich wohl von allen am längsten meine Ideen zur Mathematik anhören musste. Ein sehr großes Dankeschön, dass sie mich während meines Lebensweg immer unterstützt hat und auch in schwierigen Zeiten immer für mich da war. Ich hab dich lieb, Mama. An dieser Stelle danke ich auch meinem

Stiefvater, der mich während meines Studiums immer unterstützt und motiviert hat.
Der aller größte Dank gilt einem ganz besonderem Menschen, Stefan Heuer, der nicht nur mein Partner sondern auch mein bester Freund ist. Er musste sich in den letzten Jahren wohl am aller häufigsten immer wieder verschiedenste verrückte Gedanken zu meiner Dissertation anhören. Er hat mir immer wieder Mut gemacht und war in den letzten Jahren immer für mich da. Ich freue mich so sehr auf zukünftige gemeinsame Zeiten. Ich liebe dich.
Zu aller letzt möchte ich mich bei meinen Familienmitglieder bzw. Freunden Arnold Boschert und seiner Frau Trixi, Heinrich, Karin, Peter, Tobi und Jennifer für das Zuhören und all die schönen Zeiten bedanken.

Contents

1. Introduction **3**

2. Option pricing **17**
2.1. Option pricing models 17
2.1.1. Options 17
2.1.2. Black-Scholes Model 19
2.1.3. Heston's stochastic volatility model 22
2.1.4. Uncertain volatility model (Black-Scholes-Barenblatt equation) . 24
2.1.5. Greeks 27
2.2. Valuation of an American option 29
2.2.1. Formulation as parabolic variational inequality with Black-Scholes model 29
2.2.2. Formulation as parabolic variational inequality with Heston's model 36

3. Background Material **46**
3.1. Well-posedness and regularity results for variational inequalities 46
3.1.1. Elliptic variational inequalities 46
3.1.2. Parabolic variational inequalities 51
3.2. B-Splines 60
3.3. A priori estimates for elliptic variational inequalities 63

4. Pricing American Put and Call option with Black-Scholes model **67**
4.1. Well-posedness and regularity 67
4.2. Discretization 73
4.2.1. Semi-discretization schemes in time 73
4.2.2. B-spline Galerkin discretization 75

5. Pricing American Put option with Heston's model **79**
5.1. Well-posedness and regularity 79
5.2. Discretization 84
5.2.1. Semi-discretization schemes in time 84
5.2.2. Tensor product B-spline Galerkin discretization 88

6. Approximation of the Greeks for American options **93**
6.1. Greeks in Black-Scholes Model 94
6.2. Greeks in Heston's Model 94

7. Solution algorithms to solve discrete variational inequalities **96**
7.1. Projected iterative methods 97

7.2. Monotone multigrid method 98
7.2.1. The basic monotone multigrid algorithm 98
7.2.2. Monotone coarse grid approximation for tensor-product splines . 100
7.2.3. Convergence theory 112

8. Numerical results for variational inequalities 116
8.1. Test problem 116
8.2. Pricing American options with Black-Scholes and Heston's model 121
8.2.1. Choice of the boundary conditions for Heston's model 123
8.2.2. Influence of coinciding knots on the approximated option value and Greeks 124
8.2.3. Influence of the time discretization 131
8.2.4. Influence of the B-spline order 135
8.2.5. Convergence of the monotone Multigrid method 138

9. Pricing European option with the Black-Scholes-Barenblatt equation 144
9.1. Uncertain volatility model 144
9.2. Well-posedness 146
9.3. Time semi-discretization and spatial B-spline discretization 151
9.4. Semismooth Newton method 157
9.4.1. Slant derivative 157
9.4.2. Semismooth Newton method within a knot insertion step 158
9.5. Numerical results 163

10. Conclusion and outlook 169
10.1. Conclusion 169
10.2. Outlook 170

A. Appendix 171

B. Code documentation 174

C. Symbols 184

Bibliography 193

2. Option pricing

2.1. Option pricing models

We begin with some basic definitions and an overview of option pricing models. A detailed introduction to option pricing problems can be found in [ABM09, Ach05, Hul09, Rou13, Sey12, WDH93], only to mention a few. Also, the theses of [Bud13, Bur16, Sen15, Wei14] are recommendable for a short introduction to option pricing problems.
The organization of this chapter is as follows. In Section 2.1.1 we briefly introduce some financial option pricing instruments and in Section 2.1.2, Section 2.1.3 and Section 2.1.4 we discuss a mathematical description of the Black-Scholes, Heston stochastic volatility and Uncertain volatility model. An overview of mathematical definitions of the Greeks and their meaning on a financial market is discussed in Section 2.1.5. Finally, we introduce the mathematical formulation of pricing an American option in the Black-Scholes or Heston's model in Section 2.2.

2.1.1. Options

First of all, we present the definition of an option.

Definition 2.1. *An option is a financial contract, which allows but not obligates the holder (buyer) of an option to trade an underlying asset S at a specific point in time T (European option) or during a period of time $[0, T]$ (American option). When the holder of the option makes use of his right he pays a fixed strike price K to the seller (writer). In return the holder receives a non-negative payoff $\mathcal{H}(S)$.*

An underlying asset typically is a stock, but also parcel of shares of a company, stock indices, currencies or commodities are possible. Numerous types of options such as *Vanilla*, Barrier, Bermudan, Exotic and Asian options etc. are traded on markets as well. The mentioned options all differ in their expiration dates, payoff structures or conditions on expiration. In this thesis we are going to concentrate on Vanilla options, these are European and American options as defined in Definition 2.1. In particular, we will focus on American options in the Black-Scholes and Heston's model. For the more complicated non-linear Black-Scholes-Barenblatt equation we are going to consider European options.
At an expiration date T the holder receives a so-called *payoff*. There are numerous types of payoff structures possible. We concentrate on call, put and butterfly-spread

options. If the holder of an option is allowed to buy an underlying asset S at a specific strike price K, the option is called *call option* and the payoff is defined as

$$\mathcal{H}_{\mathrm{C}}(S) := \max(0, S - K). \tag{2.1}$$

If the holder of an option is allowed to sell an underlying asset S at a specific strike price K, the option is called *put option* and the payoff is defined as

$$\mathcal{H}_{\mathrm{P}}(S) := \max(0, K - S). \tag{2.2}$$

A graphical illustration of a put and call option is presented in Figure 2.1. Options can be considered as an insurance against future price fluctuations and can be employed for hedging transactions. From the holder's point of view, a call option can be considered as a kind of insurance against future rising underlying prices and a put option as a insurance against falling underlying prices.

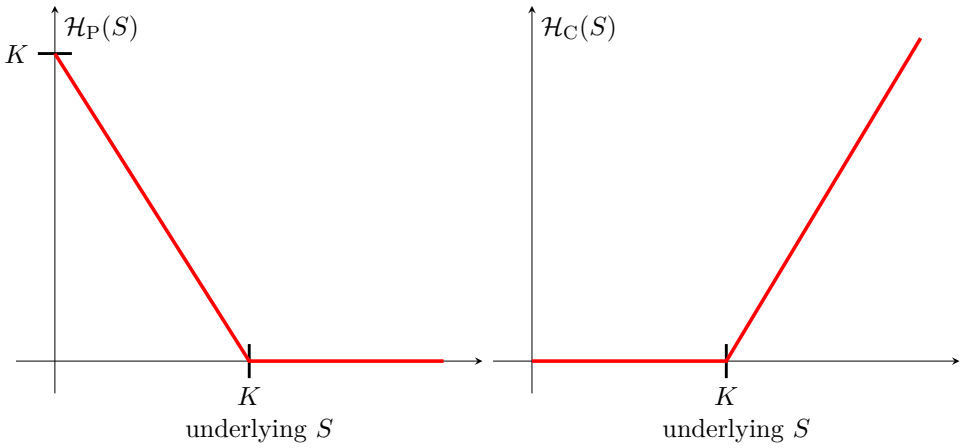

Figure 2.1.: Payoff of a put (left) and call (right) option out of the perspective of the holder.

There are also trading strategies possible where purchases and sales of put and call options are combined such as bull spread, bear spread, strangle and butterfly spread options (see for example [Sey12, p.72] or [Hul09, pp. 219-234]). We will only discuss the *butterfly-spread option*, since this strategy will be considered in the following chapters. A butterfly-spread option is a trading strategy, where positions with three different strike prices K_1, K and K_2 are taken. It can be constructed with call and put options, and both possibilities are leading to the same payoff function. A combination of call options consists of buying a call option at a strike price K_1, buying a call option at a higher strike price K_2 with $K_1 < K_2$ and selling two call options at strike price $K := (K_1 + K_2)/2$, where the strike price K is normally close to the current stock price. Due to the put-call parity (c.f. [Sey12, p. 5]), a butterfly spread can also be created by buying a put option at strike price K_1, buying a put option at higher strike price $K_2 > K_1$ and selling two put options at strike price K.

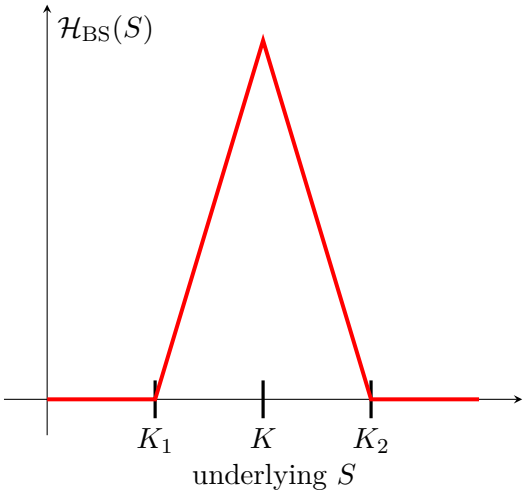

Figure 2.2.: Payoff of a butterfly-spread option out of the perspective of an investor.

At the expiration date T the value of an butterfly-spread option from the investor's point of view is given by the following payoff

$$\begin{aligned}\mathcal{H}_{\mathrm{BS}}(S) :&= \max(0, S-K_1) - 2\max(0, S-K) + \max(0, S-K_2) \\ &= \begin{cases} 0 & S < K_1,\ S > K_2 \\ S - K_1 & K_1 \leq S \leq K \\ K_2 - S & K \leq S \leq K_2. \end{cases} \end{aligned} \tag{2.3}$$

Graphically, the payoff of a butterfly-spread option can be presented as in Figure 2.2. As can be seen, a butterfly-spread option leads to a high profit if the stock price S is close to the strike price K, but the profit is increasing if the stock price tends to the strike prices K_1 or K_2. Therefore a butterfly-spread option is a recommendable strategy for an investor, who relies on barely changing stock prices around the strike price K. In other words, a butterfly-spread option leads to a limited high profit for an investor when the future volatility of the stock price is expected to be lower than the implied volatility.

As one can see, what all these payoff functions have in common is that they are piecewise linear. In particular, they are not differentiable when the underlying asset is equal to the strike price. This plays an important role for further numerical computations in this thesis.

2.1.2. Black-Scholes Model

In this section we briefly describe the derivation of the Black-Scholes equation. This partial differential equation of parabolic type was first introduced by [BS73]. There is a lot of literature about the derivation of the Black-Scholes equation available; we

follow here the approach of [Hul09, Wil01]. Regarding the valuation of options with the Black-Scholes model it is very important to consider the underlying assumption from which the resulting partial differential equation can be derived. The assumptions of the Black-Scholes model are not all empirically valid and it is necessary to understand its limitations for a proper application.

Assumption 2.2. *(Model assumptions)*
The assumptions on the assets are:

i) The underlying asset (or stock price) is modeled by a geometric Brownian motion

$$dS = (\mu - D_0)Sdt + \sigma SdW \tag{2.4}$$

with constant drift μ, constant volatility σ and Wiener process W. Assuming continuous dividend payments of the underlying asset S leads to a decrease of S in each time interval by the amount of $dD = D_0 Sdt$ with constant dividend yields D_0 (see [Sey12, p. 156]).

ii) The interest rate r on the riskless asset is constant.

The assumptions on the market are:

iii) There is no arbitrage opportunity on the market, that means it is not possible to make a riskless profit.

iii) It is possible to borrow and lend any amount of cash at the riskless interest rate r at any time.

iv) Everyone has the ability to buy and sell any amount of the stock (securities). This also includes short selling of stocks.

v) The market is frictionless, that means that the market is free of any transactions costs or taxes. Securities are perfectly divisible.

The assumptions can be relaxed by assuming the volatility or interest rate to be a deterministic function over time [Mer73]. Let $V(t, S)$ the price of an option. Applying Itô's formula results in

$$dV(t,S) = \left(\frac{\partial V(t,S)}{\partial t} + (\mu - D_0)S\frac{\partial V(t,S)}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V(t,S)}{\partial S^2}\right)dt + \left(\sigma S \frac{\partial V(t,S)}{\partial S}\right)dW. \tag{2.5}$$

As one can see, both SDE's in (2.4) and (2.5) are driven by the random terms dW. In the next step we eliminate the random terms to express the value of an option as a fully deterministic partial differential equation. To do so, a portfolio Π of Δ_1 stocks S and Δ_2 options $V(t, S)$ is constructed as follows

$$\Pi := \Delta_1 S + \Delta_2 V(t, S). \tag{2.6}$$

Therefore the change of the portfolio over the time interval dt results in

$$d\Pi = \Delta_1(dS + dD) + \Delta_2 dV(t, S), \tag{2.7}$$

where the term $\Delta_1 dS$ denotes the change of the underlying stock price S and $\Delta_1 dD(t)$ corresponds to the amount of dividend payments of the stock S over the time interval dt. To eliminate the random terms dW, Δ_1 and Δ_2 are choosen as

$$\Delta_1 := \frac{\partial V(t,S)}{\partial S} \qquad \Delta_2 := -1. \tag{2.8}$$

Inserting the SDE's (2.4) and (2.5) in $d\Pi$ provides

$$\begin{aligned} d\Pi =& \Delta_1(dS + dD) + \Delta_2 dV(t,S) \\ =& -\left(\frac{\partial V(t,S)}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V(t,S)}{\partial S^2}\right) dt + D_0 S \frac{\partial V(t,S)}{\partial S} dt. \end{aligned} \tag{2.9}$$

Since no random terms in the change of the portfolio $d\Pi$ over dt appears, the portfolio is riskless. Then the no-arbitrage-principle implies that the portfolio must earn the same rate of return as investments in other riskless securities. Therefore, the following formula is valid

$$d\Pi = r\Pi dt. \tag{2.10}$$

Finally, by applying (2.9) and (2.6) to (2.10), the following linear parabolic partial differential equation with a diffusion, convection and reaction term is derived

$$\frac{\partial V(t,S)}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V(t,S)}{\partial S^2} + (r - D_0) S \frac{\partial V(t,S)}{\partial S} - rV = 0. \tag{2.11}$$

The above equation is called *Black-Scholes* or *Black-Scholes-Merton* equation. To shorten the notation we define

$$\mathcal{L}^B V(t,S) := \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V(t,S)}{\partial S^2} + (r - D_0) S \frac{\partial V(t,S)}{\partial S} - rV, \tag{2.12}$$

and the equation (2.11) can be rewritten as

$$\frac{\partial V(t,S)}{\partial t} + \mathcal{L}^B V(t,S) = 0. \tag{2.13}$$

In the case of European options, an exact closed-form solution formula for the Black-Scholes equation for given terminal and boundary data has been established in [BS73]. This means that a European option can be priced by an easy formula. Due to its simple handling, the so-called Black-Scholes formula became very famous. In the case of an American put option or call option with dividend payments ($D_0 > 0$) the problem leads to a more complex free boundary problem, to which until today no exact solution formula is available and numerical computation is needed. The resulting problem for pricing an American put and call option with the Black-Scholes model is discussed in Section 2.2.1.

2.1.3. Heston's stochastic volatility model

In the following section we briefly describe the derivation of the Heston equation, which was first introduced by [Hes93] in 1993. The content of this section is also based on the following literature [AMST07, Bur16, Gat06, ABM09, Rou13].
One of the main limitations of the Black-Scholes model is the assumption on the constant volatility in the stock price process. Here the stock return is assumed to be normally distributed with a constant volatility violating most of the real market situations. An empirical determination of the volatility (*implied volatility*) from observed option prices for short maturities shows that the volatility in dependency of the strike prices can graphically be represented as a so-called *volatility smile* because when implied volatilities are plotted against the strike prices K, the resulting graph looks like a smile. The implied volatility is minimal when the strike price K is equal to the stock price S. For long maturities one often observe a volatility skew. One of the first models that was able to explain the volatility smile or skewness effect for different strike prices K was the Heston model. Another advantage of the Heston equation is that it provides closed-form solutions for a European call or put options for given terminal-boundary conditions. Due to this reasons the Heston model became very popular over the last decades (cf. [Rou13]).
To derive the Heston equation, it is required to make some assumptions on the asset and the market. The assumptions on the market are the same as in the Black-Scholes model, as provided in Assumption 2.2, but what is different is the stock price process, where the variance v, or volatility $\sqrt{v}$, is modeled as a CIR-process (*Cox-Ingersoll-Ross-process*)

$$dS = \mu S dt + \sqrt{v} S dW_1, \qquad dv = \kappa(\gamma - v)dt + \xi\sqrt{v}dW_2.$$

W_1 and W_2 are two different Wiener processes and they are correlated by a constant correlation $|\varrho| < 1$, $\kappa > 0$ denotes the mean reversion rate, $\gamma > 0$ is the mean reversion level and $\xi > 0$ is the volatility of the process. For the sake of simplicity, we also assume that the underlying asset does not pay dividends.
One important property of the CIR-process is that the variance remains always positive. To ensure that the volatility process is strictly positive it is often assumed in the literature that the parameters satisfy the following *Feller condition*

$$2\kappa\gamma > \xi^2, \tag{2.14}$$

otherwise the process degenerates to a deterministic function at the time when the volatility is zero. When calibrating the Heston model to real market option prices, the Feller condition can violate this condition, but in most experiments the parameters are selected so that the condition is satisfied.

In the next step we present the derivation of the Heston equation closely based on [Bur16, Gat06, Rou13].
We start by constructing a portfolio Π with one option $V := V(t, S, v)$, Δ_1 units of the underlying asset S and Δ_2 units of another underlying asset U. Then the value of the portfolio is given as follows

$$\Pi := V + \Delta_1 S + \Delta_2 U. \tag{2.15}$$

Therefore, the change of the portfolio over the time interval dt is given by

$$d\Pi = dV + \Delta_1 dS + \Delta_2 dU. \tag{2.16}$$

In the next step we apply the multidimensional version of Itô's formula to dV and dU ([KS91, Theorem 3.6]), which yields

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2}dt + \frac{1}{2}\xi^2 v\frac{\partial^2 V}{\partial v^2}dt + \xi\rho vS\frac{\partial^2 V}{\partial S\partial v}dt, \tag{2.17}$$

$$dU = \frac{\partial U}{\partial t}dt + \frac{\partial U}{\partial S}dS + \frac{\partial U}{\partial v}dv + \frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2}dt + \frac{1}{2}\xi^2 v\frac{\partial^2 U}{\partial v^2}dt + \xi\rho vS\frac{\partial^2 U}{\partial S\partial v}dt. \tag{2.18}$$

Inserting the SDE's (2.17) and (2.18) in $d\Pi$ provides

$$\begin{aligned} d\Pi = &\left(\frac{\partial V}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2 v\frac{\partial^2 V}{\partial v^2} + \rho\xi vS\frac{\partial^2 V}{\partial S\partial v}\right)dt \\ &+ \Delta_2\left(\frac{\partial U}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2} + \frac{1}{2}\xi^2 v\frac{\partial^2 U}{\partial v^2} + \xi\rho vS\frac{\partial^2 U}{\partial S\partial v}\right)dt \\ &+ \left(\frac{\partial V}{\partial S} + \Delta_2\frac{\partial U}{\partial S} + \Delta_1\right)dS + \left(\frac{\partial V}{\partial v} + \Delta_2\frac{\partial U}{\partial v}\right)dv. \end{aligned} \tag{2.19}$$

In the next step we eliminate the random terms dS and dv. Therefore we choose Δ_1 and Δ_2 as

$$\Delta_2 := -\frac{\frac{\partial V}{\partial v}}{\frac{\partial U}{\partial v}}, \qquad \Delta_1 = -\frac{\partial V}{\partial S} - \Delta_2\frac{\partial U}{\partial S}. \tag{2.20}$$

As in the Black-Scholes model the no arbitrage principle implies the formula $d\Pi = r\Pi dt$. Inserting $d\Pi$ from (2.19) and Π from (2.15) with Δ_1 and Δ_2 as above in $d\Pi = r\Pi dt$ yields

$$\begin{aligned} d\Pi = &\left(\frac{\partial V}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2 v\frac{\partial^2 V}{\partial v^2} + \rho\xi vS\frac{\partial^2 V}{\partial S\partial v}\right)dt \\ &+ \Delta_2\left(\frac{\partial U}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2} + \frac{1}{2}\xi^2 v\frac{\partial^2 U}{\partial v^2} + \xi\rho vS\frac{\partial^2 U}{\partial S\partial v}\right)dt \\ &= r\Pi dt = r(V + \Delta_1 S + \Delta_2 U). \end{aligned} \tag{2.21}$$

Rearranging leads to

$$\frac{\frac{\partial V}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2 v\frac{\partial^2 V}{\partial v^2} + \rho\xi vS\frac{\partial^2 V}{\partial S\partial v} + rS\frac{\partial V}{\partial S} - rV}{\frac{\partial V}{\partial v}}$$
$$= \frac{\frac{\partial U}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2} + \frac{1}{2}\xi^2 v\frac{\partial^2 U}{\partial v^2} + \xi\rho vS\frac{\partial^2 U}{\partial S\partial v} + rS\frac{\partial U}{\partial S} - rU}{\frac{\partial U}{\partial v}}. \tag{2.22}$$

Since the left-hand side is a function of V and the right-hand side is a function of U, equality is only possible if both sides are equal to a function. As suggested by [Hes93] the function is choosen as $f(t,S,v) = -\kappa(\gamma - v) + \lambda v$, where $\lambda(S,v,t) \geq 0$ represents the market price of volatility risk and is proportional to v, hence $\lambda(v) = \lambda v$. Then the *Heston partial differential equation* is given by

$$\frac{\partial V}{\partial t} + \mathcal{L}^H V(t,S,v) := \frac{\partial V}{\partial t} + \frac{1}{2}\left(S^2 v\frac{\partial^2 V}{\partial S^2} + 2\varrho\xi vS\frac{\partial^2 V}{\partial S\partial v} + \xi^2 v\frac{\partial^2 V}{\partial v^2}\right)$$
$$+rS\frac{\partial V}{\partial S} + (\kappa(\gamma - v) - \lambda(t,S,v))\frac{\partial V}{\partial v} - rV = 0. \tag{2.23}$$

The Heston equation is a parabolic PDE with a diffusion, a convection and a reaction term. In the case of a European call or put option, Heston [Hes93] proposed a closed-form solution for suitable terminal-boundary conditions. Since this closed-form solution includes complex integrals that can be only computed with quadrature rules, the solution is also called semi-analytical. In the case of an American put option, such a closed form solution is not available to date and a numerical discretization of a free boundary problem is needed. Regarding an American call option it can be shown that an early exercise decision makes no sense when the underlying asset pays no dividends. That is why the closed form solution for a European call option can also be used for an American call option without dividends. This result was first established by [Mer73] for the Black-Scholes model without dividends and it can also be applied for the Heston model. The derivation of the free boundary problem for the valuation of an American put option with Heston's model and its weak formulation as parabolic variational inequality can be found in Section 2.2.2.

2.1.4. Uncertain volatility model (Black-Scholes-Barenblatt equation)

In this section, we recall a simplified derivation of the Black-Scholes-Barenblatt equation from [Sey12]. The Black-Scholes-Barenblatt equation was first introduced by [ALP95, Lyo95]. In this model the volatility is a priori unknown and it is only assumed to lie within a range of extreme values. With this extreme values, the option value can be computed in a worst and best-case scenario. In the original paper of [ALP95], the Black-Scholes-Barenblatt equation is derived from a conditional expectation operator with a stochastic control variable $\sigma(t)$, a so-called stochastic control problem. To avoid the use of the theory for stochastic control problems, we follow the simplified arguments of [Sey12].
In the uncertain volatility model the assumptions are the same as in the Black-Scholes

model, but what is different is the stochastic process of the underlying asset. For simplicity we assume that the underlying asset pays no dividends. The underlying asset is assumed to follow the stochastic process

$$dS = \mu(t)Sdt + \sigma(t)SdW \tag{2.24}$$

where the stochastic variable $\sigma(t)$ lies in the range of

$$\sigma_{\min} \leq \sigma(t) \leq \sigma_{\max}. \tag{2.25}$$

Similar to the Black-Scholes equation we construct a portfolio Π with Δ_1 underlying assets S and Δ_2 options $V(t,S)$

$$\Pi := \Delta_1 S + \Delta_2 V(t,S). \tag{2.26}$$

Therefore, the change of the portfolio depending on the stochastic variable $\sigma(t)$, over the time interval dt is as follows

$$\begin{aligned} d\Pi(\sigma) =& \Delta_1 dS + \Delta_2 dV(t,S) \\ =& \Delta_1 \left(\mu(t)Sdt + \sigma(t)SdW\right) \\ &+ \Delta_2 \left(\frac{\partial V(t,S)}{\partial t} + \mu(t)S\frac{\partial V(t,S)}{\partial S} + \frac{1}{2}\sigma(t)^2 S^2 \frac{\partial^2 V(t,S)}{\partial S^2}\right) dt \\ &+ \Delta_2 \left(\sigma(t)S\frac{\partial V(t,S)}{\partial S}\right) dW. \end{aligned} \tag{2.27}$$

To eliminate the random terms dW we choose

$$\Delta_1 := -\frac{\partial V(t,S)}{\partial S}, \qquad \Delta_2 = 1, \tag{2.28}$$

that leads to

$$d\Pi(\sigma) = \left(\frac{\partial V(t,S)}{\partial t} + \frac{1}{2}\sigma(t)^2 S^2 \frac{\partial^2 V(t,S)}{\partial S^2}\right) dt \tag{2.29}$$

with the unknown stochastic variable $\sigma(t)$. Now the stochastic variable $\sigma(t)$ is choosen such that the return of $d\Pi(\sigma)$ increases by the maximal or minimal amount. Mathematically, this can be formulated as follows:

- Choose $\underline{\sigma}$ such that $d\Pi(\underline{\sigma})$ is the greatest lower bound. *(Worst case scenario)*
- Choose $\overline{\sigma}$ such that $d\Pi(\overline{\sigma})$ is the least upper bound. *(Best case scenario)*

Finally, applying the no arbitrage principle to the two cases above implies

$$\inf_{\underline{\sigma}\in\Sigma} \left(d\Pi(\underline{\sigma})\right) = r\Pi dt, \qquad \textit{(Worst case scenario)} \tag{2.30}$$

$$\sup_{\overline{\sigma}\in\Sigma} \left(d\Pi(\overline{\sigma})\right) = r\Pi dt \qquad \textit{(Best case scenario)} \tag{2.31}$$

with $\Sigma := [\sigma_{\min}, \sigma_{\max}]$. Exploiting (2.26) and (2.27) and rearranging of the two equations above leads to the following partial differential equations of Hamilton-Jacobi-Bellman type

$$\frac{\partial V(t,S)}{\partial t} + \inf_{\underline{\sigma}\in\Sigma}\left(\frac{1}{2}\underline{\sigma}^2 S^2 \frac{\partial^2 V(t,S)}{\partial S^2}\right) + rS\frac{\partial V(t,S)}{\partial S} - rV = 0 \tag{2.32}$$

and

$$\frac{\partial V(t,S)}{\partial t} + \sup_{\overline{\sigma}\in\Sigma}\left(\frac{1}{2}\overline{\sigma}^2 S^2 \frac{\partial^2 V(t,S)}{\partial S^2}\right) + rS\frac{\partial V(t,S)}{\partial S} - rV = 0. \tag{2.33}$$

These two equations are known as *Black-Scholes-Barenblatt equation* [ALP95, Lyo95]. Assuming that the asset pays out a dividend D_0, the Black-Scholes-Barenblatt equation can be rewritten in the following compact form for the worst case scenario

$$\frac{\partial V(t,S)}{\partial t} + \inf_{\underline{\sigma}\in\Sigma}\left(\mathcal{L}^{\underline{\sigma}} V(t,S)\right) = 0 \tag{2.34}$$

and for the best case scenario we have

$$\frac{\partial V(t,S)}{\partial t} + \sup_{\overline{\sigma}\in\Sigma}\left(\mathcal{L}^{\overline{\sigma}} V(t,S)\right) = 0, \tag{2.35}$$

where the linear differential operator $\mathcal{L}^{\sigma} V(t,S)$ is defined as

$$\mathcal{L}^{\sigma} V(t,S) := \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V(t,S)}{\partial S^2} + (r - D_0) S \frac{\partial V(t,S)}{\partial S} - rV. \tag{2.36}$$

In the case of European options, the two extreme option values can be obtained by solving the equations above for a given terminal condition $V(T,S) = \mathcal{H}(S)$ with payoff $\mathcal{H}(S)$. In this setting $V(t,S)$ corresponds to the costs of dynamic hedging with the underlying asset under the worst case volatility path and solves (2.34) with

$$\underline{\sigma}^{-}\left[V_{SS}\right] := \begin{cases} \sigma_{\min}^2 & \text{if } V_{SS} \geq 0, \\ \sigma_{\max}^2 & \text{if } V_{SS} < 0, \end{cases} \tag{2.37}$$

and (2.35) the best-case scenario with

$$\overline{\sigma}^{+}\left[V_{SS}\right] := \begin{cases} \sigma_{\max}^2 & \text{if } V_{SS} \geq 0, \\ \sigma_{\min}^2 & \text{if } V_{SS} < 0. \end{cases} \tag{2.38}$$

Since the volatility function $\overline{\sigma}^{+}\left[V_{SS}\right]$ or $\underline{\sigma}^{-}\left[V_{SS}\right]$ depends on the second derivatives of the solution for each $t \in [0,T)$, the Black-Scholes-Barenblatt equation is a non-linear partial differential equation. For convex payoff functions, as for a put or call option, this problem reduces to the Black-Scholes equation with one of the extreme values of the volatility. In the simple case of a European option for initial condition given as

a non-convex payoff function, no closed-form solution of the Black-Scholes-Barenblatt equation is available and numerical experiments for non-linear PDE's are needed.

2.1.5. Greeks

To determine optimal risk strategies in finance, the focus of interest is not only in the solution of the partial differential equation or of the free boundary problem (i.e. the European or American option price), but also on its partial derivatives up to order two, the so-called *Greeks*. The name is used since most of these sensitivities are denoted in Greek letters. One of the main objectives of this thesis is to find an pointwise approximation with cubic (tensor product) B-splines of these sensitivities, in particular for the more challenging case of an American option. In this section we give a short overview of the most relevant Greeks and their meaning on a financial market. The content of the next section is based on [Hul09, p.247, pp. 349–376]. A short introduction can also be found in [Bos15, Wei14].
The *Delta* of an option measures the change of the option value V with respect to a small change of the underlying asset S. Mathematically, Delta is the partial derivative of the option value with respect to the underlying asset S

$$\Delta := \frac{\partial V}{\partial S} \tag{2.39}$$

and it satisfies $0 \leq \Delta \leq 1$ for a call option and $-1 \leq \Delta \leq 0$ for a put option. Delta is the amount of underlying assets the owner of one option should buy or sell to create a riskless portfolio. This is also known as Delta-hedging. Adjusting the amount of sales or purchases of the underlying asset such that the Delta of the portfolio sums to zero is called delta neutral. Since Delta is changing over time, the owner of the portfolio remains delta neutral for only a relatively small period of time and the hedge has to be adjusted for each period. In order to control the adjustments to remain delta neutral, the sensitivity *Gamma* is highly relevant. Gamma is the change of Delta Δ with respect to a small change of the underlying asset S. Formally speaking, Gamma is the second partial derivative with respect to the underlying price S

$$\Gamma := \frac{\partial^2 V}{\partial S^2} \tag{2.40}$$

and is positive for a put and call option. If the absolute value of Gamma is small, Delta changes slowly, and the hedge has to adjust infrequently to remain delta neutral. Otherwise, if the absolute value of Γ is high, Delta is very sensitive to a small change of the underlying asset, and the hedge has to adjust frequently to remain delta neutral. As discussed earlier, the option value depends also on the volatility. Therefore, the sensitivity *Vega* measuring the change of the option value to a small change of the volatility $\sqrt{v}$, is highly relevant. In Heston's model, where the volatility is not constant, Vega can be obtained by the first partial derivative of the option value with respect

to the volatility $\sqrt{v}$. For the sake of simplicity the partial derivative is often computed with respect to the variance v

$$\nu := \frac{\partial V}{\partial v}. \tag{2.41}$$

For European and American vanilla options, Vega is always non-negative, $\nu \geq 0$. If the value of Vega is high, volatility changes have a great impact on the option value. Otherwise, if the value of Vega is low, the option value is not sensitive to a small change of the volatility. In order to manage the volatility risk, it is also possible to hedge a portfolio against the volatility of the underlying asset. Therefore, a portfolio with an Vega of ν_Π can be made vega neutral by including $-\nu_\Pi/\nu$ of one option with an Vega of ν. Explicit examples for hedging strategies can be found in [Hul09, pp. 349–367]

2.2. Valuation of an American option

2.2.1. Formulation as parabolic variational inequality with Black-Scholes model

In this section we consider the valuation of an American put and call option in the *Black-Scholes model*. The option price $V := V(t, S)$ is assumed to depend on an underlying asset $S \in \mathbb{R}^+$, time t, volatility σ, interest rate r, dividend yields D_0 and is subject to a strike price K. In the case of a put option the payoff is given by (2.2) and in the case of call option by (2.1).
In comparison to a European option, a fundamental difficulty arises from the fact that American option can be exercised during a time period $[0, T]$ and not only at a specific end time T. Since the optimal underlying price $S_f := S_f(t) : [0, T] \to \mathbb{R}^+$ depending on time $t \in [0, T]$, such that exercising the option is optimal, is a priori not known, the valuation of an American option leads to a *free boundary problem*. Usually, a log-transformation $x = \log(S/K)$ is introduced to eliminate the degenerate part in S transforming the Black-Scholes equation to a non-degenerate equation on the whole axis of $\mathbb{R}$. For the numerical computation the domain is truncated such that $x_{\min} < x < x_{\max}$ with $x_{\min} < 0$ and $x_{\max} > 0$ and artificial boundary conditions are set. For an American put option the solution of the free boundary problem is given by the obstacle $\mathcal{H}_{\mathrm{P}}(S)$ for all $S \leq S_f(t)$, such that the truncation on the left interval, for small enough $x_{\min}$ and standard parameter r, σ and K used in the literature, has no truncation effect to the numerical solution. Since for an American call option the solution is modeled by the Black-Scholes equation for all $S \leq S_f(t)$, truncation on the left interval has a direct influence on the numerical solution and leads to truncation errors. As a result, we present a weak formulation as a *parabolic variational inequality in a weighted Sobolev space* on a bounded domain without changing the solution, which has been introduced in [AP05, DL19a, DL]. This allows us to compute an accurate numerical approximation of the American call option. Therefore, it is assumed that the coefficients of the Black-Scholes equation in (2.11) satisfy the following assumptions

$$\sigma > 0, \quad r > 0 \text{ and } D_0 \geq 0, \tag{2.42}$$

with $D_0 \geq 0$ for an American put option and $D_0 > 0$ for an American call option.

The free boundary problem for pricing an American option and the reformulation as a *linear complementarity problem* is well known and can be found in [WDH93].

Valuation of an American put option

First, we consider the valuation of an American put option with payoff $\mathcal{H}_{\mathrm{P}}$ as defined in (2.2). Let S_f be the unique a priori not known underlying, such that for all $S \leq S_f \leq K$ exercising the option is optimal and the holder will receive a positive profit $\mathcal{H}_{\mathrm{P}}(S) > 0$, whereas for all $S > S_f$ holding the option and speculating on future falling underlying prices is a better strategy. For $t = T$ the optimal underlying price is $S_f(T) = K$. Based

on this consideration, the option price $V(t,S)$ with $(t,S) \in [0,T) \times (0,\infty)$ is given by the payoff function

$$V(t,S) = \mathcal{H}_{\mathrm{P}}(S) \qquad \text{for all } S \leq S_f(t) \tag{2.43}$$

and is modeled by the Black-Scholes equation

$$\frac{\partial V(t,S)}{\partial t} + \mathcal{L}^B V(t,S) = 0 \qquad \text{for all } S > S_f(t) \tag{2.44}$$

with $\mathcal{L}^B$ defined in (2.12). At the end of the time period, there is no other possibility for the holder than to make use of his right or let the option expire. Thus, the option value at its expiration date $t = T$ is given by the following end condition

$$V(T,S) = \mathcal{H}_{\mathrm{P}}(S). \tag{2.45}$$

In order to find a discrete formulation of the free boundary problem we truncate the infinite domain $\mathbb{R}^+$ to $I := (0, S_{\max})$, choosing $S_{\max}$ such that $0 < K < S_{\max} < \infty$. If $S_{\max}$ is large enough the Dirichlet boundary condition is given by

$$V(t,S_{\max}) = \mathcal{H}_{\mathrm{P}}(S_{\max}) = 0, \tag{2.46}$$

since as for a large underlying price the American put option becomes worthless. In summary we can formulate the following free boundary problem in a finite domain:

Problem 2.3. *(Localized free boundary problem – Valuation of American put option) Find $V(t,S)$ with $(t,S) \in [0,T) \times I$ such that*

$$V = \mathcal{H}_P(S) \quad \textit{for } S \leq S_f(t), \tag{2.47}$$

$$\frac{\partial V}{\partial t} + \mathcal{L}^B V = 0 \quad \textit{for } S > S_f(t) \tag{2.48}$$

with boundary and end condition

$$V(t,S_{\max}) = \mathcal{H}_P(S_{\max}), \qquad V(T,S) = \mathcal{H}_P(S), \tag{2.49}$$

where the payoff function is given by $\mathcal{H}_P(S) := \max\{0, K - S\}$.

The free boundary for the valuation of an American put option lies in the range

$$0 < S_f(t) < K \quad \text{for a.e. } \tau \in (0,T] \tag{2.50}$$

(see e.g. [Ach05]) and is a continuous monotonically decreasing function for $t \to 0$.

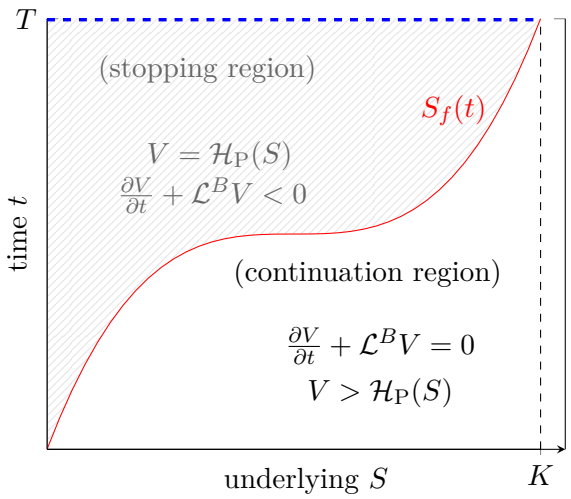

Figure 2.3.: Free boundary problem for the valuation of American put option with the Black-Scholes model. The figure is reproduced from [Sey09].

In order to obtain a formulation where the a priori unknown free boundary does not explicitly appear, the problem is typically reformulated as a linear complementarity problem (c.f. [WDH93]). Therefore, (2.47) implies that the option value is given by the payoff function $V(t,S) = \mathcal{H}_P(S) = K - S$ for $0 \leq S \leq S_f < K$, applying the Black-Scholes equation to the payoff leads to

$$\begin{aligned}\frac{\partial V}{\partial t} + \mathcal{L}^B V &= \frac{\partial (K-S)}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 (K-S)}{\partial S^2} + (r - D_0) S \frac{\partial (K-S)}{\partial S} - r(K-S) \\ &= -D_0 S - rK \leq -(D_0 + r)K < 0 \quad \text{for } D_0 \geq 0, r > 0. \end{aligned} \tag{2.51}$$

For $S > S_f$ the option value must be greater than the payoff function

$$V(t,S) > \mathcal{H}_P(S) \tag{2.52}$$

because otherwise the holder can immediately exercise the option and receives a positive profit, which contradicts the no arbitrage principle. A graphical illustration of this formulation can be found in Figure 2.3. In summary, together with (2.47) and (2.48), the following linear complementarity problem is derived:

Problem 2.4. *(Complementarity problem – Valuation of American put option) Find $V(t,S)$ with $(t,S) \in [0,T) \times I$ such that*

$$\begin{aligned}\left(\frac{\partial V}{\partial t} + \mathcal{L}^B V\right)(V - \mathcal{H}_P) &= 0 \\ \frac{\partial V}{\partial t} + \mathcal{L}^B V &\leq 0 \\ V - \mathcal{H}_P &\geq 0\end{aligned}$$

and $V(t, S_{\max}) = \mathcal{H}_P(S_{\max})$, $V(T,S) = \mathcal{H}_P(S)$, where the payoff function is given by $\mathcal{H}_P(S) := \max\{0, K - S\}$.

To simplify further theoretical considerations, we introduce the time transformation $\tau := T - t$, hence we have the following *time transformed option price* and *free boundary*

$$y(\tau, S) := V(T - t, S), \qquad \widehat{S}_f(\tau) := S_f(T - t). \tag{2.53}$$

Here τ can be interpreted as the time until the option expires. The transformation leads to the following *time transformed Black-Scholes equation*

$$\frac{\partial y}{\partial \tau} - \mathcal{L}^B y = \frac{\partial y}{\partial \tau} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 y}{\partial S^2} + (D_0 - r)S\frac{\partial y}{\partial S} + ry = 0. \tag{2.54}$$

To this end, the time transformed solution satisfies the following complementarity problem:

Problem 2.5. *(Time transformed complementarity problem – put option)*
Find $y(\tau, S)$ with $(\tau, S) \in (0, T] \times I$ such that

$$\left(\frac{\partial y}{\partial \tau} - \mathcal{L}^B y\right)(y - \mathcal{H}_P) = 0 \tag{2.55}$$

$$\frac{\partial y}{\partial \tau} - \mathcal{L}^B y \geq 0 \tag{2.56}$$

$$y - \mathcal{H}_P \geq 0 \tag{2.57}$$

and $y(0, S) = \mathcal{H}_P(S)$ and $y(\tau, S_{\max}) = 0$, where the payoff function is given by $\mathcal{H}_P(S) := \max\{0, K - S\}$.

Valuation of an American call option

Now, we present the valuation of an American call option with payoff function $\mathcal{H}_C(S)$ as defined in (2.1) on a continuous dividend paying asset. In order to induce an early exercised decision of the holder, the value of an American call option must be deeper in the money. An option is said to be deep in the money when the corresponding payoff is strictly positive. Hence, the holder will exercise the American call option for all $S \geq S_f(t) > K$, whereas for all $S < S_f(t)$ the continuation of the option and speculating on future increasing underlying prices will be a better strategy. Exploiting these arguments the value of an American call option $V(t, S)$ with $(t, S) \in [0, T) \times (0, \infty)$ is given by

$$V(t, S) = \mathcal{H}_C(S) \quad \text{for all } S \geq S_f(t) \tag{2.58}$$

and is modelled by the Black-Scholes equation

$$\frac{\partial V}{\partial t} + \mathcal{L}^B V(t, S) = 0 \quad \text{for all } S < S_f(t). \tag{2.59}$$

At the end of the period the holder will receive the payoff, such that the end condition is given by the payoff

$$V(T, S) = \mathcal{H}_C(S). \tag{2.60}$$

For the numerical computation, we truncate the domain $(0, \infty)$ to a finite interval $(0, S_{\max})$. It can be shown (see [DL]) that for large $S_{\max}$, more precisely,

$$S_{\max} = \frac{\mathfrak{a}}{\mathfrak{a}-1} K, \ \mathfrak{a} > 1 \tag{2.61}$$

the Dirichlet boundary condition $V(t, S_{\max}) = \mathcal{H}_{\mathrm{C}}(S_{\max})$ is an exact condition. Finally, the following free boundary problem can be formulated for the valuation of an American call option:

Problem 2.6. *(Localized free boundary problem – Valuation of American call option) Find $V(t, S)$ with $(t, S) \in [0, T) \times I$ such that*

$$V = \mathcal{H}_C(S) \quad \text{for } S \geq S_f(t), \tag{2.62}$$

$$\frac{\partial V}{\partial t} + \mathcal{L}^B V = 0 \quad \text{for } S < S_f(t) \tag{2.63}$$

with boundary and end condition

$$V(t, S_{\max}) = \mathcal{H}_C(S_{\max}), \qquad V(T, S) = \mathcal{H}_C(S), \tag{2.64}$$

where the payoff function is given by $\mathcal{H}_C(S) := \max\{0, S - K\}$.

To formulate the free boundary problem as complementarity problem for an American call option, we apply the Black-Scholes operator to the payoff function $\mathcal{H}_{\mathrm{C}}(S) := \max\{0, S - K\}$ for $S \geq S_f(t)$

$$\begin{aligned} \frac{\partial V}{\partial t} + \mathcal{L}^B V &= \frac{\partial (S-K)}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 (S-K)}{\partial S^2} + (r - D_0) S \frac{\partial (S-K)}{\partial S} - r(S-K) \\ &= -D_0 S + rK \leq -(D_0 + r)K < 0 \end{aligned} \tag{2.65}$$

for $D_0 > 0, r > 0$ and $S \geq S_f(t) > S_f(T) = K \max\{1, \frac{r}{q}\}$. Due to the no arbitrage principle we have

$$V(t, S) > \mathcal{H}_{\mathrm{C}}(S). \tag{2.66}$$

The two equations above, along with (2.62) and (2.63), imply the following complementarity problem for an American call option:

Problem 2.7. *(Complementarity problem – Valuation of American call option) Find $V(t, S)$ with $(t, S) \in [0, T) \times I$ such that*

$$\begin{aligned} \left(\frac{\partial V}{\partial t} + \mathcal{L}^B V \right) (V - \mathcal{H}_C) &= 0 \\ \frac{\partial V}{\partial t} + \mathcal{L}^B V &\leq 0 \\ V - \mathcal{H}_C &\geq 0 \end{aligned}$$

and $V(t, S_{\max}) = 0$, $V(T, S) = \mathcal{H}_C(S)$, where the payoff function is given by $\mathcal{H}_C(S) := \max\{0, S - K\}$.

Finally, a transformation in time as described in (2.53) leads to the following time transformed problem:

Problem 2.8. *(Time transformed complementarity problem – American call option) Find $y(\tau, S)$ with $(\tau, S) \in (0, T] \times I$ such that*

$$\left(\frac{\partial y}{\partial \tau} - \mathcal{L}^B y\right)(y - \mathcal{H}_C) = 0 \tag{2.67}$$

$$\frac{\partial y}{\partial \tau} - \mathcal{L}^B y \geq 0 \tag{2.68}$$

$$y - \mathcal{H}_C \geq 0 \tag{2.69}$$

and $y(0, S) = \mathcal{H}_C(S)$ and $y(\tau, S_{\max}) = 0$, where the payoff function is given by $\mathcal{H}_C(S) := \max\{0, S - K\}$.

Formulation as a parabolic variational inequality

Now, we pose Problem 2.5 for an American put option and Problem 2.8 for an American call option in variational form. Since the complementarity problem for an American call and put option only differs in the payoff $\mathcal{H}_{\mathrm{P}}(S)$ or $\mathcal{H}_{\mathrm{C}}(S)$, the derivation of the variational formulation is the same. To simplify the notation, we will write $\mathcal{H}(S)$ for both type of options. We denote by $H := L^2(I)$ the space of square integrable functions and we define a weighted Sobolev space by

$$\mathcal{V} := \left\{\varphi \in H : S\frac{\partial \varphi}{\partial S} \in H\right\} \tag{2.70}$$

equipped with the norm

$$\|\varphi\|_{\mathcal{V}}^2 := \int_0^{S_{\max}} \left(\varphi^2 + \left(S\frac{\partial \varphi}{\partial S}\right)^2\right) \mathrm{d}S. \tag{2.71}$$

Let

$$\mathcal{V}^0 := \left\{\varphi \in H : S\frac{\partial \varphi}{\partial S} \in H, \varphi(S_{\max}) = 0\right\} \tag{2.72}$$

the space with zero boundary condition. We denote by $\mathcal{V}^*$ the dual space of $\mathcal{V}^0$. The embeddings $\mathcal{V}^0 \subset H = H^* \subset \mathcal{V}^*$ are continuous and dense (c.f. [AP05, p.30]), the inner product $(\cdot, \cdot)$ of H is used to define the duality pairing between $\mathcal{V}^*$ and $\mathcal{V}$. A norm of the dual space $\mathcal{V}^*$ is given by the operator norm

$$\|f\|_{\mathcal{V}^*} := \sup_{\eta \in \mathcal{V}^0 \setminus 0} \frac{(f, \eta)}{\|\eta\|_{\mathcal{V}}}. \tag{2.73}$$

We also need the following convex and closed set

$$\mathcal{K} := \{\varphi \in \mathcal{V} : \varphi \geq \mathcal{H} \text{ in } I,\ \varphi(S_{\max}) = \mathcal{H}(S_{\max})\}. \tag{2.74}$$

For the derivation of the variational inequality, let $\varphi \in \mathcal{K}$, which ensures the positivity of $\varphi - \mathcal{H} \geq 0$. Multiplying (2.56) in the case of a put option (or (2.68) in the case of a call option) with $\varphi - \mathcal{H} \geq 0$ and integrating over I leads to

$$\int_I \left(\frac{\partial y}{\partial \tau} - \mathcal{L}^B y \right) (\varphi - \mathcal{H}) \mathrm{d}S \geq 0 \quad \text{and} \quad \int_I \left(\frac{\partial y}{\partial \tau} - \mathcal{L}^B y \right) (y - \mathcal{H}) \mathrm{d}S = 0.$$

Subtraction results in

$$\int_I \left(\frac{\partial y}{\partial \tau} - \mathcal{L}^B y \right) (\varphi - y) \mathrm{d}S \geq 0. \tag{2.75}$$

Finally, integration by parts for the diffusion term and $y \in \mathcal{K}$,

$$\begin{aligned} &\int_I \left(-\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 y}{\partial S^2} \right) (\varphi - y) \mathrm{d}S \\ &= \int_I \left(\frac{\sigma^2}{2} S^2 \frac{\partial y}{\partial S} \frac{\partial (\varphi - y)}{\partial S} \right) + \sigma^2 S \frac{\partial y}{\partial S} (\varphi - y) \mathrm{d}S, \end{aligned} \tag{2.76}$$

leads to a parabolic variational inequality. This formulation can also be found in [AP05, p. 187-188] or [DL19a, DL].

Problem 2.9. *(Variational inequality with inhomogeneous boundary conditions) Find $y \in L^2(0,T;\mathcal{K})$ such that $\frac{\partial y}{\partial \tau} \in L^2(0,T;H)$,*

$$\left(\frac{\partial y}{\partial \tau}, \varphi - y \right) + a^B(y, \varphi - y) \geq 0 \quad \textit{for all } \varphi \in \mathcal{K} \tag{2.77}$$

and $y(0) = \mathcal{H}(S)$, where $a^B(\cdot,\cdot) := \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ defines the bilinear form

$$a^B(y, \varphi) := a_0^B(y, \varphi) - a_1^B(y, \varphi) \tag{2.78}$$

with

$$a_0^B(y, \varphi) := \int_I \left(\frac{\sigma^2}{2} S^2 \frac{\partial y}{\partial S} \frac{\partial \varphi}{\partial S} \right) \mathrm{d}S \tag{2.79}$$

and

$$a_1^B(y, \varphi) := -\int_I ry\varphi + \left(\sigma^2 + D_0 - r \right) S \frac{\partial y}{\partial S} \varphi \mathrm{d}S. \tag{2.80}$$

The payoff is given by $\mathcal{H}(S) := \mathcal{H}_P(S)$ for a put option or $\mathcal{H}(S) := \mathcal{H}_C(S)$ for a call option.

For subsequent numerical computation, it is convenient to transform Problem 2.9 into a variational inequality with homogeneous Dirichlet boundary condition. Observe that $\mathcal{H}(S) \in \mathcal{V}$ yields $\mathcal{K} = \mathcal{H} + \mathcal{K}^0$, where the convex set $\mathcal{K}^0$ is defined as

$$\mathcal{K}^0 := \{ \varphi \in \mathcal{V}^0 : \varphi \geq 0 \text{ in } \Omega \}. \tag{2.81}$$

Setting $y = u + \mathcal{H}$ leads to the following parabolic variational inequality with homogeneous boundary condition:

Problem 2.10. *(Variational inequality with homogeneous boundary conditions) Find $u \in L^2(0,T;\mathcal{K}^0)$ such that $\frac{\partial u}{\partial \tau} \in L^2(0,T;H)$,*

$$\left(\frac{\partial u}{\partial \tau}, \varphi - u\right) + a^B(u + \mathcal{H}, \varphi - u) \geq 0 \quad \text{for all } \varphi \in \mathcal{K}^0 \tag{2.82}$$

and $u(0) = 0$, where $a^B(\cdot,\cdot)$ is defined as in (2.78). The payoff is given by $\mathcal{H}(S) := \mathcal{H}_P(S)$ for a put option or $\mathcal{H}(S) := \mathcal{H}_C(S)$ for a call option.

2.2.2. Formulation as parabolic variational inequality with Heston's model

In this section we consider the valuation of an American put option in *Heston's model*. Therefore, similar considerations as for the valuation of an American option in the Black-Scholes model are used to derive a free boundary problem and the corresponding complementarity problem. Similar results can also be found for instance in [Bos15, Bur16, IT09, KSW12, Oos03, Wei14].
Since many different boundary conditions for the valuation of an American put option in Heston's model are used in the literature, we present an short overview of the boundary conditions. For simplicity, we also use a log-transformation of the underlying price to avoid a degenerate coefficient in S. To this end, we derive a weak formulation for the valuation of an American put option in Heston's model. The derivation of the weak form in this thesis differs from [Bos15, Bur16, KSW12, Wei14] due to an error, as indicated below.
As discussed before, the option price $V(t,S,v)$ in Heston's model is assumed to depend on time $t \in [0,T)$, underlying price $S \in \mathbb{R}^+$ and variance $v \in \mathbb{R}^+$. To avoid a degenerate coefficient in v it is assumed that the *Feller*-condition

$$2\kappa\gamma > \xi^2 \text{ and } v \geq v_{\min} > 0 \tag{2.83}$$

is satisfied such that the variance is strictly positive. It is also assumed that the coefficients satisfy the following assumptions

$$|\rho| < 1, \quad r > 0. \tag{2.84}$$

As described before, American options can be exercised at any time $t \in [0,T]$. The a priori unknown underlying price to which exercising the option is a better strategy depends on the variance v and time t. Due to this fact, the optimal exercising price in Heston's model is mathematically expressed as a surface, referred to as the free boundary, $S_f(t,v) : [0,T] \times \mathbb{R}_+ \to \mathbb{R}_+$. Since for all $S \leq S_f(t,v)$ exercising the option is a better strategy, the option value is given by

$$V(t,S,v) = \mathcal{H}_{\mathrm{P}}(S) \;\; \text{for all } S \leq S_f(t,v). \tag{2.85}$$

For all $S > S_f(t,v)$ speculating on future falling underlying prices and holding the option is a better strategy, therefore the option value is modeled by the Heston equation

$$\frac{\partial V}{\partial t} + \mathcal{L}^H V = 0 \quad \text{for all } S > S_f(t,v). \tag{2.86}$$

Several variations of boundary conditions for the valuation of an American option in Heston's model can be found in the literature. For a detailed summary of the boundary condition used in the literature, we also refer to [Bur16, pp. 24–26]. We start by introducing the boundary conditions in the S-direction derived from some financial considerations. For small underlying prices $S \to 0$, the option is at the money and the holder of the option will receive a positive payoff $\mathcal{H}_\mathrm{P}(S)$. As a result, the following boundary condition is used in the literature [CP99, Oos03, IT09, KSW12]

$$\lim_{S\to 0} V(t,S,v) = \lim_{S\to 0} \mathcal{H}_\mathrm{P}(S) = K. \tag{2.87}$$

In the case of high underlying prices $S \to \infty$, the option is worthless and the holder will let the selling right expire. This is modeled by the following Dirichlet boundary condition in [CP99, KSW12, Che11]

$$\lim_{S\to \infty} V(t,S,v) = \lim_{S\to \infty} \mathcal{H}_\mathrm{P}(S) = 0. \tag{2.88}$$

The choice of the boundary condition for $v \to 0$ is still subject to discussion. It is well known, that if Feller's condition $2\kappa\gamma > \xi^2$ is fulfilled, the process of the volatility will not attain the boundary $v = 0$ (see e.g. [AMST07]). An extensive discussion from the financial and the mathematical point of view in [Che11, ZC11] supports that for a small variance $v \to 0$ the option price behaves almost deterministic and the price is equal to the payoff $\mathcal{H}_\mathrm{P}(S)$. The same boundary condition can be found in [Oos03, KSW12, Che11]. Since in this thesis it is assumed that $v \geq v_\mathrm{min} > 0$, we write $v \to v_\mathrm{min}$ instead of $v \to 0$

$$\lim_{v\to v_\mathrm{min}} V(t,S,v) = \mathcal{H}_\mathrm{P}(S). \tag{2.89}$$

In [CP99] the authors argue that due to the degeneracy for $v \to 0$ the Heston equation becomes increasingly hyperbolic and imposing a Dirichlet boundary condition is not appropriate. They recommend a Neumann boundary condition for $v \to 0$

$$\lim_{v\to v_\mathrm{min}} \frac{\partial V(t,S,v)}{\partial v} = 0. \tag{2.90}$$

It was found out that the partial derivative of the solution with respect to the variance, Vega, is stable when implemented as a one-sided difference quotient. We make the same observation for a cubic B-spline discretization in Section 8.2.1.
At least it is known that for increasing volatility the fluctuations of the option price decrease. This is modelled by the following Neumann boundary condition in [CP99, Oos03, IT09, Che11, KSW12]

$$\lim_{v\to \infty} \frac{\partial V(t,S,v)}{\partial v} = 0. \tag{2.91}$$

Note, that in [Oos03, IT09], instead of (2.88) a Neumann boundary condition is used. In [FLMN11] the authors use

$$\lim_{v\to\infty} V(t,S,v) = \mathcal{H}_{\mathrm{P}}(S). \tag{2.92}$$

instead of (2.91), but they do not mention why.
In this thesis, we study two different boundary conditions for $v \to v_{\min}$. The boundary conditions under consideration are

$$\begin{aligned} &\lim_{S\to 0} V(t,S,v) = K, \quad \lim_{v\to v_{\min}} V(t,S,v) = \mathcal{H}_{\mathrm{P}}(S), \\ &\lim_{S\to\infty} V(t,S,v) = 0, \quad \lim_{v\to\infty} \frac{\partial V(t,S,v)}{\partial v} = 0 \end{aligned} \tag{2.93}$$

and

$$\begin{aligned} &\lim_{S\to 0} V(t,S,v) = K, \quad \lim_{v\to v_{\min}} \frac{\partial V(t,S,v)}{\partial v} = 0, \\ &\lim_{S\to\infty} V(t,S,v) = 0, \quad \lim_{v\to\infty} \frac{\partial V(t,S,v)}{\partial v} = 0. \end{aligned} \tag{2.94}$$

Note that the difference just lies in the boundary condition for $v \to v_{\min}$. In particular, we will show in Section 8.2.1 that choosing a Dirichlet boundary condition as the payoff for $v_{\min}$ leads to unstable approximation of Vega and Gamma. Thus, as pointed out by [CP99] the Neumann boundary condition is preferable for $v \to v_{\min}$.

Problem 2.11. *(Free boundary problem)*
Find $V := V(t,S,v)$ *with* $(t,S,v) \in [0,T) \times \mathbb{R}^+ \times \mathbb{R}^+$ *such that*

$$\frac{\partial V}{\partial t} + \mathcal{L}^H V = 0 \qquad \text{for all } S > S_f(t,v), \tag{2.95}$$

$$V = \mathcal{H}_P(S) \quad \text{for all } S \le S_f(t,v), \tag{2.96}$$

with appropriate boundary conditions and end condition $V(T,S,v) = \mathcal{H}_P(S)$.

For the numerical treatment of the free boundary problem, it is necessary to derive an equivalent formulation in which the a priori unknown free boundary $S_f(t,v)$ does not explicitly appear. Therefore, it is known from (2.96) that in the stopping area $S \le S_f(t,v) < K$ for a.e. $t \in [0,T)$ the option value $V(t,S,v)$ is determined by the payoff function $\mathcal{H}_{\mathrm{P}}(S)$. Applying the Heston equation to the payoff function for $S \le S_f(t,v) < K$ leads to

$$\frac{\partial V}{\partial t} + \mathcal{L}^H V = \frac{\partial (K-S)}{\partial t} + \mathcal{L}^H (K-S) = -rK < 0 \quad \text{for } r > 0. \tag{2.97}$$

Due to the no arbitrage principle (see e.g. [Sey09], appendix 7) we have $V(t,S,v) > \mathcal{H}_{\mathrm{P}}$, because otherwise the holder of the option could immediately exercise the option and receive a positive profit $\mathcal{H}(S) - V(t,S,v) > 0$. Multiplying (2.95) with (2.96) one derives the following complementarity problem:

Problem 2.12. *(Complementarity problem)*
Find $V(t,S,v)$ *with* $(t,S,v) \in [0,T) \times \mathbb{R}^+ \times \mathbb{R}^+$ *such that*

$$\left(\frac{\partial V}{\partial t} + \mathcal{L}^H V\right)(V - \mathcal{H}_P(S)) = 0, \tag{2.98}$$

$$\left(\frac{\partial V}{\partial t} + \mathcal{L}^H V\right) \leq 0, \tag{2.99}$$

$$V - \mathcal{H}_P(S) \geq 0 \tag{2.100}$$

with the set of boundary conditions in (2.93) or, alternatively, with the set of boundary conditions in (2.94) and end condition $V(T,S,v) = \mathcal{H}(S)$.

A similar formulation can be found in [Bos15, Bur16, KSW12] or with some modification of the boundary condition in [CP99, Oos03, IT09].
To eliminate the degeneracy in the Heston equation (2.23) a standard way is to transform the underlying price S by $x = \ln\left(\frac{S}{K}\right)$. To simplify further calculations we transform the backward problem to a forward problem by a new time coordinate $\tau := T - t$. Then the option price in new coordinates is defined by

$$y := y(\tau, x, v) = V(T - t, Ke^x, v). \tag{2.101}$$

Following [AWW01, Bos15, Bur16, KSW12, Wei14], the Heston equation can be written as

$$-\frac{\partial y}{\partial \tau} + \mathcal{Z}^H y := -\frac{\partial y}{\partial \tau} + \nabla \cdot A \nabla y - \mathbf{b} \cdot \nabla y - ry, \tag{2.102}$$

where $\nabla y := \left(\frac{\partial y}{\partial x}, \frac{\partial y}{\partial v}\right)^T \in \mathbb{R}^2$ denotes the gradient of y and $\nabla \cdot c(x,v) := \frac{\partial c_1(x,v)}{\partial x} + \frac{\partial c_2(x,v)}{\partial v}$ the divergence of the vector field $c(x,v) : \mathbb{R}^2 \to \mathbb{R}^2$. The matrix $A \in \mathbb{R}^{2\times 2}$ and the vector $\mathbf{b} \in \mathbb{R}^2$ are defined by

$$A := \frac{1}{2} v \begin{pmatrix} 1 & \varrho\xi \\ \varrho\xi & \xi^2 \end{pmatrix}, \mathbf{b} := \begin{pmatrix} \frac{1}{2}v + \frac{1}{2}\varrho\xi - r \\ \kappa(v - \gamma) + \lambda(v) + \frac{1}{2}\xi^2 \end{pmatrix}. \tag{2.103}$$

The transformed payoff function is denoted by

$$g(x) = g(x,v) := \mathcal{H}_{\mathrm{P}}(Ke^x, v) = \max\{K(1 - e^x), 0\}. \tag{2.104}$$

The time interval and the spatial domain in the new coordinates are given by

$$(\tau, x, v) \in (0, T] \times \mathbb{R} \times \mathbb{R}^+. \tag{2.105}$$

Note that the transformed underlying price is now on the whole axis of real numbers and the boundary condition in (2.93) reads as

$$\begin{aligned} \lim_{x \to -\infty} y(\tau, x, v) &= K, \quad \lim_{v \to v_{\min}} y(\tau, x, v) = g(x), \\ \lim_{x \to \infty} y(\tau, x, v) &= 0, \quad \lim_{v \to \infty} \frac{\partial y(\tau, x, v)}{\partial v} = 0, \end{aligned} \tag{2.106}$$

and for the boundary conditions in (2.94) one has

$$\begin{aligned} \lim_{x\to-\infty} y(\tau,x,v) = K, \quad \lim_{v\to v_{\min}} \frac{\partial y(\tau,x,v)}{\partial v} = 0, \\ \lim_{x\to\infty} y(\tau,x,v) = 0, \quad \lim_{v\to\infty} \frac{\partial y(\tau,x,v)}{\partial v} = 0. \end{aligned} \tag{2.107}$$

The free boundary $S_f(t,v)$ transforms into $x_f(\tau,v) := \log\left(\frac{S_f(t,v)}{K}\right)$. With this notation we can reformulate Problem 2.12 to the following complementarity problem in new coordinates:

Problem 2.13. *(Transformed complementarity problem)*
Find $y(\tau,x,v)$ with $(\tau,x,v) \in (0,T] \times \mathbb{R} \times \mathbb{R}^+$ such that

$$\begin{aligned} \left(-\frac{\partial y}{\partial \tau} + \mathcal{Z}^H y\right)(y-g) &= 0 \\ -\frac{\partial y}{\partial \tau} + \mathcal{Z}^H y &\le 0 \\ y - g &\ge 0 \end{aligned}$$

with the set of boundary conditions in (2.106) or, alternatively, with the set of boundary conditions in (2.107) and initial condition $y(0,x,v) = g(x)$.

In order to enable numerical computations, we localize the domain as

$$(0,T] \times \Omega := (0,T] \times (x_{\min}, x_{\max}) \times (v_{\min}, v_{\max}). \tag{2.108}$$

with $-\infty < x_{\min} < 0 < x_{\max} < \infty$ and $0 < v_{\min} < v_{\max}$. Therefore, the truncated boundaries are defined as follows:

$$\begin{aligned} \Upsilon_1 &:= \{v \in (v_{\min}, v_{\max}) : x = x_{\min}\}, & \Upsilon_2 &:= \{v \in (v_{\min}, v_{\max}) : x = x_{\max}\}, \\ \Upsilon_3 &:= \{x \in (x_{\min}, x_{\max}) : v = v_{\min}\}, & \Upsilon_4 &:= \{x \in (x_{\min}, x_{\max}) : v = v_{\max}\}. \end{aligned} \tag{2.109}$$

The boundary conditions for an American put option in (2.106) are set to

$$y(\tau,x,v) = g(x) \ \text{ on } \Upsilon := \Upsilon_1 \cup \Upsilon_2 \cup \Upsilon_3, \qquad \frac{\partial y(\tau,x,v)}{\partial v} = 0 \ \text{ on } \Upsilon_v := \Upsilon_4. \tag{2.110}$$

The alternative set of boundary conditions in (2.107) are replaced by

$$y(\tau,x,v) = g(x) \ \text{ on } \Upsilon := \Upsilon_1 \cup \Upsilon_2, \qquad \frac{\partial y(\tau,x,v)}{\partial v} = 0 \ \text{ on } \Upsilon_v := \Upsilon_3 \cup \Upsilon_4. \tag{2.111}$$

Problem 2.14. *(Localised transformed complementarity problem)*
Find $y(\tau, x, v)$ with $(\tau, x, v) \in (0, T] \times \Omega$, such that

$$\begin{aligned} \left(-\frac{\partial y}{\partial \tau} + \mathcal{Z}^H y\right)(y - g) &= 0 \\ \left(-\frac{\partial y}{\partial \tau} + \mathcal{Z}^H y\right) &\leq 0 \\ y - g &\geq 0 \end{aligned} \tag{2.112}$$

with $y = g$ on Υ, $\frac{\partial y}{\partial v} = 0$ on Υ_v and $y(0, x, v) = g(x)$ for all $(x, v) \in \Omega$.

The same formulation of the localised transformed complementarity problem with boundary condition as defined in (2.110) can be found in [Bos15, KSW12, Wei14], or with both alternatives (2.110) and (2.111), in [Bur16].
In the next step we formulate the problem as a parabolic variational inequality.

Remark 2.15. *It should be mentioned here that the bilinear form derived in this thesis differs from the bilinear form derived in [Bos15, Bur16, KSW12, Wei14]. This is due to the fact that after applying integration by parts the boundary integral on the Neumann boundary Υ_v, where y_x is unknown, does not vanish. To derive a bilinear form without a boundary integral, we apply integration by parts to the mixed term $\int v\rho\xi y_{xv}(\varphi - y)\mathrm{d}\Omega$ only with respect to x, which leads to a different bilinear form with an asymmetric matrix $\tilde{A}$. Due to this different result, it is impossible to use the same weighted Sobolev space $\|\cdot\|_A$ as in [Bur16, KSW12] citing [AWW01] induced by the symmetric matrix A defined in (2.103), to show the existence and uniqueness of the solution.*

In order to derive a variational formulation of Problem 2.14, we introduce the following spaces: Let $H := L^2(\Omega)$ the space of square integrable function. We define by $\mathcal{V}$ the Sobolev space of once weak differentiable function as follows

$$\mathcal{V} := \left\{\varphi \in H : \frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial v} \in H\right\} \tag{2.113}$$

equipped with the norm

$$\|\varphi\|_{\mathcal{V}}^2 := \int_\Omega \left(\left(\frac{\partial \varphi}{\partial x}\right)^2 + \left(\frac{\partial \varphi}{\partial v}\right)^2 + \varphi^2\right) \mathrm{d}\Omega. \tag{2.114}$$

Let further

$$\mathcal{V}^0 := \{\varphi \in \mathcal{V} : \varphi = 0 \text{ on } \Upsilon \subset \partial\Omega\} \tag{2.115}$$

be the space with zero boundary conditions on Υ and $\mathcal{V}^*$ the dual space of $\mathcal{V}^0$. We know from [DB06] that the space $D(\Omega) := C^\infty(\overline{\Omega}) \cap \mathcal{V}^0$ is dense in $\mathcal{V}^0$. Since $D(\Omega)$ is dense in $\mathcal{V}^0$ and $D(\Omega) \subset L^2(\Omega)$, the space $\mathcal{V}^0$ is also dense in $L^2(\Omega)$ as well. It is obvious, that the embedding $\mathcal{V}^0 \subset L^2(\Omega)$ is continuous because the $\mathcal{V}$-norm is by definition stronger than the $L^2(\Omega)$-norm. In summary, we conclude that $\mathcal{V}^0 \subset H = H^* \subset \mathcal{V}^*$ with $H := L^2(\Omega)$ is a continuous and dense embedding.

The inner product $(\cdot,\cdot)$ of H can be considered as an extension to the pairing between $\mathcal{V}^*$ and $\mathcal{V}^0$,

$$\|f\|_{\mathcal{V}^*} := \sup_{\eta\in\mathcal{V}^0\setminus 0} \frac{(f,\eta)}{\|\eta\|_{\mathcal{V}}}. \tag{2.116}$$

We also define the closed convex set by

$$\mathcal{K} := \{\varphi := \varphi(x,v) \in \mathcal{V} : \varphi(x,v) \geq g(x) \text{ for all } (x,v)\in\Omega,\ \varphi = g \text{ on } \Upsilon\} \subset \mathcal{V}. \tag{2.117}$$

Now, we present the derivation of the parabolic variational inequality. Let $\varphi \in \mathcal{K}$, multiplication of the first and second equation in Problem 2.14 by $\varphi - g \geq 0$ and integration over the domain Ω leads to

$$\int_\Omega \left(-\frac{\partial y}{\partial \tau} + \mathcal{Z}^H y\right)(y-g)\mathrm{d}\Omega = 0 \text{ and } \int_\Omega \left(-\frac{\partial y}{\partial \tau} + \mathcal{Z}^H y\right)(\varphi-g)\mathrm{d}\Omega \leq 0 \tag{2.118}$$

for all $\varphi \in \mathcal{K}$. Subtraction of the inequality and the equation above gives us

$$\int_\Omega \left(-\frac{\partial y}{\partial \tau} + \mathcal{Z}^H y\right)(\varphi-y)\mathrm{d}\Omega \leq 0 \quad \text{for all } \varphi\in\mathcal{K}. \tag{2.119}$$

We first present the derivation of the weak formulation in $\mathcal{V}$ associated with the operator $\mathcal{Z}^H$, where the boundary integral does not vanish due to the Neumann boundary condition on Υ_v. More precisely, let us consider the diffusion term of the Heston equation defined in (2.102), applying integration by parts to $\int \frac{1}{2}v(y_{xx}+\varrho\xi y_{xv})(\varphi-y)\mathrm{d}\Omega$ with respect to x and to $\int \frac{1}{2}v\left(\varrho\xi y_{xv}+\xi^2 y_{vv}\right)(\varphi-y)\mathrm{d}\Omega$ with respect to v leads to

$$\begin{aligned}
&\int_\Omega \nabla\cdot A\nabla y(\varphi-y)d\Omega = \int_\Omega \frac{1}{2}v\left(y_{xx}+\varrho\xi y_{xv}\right)(\varphi-y)d\Omega \\
&+\int_\Omega \frac{1}{2}v\left(\varrho\xi y_{xv}+\xi^2 y_{vv}\right)(\varphi-y)d\Omega + \int_\Omega \frac{1}{2}\left(\varrho\xi y_x+\xi^2 y_v\right)(\varphi-y)d\Omega \\
=&-\int_\Omega \frac{1}{2}v\begin{pmatrix} y_x+\varrho\xi y_v \\ \varrho\xi y_x+\xi^2 y_v\end{pmatrix}\cdot\begin{pmatrix}(\varphi-y)_x\\(\varphi-y)_v\end{pmatrix}d\Omega - \int_\Omega \frac{1}{2}\left(\varrho\xi y_x+\xi^2 y_v\right)(\varphi-y)d\Omega \\
&+\int_{\partial\Omega} \frac{1}{2}v\begin{pmatrix} y_x+\varrho\xi y_v \\ \varrho\xi y_x+\xi^2 y_v\end{pmatrix}\cdot\mathbf{n}(\varphi-y)d\partial\Omega + \int_\Omega \frac{1}{2}\left(\varrho\xi y_x+\xi^2 y_v\right)(\varphi-y)d\Omega \\
=&-\int_\Omega (A\nabla y)\cdot\nabla(\varphi-y)d\Omega + \int_{\partial\Omega}(A\nabla y)\cdot\mathbf{n}(\varphi-y)d\partial\Omega
\end{aligned} \tag{2.120}$$

for all $\varphi\in\mathcal{K}$, where $\mathbf{n}$ is the outward pointing unit normal vector of the boundary $\partial\Omega$. First we consider the boundary integral for the set of boundary conditions defined in (2.110). Due to the boundary conditions and the choice of the test function space $\varphi = y = g$ on $\Upsilon := \Upsilon_1\cup\Upsilon_2\cup\Upsilon_3$, the boundary integrals on Υ are equal to zero. Together with the Neumann boundary condition on $\Upsilon_v := \Upsilon_4$ this yields

$$\int_{\partial\Omega}(A\nabla y)\cdot\mathbf{n}(\varphi-y)d\partial\Omega = \int_{\Upsilon_4}(A\nabla y)\cdot\begin{pmatrix}0\\1\end{pmatrix}(\varphi-y)d\Upsilon_4 = \int_{\Upsilon_4}\frac{1}{2}v(\varrho\xi y_x)(\varphi-y)d\Upsilon_4. \tag{2.121}$$

For the alternative set of boundary conditions defined in (2.111) one has

$$\begin{aligned}
&\int_{\partial\Omega}(A\nabla y)\cdot\mathbf{n}(\varphi-y)d\partial\Omega\\
&=\int_{\Upsilon_3}(A\nabla y)\cdot\begin{pmatrix}0\\-1\end{pmatrix}(\varphi-y)d\Upsilon_3+\int_{\Upsilon_4}(A\nabla y)\cdot\begin{pmatrix}0\\1\end{pmatrix}(\varphi-y)d\Upsilon_4\\
&=\int_{\Upsilon_4}\frac{1}{2}v(\varrho\xi y_x)(\varphi-y)d\Upsilon_4-\int_{\Upsilon_3}\frac{1}{2}v(\varrho\xi y_x)(\varphi-y)d\Upsilon_3. \qquad (2.122)
\end{aligned}$$

Consequently, for both sets of boundary conditions the boundary integral on Υ_v does not vanish.
Using one of the formulations above to show existence and uniqueness of a weak solution causes difficulties because of the boundary integral, where y_x is unknown. To avoid this problem, we apply the integration by parts theorem to $\int_\Omega \frac{1}{2}v\left(y_{xx}+2\varrho\xi y_{xv}\right)(\varphi-y)\mathrm{d}\Omega$ only with respect to x and to $\int\frac{1}{2}v\xi^2y_{vv}(\varphi-y)d\Omega$ with respect to v. Due to the boundary condition on Υ_v and the choice of the test function space $\varphi=y=g$ on Υ all boundary integrals vanish, i.e. this yields

$$\begin{aligned}
&\int_\Omega\nabla\cdot A\nabla y(\varphi-y)d\Omega=\int_\Omega\frac{1}{2}v\left(y_{xx}+2\varrho\xi y_{xv}\right)(\varphi-y)\mathrm{d}\Omega\\
&+\int_\Omega\frac{1}{2}v\xi^2y_{vv}(\varphi-y)d\Omega+\int_\Omega\frac{1}{2}\left(\varrho\xi y_x+\xi^2y_v\right)(\varphi-y)d\Omega\\
=&-\int_\Omega\frac{1}{2}v\begin{pmatrix}y_x+2\varrho\xi y_v\\ \xi^2y_v\end{pmatrix}\cdot\begin{pmatrix}(\varphi-y)_x\\(\varphi-y)_v\end{pmatrix}d\Omega-\int_\Omega\frac{1}{2}\xi^2y_v(\varphi-y)d\Omega\\
&+\int_{\partial\Omega}\frac{1}{2}v\begin{pmatrix}y_x+2\varrho\xi y_v\\ \xi^2y_v\end{pmatrix}\cdot\mathbf{n}(\varphi-y)ds+\int_\Omega\frac{1}{2}\left(\varrho\xi y_x+\xi^2y_v\right)(\varphi-y)d\Omega\\
=&-\int_\Omega\tilde{A}\nabla y\cdot\nabla(\varphi-y)d\Omega+\int_\Omega\frac{1}{2}\varrho\xi y_x(\varphi-y)d\Omega, \qquad (2.123)
\end{aligned}$$

where $\tilde{A}\in\mathbb{R}^{2\times2}$ is defined as

$$\tilde{A}:=\frac{1}{2}v\begin{pmatrix}1&2\varrho\xi\\0&\xi^2\end{pmatrix}. \qquad (2.124)$$

Inserting (2.123) in the spatial operator of (2.119) we get

$$\begin{aligned}
(\mathcal{Z}^Hy,\varphi-y)=&-\int_\Omega\tilde{A}\nabla y\cdot\nabla(\varphi-y)\mathrm{d}\Omega+\int_\Omega\frac{1}{2}\varrho\xi y_x(\varphi-y)\mathrm{d}\Omega\\
&-\int_\Omega\mathbf{b}\cdot\nabla y(\varphi-y)\mathrm{d}\Omega-\int_\Omega ry(\varphi-y)\mathrm{d}\Omega\\
=&-\int_\Omega\tilde{A}\nabla y\cdot\nabla(\varphi-y)\mathrm{d}\Omega+\int_\Omega\frac{1}{2}\varrho\xi y_x(\varphi-y)\mathrm{d}\Omega\\
&-\int_\Omega\left(\frac{1}{2}v+\frac{1}{2}\rho\xi-r\right)y_x(\varphi-y)\mathrm{d}\Omega\\
&-\int_\Omega\left(\kappa(v-\gamma)+\lambda(v)+\frac{1}{2}\xi^2\right)y_v(\varphi-y)\mathrm{d}\Omega-\int_\Omega ry(\varphi-y)\mathrm{d}\Omega
\end{aligned}$$

$$
\begin{aligned}
&= -\int_\Omega \tilde{A}\nabla y \cdot \nabla(\varphi - y) + \left(\frac{1}{2}v - r\right) y_x(\varphi - y)\mathrm{d}\Omega \\
&\quad - \int_\Omega \left(\kappa(v-\gamma) + \lambda(v) + \frac{1}{2}\xi^2\right) y_v(\varphi - y) + ry(\varphi - y)\mathrm{d}\Omega \\
&=: -\int_\Omega \tilde{A}\nabla y \cdot \nabla(\varphi - y) + \tilde{\mathbf{b}} \cdot \nabla y(\varphi - y) + ry(\varphi - y)\mathrm{d}\Omega, \qquad (2.125)
\end{aligned}
$$

where $\tilde{\mathbf{b}} \in \mathbb{R}^2$ is defined as

$$
\tilde{\mathbf{b}} := \begin{pmatrix} \tilde{b}_1(v) \\ \tilde{b}_2(v) \end{pmatrix} := \begin{pmatrix} \frac{1}{2}v - r \\ \kappa(v-\gamma) + \lambda(v) + \frac{1}{2}\xi^2 \end{pmatrix}. \qquad (2.126)
$$

In summary, we have proved:

Lemma 2.16. *The weak formulation of $\mathcal{Z}^H y$ defined in (2.102) with the set of boundary conditions in (2.106) or (2.107) is*

$$
-\int_\Omega \tilde{A}\nabla y \cdot \nabla(\varphi - y) + \tilde{\mathbf{b}} \cdot \nabla y(\varphi - y) + ry(\varphi - y)\mathrm{d}\Omega. \qquad (2.127)
$$

Inserting (2.125) in (2.119) and multiplication of the inequality by -1 leads to the following parabolic variational inequality:

Problem 2.17. *(Variational inequality with non-homogeneous boundary condition) Find $y \in L^2(0,T;\mathcal{K})$ such that $\frac{\partial y}{\partial \tau} \in L^2(0,T;H)$*

$$
\left(\frac{\partial y}{\partial \tau}, \varphi - y\right) + a^H(y, \varphi - y) \geq 0 \quad \text{for all } \varphi \in \mathcal{K}
$$

and $y(0,x,v) := g$ is satisfied, where $a^H(\cdot,\cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ defines the bilinear form

$$
a^H(y, \varphi - y) := \int_\Omega \left(\tilde{A}\nabla y \cdot \nabla(\varphi - y) + (\tilde{\mathbf{b}} \cdot \nabla y + ry)(\varphi - y)\right) d\Omega \qquad (2.128)
$$

with $\tilde{A} \in \mathbb{R}^{2\times 2}$ and $\tilde{\mathbf{b}} \in \mathbb{R}^2$ as defined in (2.124) and (2.126).

Now, we transform Problem 2.17 into a problem with homogeneous boundary conditions. The trace theory (c.f. [AF03]) indicates that a linear continuous extension operator

$$
\mathcal{T} : H^{\mathfrak{m}-\frac{1}{2}}(\Upsilon) \to H^{\mathfrak{m}}(\Omega) \qquad (2.129)
$$

for some $\mathfrak{m} \in \mathbb{N}$, $\mathfrak{m} \geq 1$ extends the non-homogeneous boundary condition to the interior of the domain, i.e. the extension of $g \in H^{\mathfrak{m}-\frac{1}{2}}(\Upsilon)$ is the function $u_g \in H^{\mathfrak{m}}(\Omega)$.
For the set of boundary conditions in (2.110) the Dirichlet condition for $v = v_{\min}$ is given by the non-smooth payoff function $g(x) \in H^1(\Upsilon)$ for all $x \in (x_{\min}, x_{\max})$. For the alternative set of boundary conditions in (2.111) the Dirichlet conditions are only set to the function $g(x)$ on the boundaries for $x = x_{\min} < 0$ and $x = x_{\max} > 0$, where the

function $g(x)$ is arbitrarily smooth. In other words, in this case the Dirichlet boundary conditions are only specified by the point evaluations $g(x_{\min})$ and $g(x_{\max})$. Thus, for the set of boundary conditions in (2.111) a linear extension of $g(x_{\min})$ and $g(x_{\max})$ enables $u_g \in H^2(\Omega)$.
Since the initial condition and the obstacle function are also specified by the function $g(x)$, for further numerical computations it is comfortable to choose $u_g := g(x,v) = g(x) \in H^1(\Omega)$ for both sets of boundary conditions. We set $u(\tau,x,v) := y(\tau,x,v) - g(x,v) =: y - g$ and introduce the following closed convex set

$$\mathcal{K}^0 := \{\varphi := \varphi(x,v) \in \mathcal{V}^0 : \varphi(x,v) \geq 0\}. \tag{2.130}$$

For $g \in H^1(\Omega)$ the term $a^H(g, \varphi - u)$ is well-posed and Problem 2.17 is equivalent to the following problem with homogeneous boundary conditions:

Problem 2.18. *(Variational inequality with homogeneous boundary condition)*
Find $u \in L^2(0,T;\mathcal{K}^0)$ *such that* $\frac{\partial u}{\partial \tau} \in L^2(0,T;H)$

$$\left(\frac{\partial u}{\partial \tau}, \varphi - u\right) + a^H(u, \varphi - u) \geq \langle f, \varphi - u\rangle \quad \textit{for all } \varphi \in \mathcal{K}^0$$

and $u(0,x,v) := 0$ *is satisfied, where* $a^H(\cdot,\cdot) : \mathcal{V}^0 \times \mathcal{V}^0 \to \mathbb{R}$ *defines the bilinear form in (2.128) and right side*

$$\langle f, \varphi - u\rangle := -a^H(g, \varphi - u). \tag{2.131}$$

3. Background Material

After having introduced the parabolic variational inequality for the valuation of an American option in the Black-Scholes or Heston's model, this section is devoted to study some regularity results for variational inequalities. These results are important for two reasons, namely, for the error analysis for further numerical computations and to clarify the maximal smoothness of the Greeks, the partial derivatives of the American option price with respect to the model parameters. Since the initial condition and the obstacle function for pricing American options is only once weak differentiable, the results for variational inequalities given in the literature have to be considered carefully.

3.1. Well-posedness and regularity results for variational inequalities

3.1.1. Elliptic variational inequalities

In this subsection we give an introduction to second order elliptic variational inequalities and an short overview of some regularity results. For a very good introduction to elliptic variational inequalities we refer to [KS00] and [BL82, Chapter 3.1].
We start with the introduction of the notation. We denote by $H := L^2(\Omega)$ the space of square integrable function on $\Omega \subset \mathbb{R}^d$ for $d = 1, 2$ equipped with the norm

$$\|w\|_H = (w, w)_H^{\frac{1}{2}} := \left(\int_\Omega w^2 \mathrm{d}\Omega \right)^{\frac{1}{2}} \quad \text{for all } w \in H. \tag{3.1}$$

The dual space of H is denoted by H^* and is defined by

$$H^* := \left\{ f : H \to \mathbb{R} \Big| \sup_{w \in H} \frac{|(f, w)|}{\|w\|_H} < \infty \right\}. \tag{3.2}$$

Due to the Riesz representation theorem, it is well known that H^* is isomorphic to H and one can identify H^* with H. We denote by $\mathcal{V} \subset H$ the Sobolev space (or possibly weighted Sobolev space) of once weakly differentiable functions and inner product $(\varphi, \varphi)_\mathcal{V}$, which induces the norm

$$\|\varphi\|_\mathcal{V} := (\varphi, \varphi)_\mathcal{V}^{\frac{1}{2}} \quad \text{for all } \varphi \in \mathcal{V}, \tag{3.3}$$

such that the canonical embedding is continuous, i.e.,

$$\|\varphi\|_H \lesssim \|\varphi\|_{\mathcal{V}} \quad \text{for all } \varphi \in \mathcal{V}. \tag{3.4}$$

Let further the space $\mathcal{V}^0 \subset \mathcal{V}$ denote the Sobolev space of once weakly differentiable functions with homogeneous Dirichlet boundary condition on $\Upsilon \subseteq \partial\Omega$. For any $w \in H$ the map $\varphi \to (w, \varphi)$ belongs to the dual space H^*, since by applying the Cauchy-Schwarz inequality and the canonical embedding we have

$$|(w, \varphi)_H| \leq \|w\|_H \|\varphi\|_H \lesssim \|w\|_H \|\varphi\|_{\mathcal{V}} \quad \text{for all } \varphi \in \mathcal{V}. \tag{3.5}$$

The dual space of $\mathcal{V}^0$ is defined as

$$\|f\|_{\mathcal{V}^*} := \sup_{\varphi \in \mathcal{V}^0 \setminus 0} \frac{(f, \varphi)_H}{\|\varphi\|_{\mathcal{V}}}, \tag{3.6}$$

hence we have the following dense and continuous embedding, $\mathcal{V}^0 \subset H = H^* \subset \mathcal{V}^*$. In the triple we define the linear operator $\mathcal{A} : \mathcal{V}^0 \to \mathcal{V}^*$ associated with the bilinear form

$$\langle \mathcal{A}u, \varphi \rangle_{\mathcal{V}^*, \mathcal{V}} = a(u, \varphi). \tag{3.7}$$

We also define the non-empty closed convex set

$$\mathcal{K} := \{\varphi \in \mathcal{V} : \varphi \geq \psi, \ \varphi = \psi \text{ on } \Upsilon\} \subset \mathcal{V} \tag{3.8}$$

of functions with inhomogeneous Dirichlet boundary condition ψ on $\Upsilon \subseteq \partial\Omega$.
With this notation a general elliptic variational inequality can be defined as follows:

Problem 3.1. *(Elliptic variational inequality with inhomogeneous boundary condition) Let $f \in \mathcal{V}^*$ and $\mathcal{K} \subseteq \mathcal{V}$ be a non-empty closed convex set. Find $y \in \mathcal{K}$ such that*

$$a(y, \varphi - y) \geq \langle f, \varphi - y \rangle_{\mathcal{V}^*, \mathcal{V}} \quad \textit{for all } \varphi \in \mathcal{K}, \tag{3.9}$$

where $a : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is a bilinear form.

For $f \in L^2(\Omega)$ the duality pairing $\langle \cdot, \cdot \rangle_{\mathcal{V}^*, \mathcal{V}}$ in (3.9) can be replaced by the $L^2(\Omega)$ inner product $(\cdot, \cdot)$. In order to find an equivalent formulation with homogeneous boundary conditions on $\Upsilon \subseteq \Omega$ we can set $u := y - u_\psi$, where u_ψ is the linear extension to the interior of the domain. Further we denote the modified obstacle function by $\widetilde{\psi} := \psi - u_\psi$ and introduce the closed convex subset of $\mathcal{V}^0$

$$\mathcal{K}^0 := \{\varphi \in \mathcal{V}^0 : \varphi \geq \widetilde{\psi}\}. \tag{3.10}$$

With this notation we can formulate an equivalent formulation with homogeneous boundary conditions on $\Upsilon \subseteq \partial\Omega$.

Problem 3.2. *(Elliptic variational inequality with homogeneous boundary condition) Let $\mathcal{K}^0 \subseteq \mathcal{V}^0$ be a non-empty closed convex set. Find $u \in \mathcal{K}^0$ such that*

$$a(u, \varphi - u) \geq (f, \varphi - u) - a(u_\psi, \varphi - u) \quad \text{for all } \varphi \in \mathcal{K}^0, \tag{3.11}$$

where $a : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is a bilinear form.

The existence and uniqueness result for elliptic variational inequalities has been developed by Lions and Stampacchia [LS67]. The main idea of this proof is the application of the well known Banach fixed point theorem. A different approach is proposed by Kinderlehrer and Stampacchia [KS80, KS00]. There, the existence result for a symmetric bilinear form is proved by a minimization principle and treating the general case as an perturbation of the symmetric one. The authors in [TLG81] have proved some results for variational inequalities with a penalty approach. The following theorem summarizes the result of [KS80, KS00].

Theorem 3.3. *(Existence and uniqueness for elliptic variational inequalities) Suppose $f \in \mathcal{V}^*, \tilde{\psi} \in \mathcal{V}^0$ and the following conditions*

i) $\mathcal{K}^0$ is a closed convex non-empty set in $\mathcal{V}^0$.

ii) The bilinear form $a(\cdot, \cdot)$ is bounded, i.e. there exists $C > 0$ such that

$$|a(w, \varphi)| \leq C\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} \quad \text{for all } w, \varphi \in \mathcal{V}. \tag{3.12}$$

iii) The bilinear form $a(\cdot, \cdot)$ is coercive on $\mathcal{V}^0$, i.e. there exists $\alpha_0 > 0$ such that

$$a(\varphi, \varphi) \geq \alpha_0\|\varphi\|_{\mathcal{V}}^2 \quad \text{for all } \varphi \in \mathcal{V}^0. \tag{3.13}$$

Then there exists a unique solution $u \in \mathcal{V}^0$ to Problem 3.2.

Remark 3.4. *If $\mathcal{K}^0 = \mathcal{V}^0$ Theorem 3.3 reduces to the well known Lax-Milgram Theorem.*

The authors in [KS00, pp. 106–113] or [BL82, Chapter 3.1, p. 206] have shown that for $f \in L^2(\Omega)$ and $\mathcal{A}\tilde{\psi} \in L^2(\Omega)$ the solution satisfies $u \in H^2(\Omega)$ using a method called *penalization*. As for American option pricing problems, which are considered in this thesis, the assumption on the obstacle given by the non-smooth payoff function is in general not satisfied, we present a result with less restrictive assumptions on the obstacle. Since no regularity results for the elliptic case with obstacle functions in $H^1(\Omega)$ are found in the literature, we give a proof in Lemma 3.5. The main trick in the proof (i.e. assuming (3.15) below) is based on the ideas from [Mem12] and [KTK17] proving some results for the parabolic case in the context of American options.

To do so, we use the penalty approach. The *penalty method* consists of substituting the variational inequality by a family of non-linear equations and showing some regularity results for this penalized problem. Then, it can be shown that their solutions converge to the solution of the variational inequality. Therefore, we introduce the penalty problem corresponding to Problem 3.2 as follows: For $\varepsilon > 0$ find $u^\varepsilon \in \mathcal{V}^0$ such that

$$a(u^\varepsilon, \varphi) + \varepsilon^{-1}(\mathfrak{Q}u^\varepsilon, \varphi) = \langle f, \varphi\rangle - a(u_\psi, \varphi) \tag{3.14}$$

for all $\varphi \in \mathcal{V}^0$ with penalty function $\mathfrak{Q}u^\varepsilon := \min(u^\varepsilon - \tilde{\psi}, 0)$. It is well known that the penalty problem has a unique solution $u^\varepsilon \in \mathcal{V}^0$ (see [BL82, Chapter 3.1, p. 193]).
Note that the assumption on the obstacle in (3.15) below is satisfied for convex functions.

Lemma 3.5. *Let $a(\cdot,\cdot)$ be a bounded and coercive bilinear form. Suppose $f \in L^2(\Omega)$, $u_\psi \in H^2(\Omega)$. Moreover, let $\tilde{\psi} \in \mathcal{V}^0$ such that*

$$-a(\tilde{\psi}, \mathfrak{Q}u^\varepsilon) \lesssim \|\tilde{\psi}\|_{\mathcal{V}} \|\mathfrak{Q}u^\varepsilon\|_H, \tag{3.15}$$

then the solution of Problem 3.2 satisfies $u^\varepsilon \in H^2(\Omega)$ and the following estimate holds

$$\|u^\varepsilon\|_{H^2(\Omega)} \lesssim \|\tilde{\psi}\|_{\mathcal{V}} + \|f\|_{L^2(\Omega)} + \|u_\psi\|_{H^2(\Omega)} \tag{3.16}$$

with a constant independent of the penalty parameter ε.

Proof. We start by setting $\varphi = \mathfrak{Q}u^\varepsilon \in \mathcal{V}^0$ in (3.14), which results in

$$a(u^\varepsilon, \mathfrak{Q}u^\varepsilon) + \varepsilon^{-1}\|\mathfrak{Q}u^\varepsilon\|_H^2 = \langle f, \mathfrak{Q}u^\varepsilon\rangle - a(u_\psi, \mathfrak{Q}u^\varepsilon). \tag{3.17}$$

As $a(\mathfrak{Q}u^\varepsilon, \mathfrak{Q}u^\varepsilon) = a(u^\varepsilon - \tilde{\psi}, \mathfrak{Q}u^\varepsilon)$ the relation above leads to

$$a(\mathfrak{Q}u^\varepsilon, \mathfrak{Q}u^\varepsilon) + \varepsilon^{-1}\|\mathfrak{Q}u^\varepsilon\|_H^2 = -a(\tilde{\psi}, \mathfrak{Q}u^\varepsilon) + \langle f, \mathfrak{Q}u^\varepsilon\rangle - a(u_\psi, \mathfrak{Q}u^\varepsilon). \tag{3.18}$$

Thus, assumption (3.15), coercivity and boundedness of $a(\cdot,\cdot)$, $f \in H$ and applying Cauchy-Schwarz inequality results in

$$\alpha_0\|\mathfrak{Q}u^\varepsilon\|_{\mathcal{V}}^2 + \varepsilon^{-1}\|\mathfrak{Q}u^\varepsilon\|_H^2 \lesssim \left(\|\tilde{\psi}\|_{\mathcal{V}} + \|f\|_H + \|u_\psi\|_{H^2(\Omega)}\right)\|\mathfrak{Q}u^\varepsilon\|_H, \tag{3.19}$$

which implies

$$\|\mathfrak{Q}u^\varepsilon\|_H \lesssim \varepsilon\left(\|\tilde{\psi}\|_{\mathcal{V}} + \|f\|_H + \|u_\psi\|_{H^2(\Omega)}\right). \tag{3.20}$$

It follows from the elliptic regularity theory for variational equations that for $\Omega \subset \mathbb{R}^d$

$$a(u^\varepsilon, \varphi) = \langle f, \varphi\rangle - a(u_\psi, \varphi) - \varepsilon^{-1}(\mathfrak{Q}u^\varepsilon, \varphi) \tag{3.21}$$

with $f \in H$, $u_\psi \in H^2(\Omega)$ and $\mathfrak{Q}u^\varepsilon \in H$ that there exists a unique solution and the following estimate holds

$$\|u^\varepsilon\|_{H^2(\Omega)} \lesssim \|u\|_{\mathcal{V}} + \|u_\psi\|_{H^2(\Omega)} + \varepsilon^{-1}\|\mathfrak{Q}u^\varepsilon\|_H + \|f\|_H. \tag{3.22}$$

Finally, applying (3.20) together with Theorem 3.3 gives the desired result. □

By passing to the limit in (3.16) as $\varepsilon \to 0$ it can be shown that the solution of the variational inequality satisfies the following regularity result.

Theorem 3.6. *Let $a(\cdot,\cdot)$ be a bounded and coercive bilinear form. Suppose $f \in L^2(\Omega)$, $u_\psi \in H^2(\Omega)$. Moreover let $\tilde{\psi} \in \mathcal{V}^0$ satisfy (3.15). Then the solution satisfies $u \in H^2(\Omega)$.*

Proof. The proof consists of passing to the limit as $\varepsilon \to 0$ and proving that the solution is a unique solution of the variational inequality. It can be proven analogously to [BL82, Chapter 3.1]. □

For an elliptic variational equation the smoothness of the right hand side f, the boundary condition function ψ on $\partial\Omega$ and the domain Ω are directly related to the smoothness of the solution u, i.e. if $f \in H^{\mathfrak{m}-2}(\Omega)$, $\psi \in H^{\mathfrak{m}-1/2}(\partial\Omega)$, and the boundary of Ω is in $C^{\mathfrak{m}}$, then the solution satisfies $u \in H^{\mathfrak{m}}(\Omega)$. But the statement is not true for elliptic variational inequalities in general. The threshold of smoothness for an elliptic variational inequality with zero boundary condition has been established by Brézis [Bré71] and is provided in Theorem 3.7.

Theorem 3.7. *(Maximal smoothness for elliptic variational inequality)*
If $f \in C^1(\overline{\Omega})$, $\tilde{\psi} \in C^3(\overline{\Omega})$ and if the boundary of Ω is sufficiently smooth, the solution satisfies

$$u \in W^{\mathfrak{s},p}(\Omega) \cap W^{2,\infty}(\Omega) \quad \text{for all } \mathfrak{s} < 2 + \frac{1}{p} \quad \text{and} \quad 1 < p < \infty, \tag{3.23}$$

where $W^{\mathfrak{s},p}(\Omega)$ denotes the Sobolev-Slobodeckij space with non-integer $\mathfrak{s} > 0$.

For integer $\mathfrak{s}, \mathfrak{m} \in \mathbb{N}$ the Sobolev-Slobodeckij space reduces to the well known Sobolev space $W^{\mathfrak{m},p}$ of $\mathfrak{m}$'th weak differentiable function in the Lebesgue space $L^p(\Omega)$. A proof for $W^{2,\infty}(\Omega)$ regularity assuming $f \in C^1(\overline{\Omega})$ and $\psi \in C^2(\overline{\Omega})$ was also established by [KS00, Chapter IV.6].
As regards the valuation of American option we are interested in the approximation of the partial derivatives up to order two, we also consider the behavior of the classical derivatives of the solution. At this point, it should be mentioned that the smoothness assumption in Theorem 3.7 for variational inequalities arising in the valuation of American options for the obstacle function, given by a non-smooth piecewise linear payoff function, is not satisfied. But it is still helpful to explain, which regularity for an variational inequality can be expected in the best case.
Under the presented assumption the conclusion is that the solution of an elliptic variational inequality lies in the Hölder space $C^{1,1}(\overline{\Omega})$, because the space $W^{2,\infty}(\overline{\Omega})$ is defined to be the Hölder space $C^{1,1}(\overline{\Omega})$. Moreover, this implies that the solution is once classical differentiable for the presented assumption.
$W^{2,\infty}(\Omega)$ regularity for the solution of a variational inequality also implies that the second derivative of the solution is bounded and exists on some neighborhood in Ω.

3.1.2. Parabolic variational inequalities

In this section, we consider time dependent problems on the time interval $[0,T] \subset \mathbb{R}$. Let $H, \mathcal{V}$ and $\mathcal{V}^0$ as defined in Section 3.1.1. For the formulation of parabolic variational inequality, we introduce the definition of Bochner spaces. A detailed explanation of Bochner spaces can be found in [Eva98].
The Bochner space $L^p(0,T;X)$ contains all real vector-valued functions $u : (0,T) \to X$, where X is a real Banach space, such that the norms

$$\|u\|_{L^p(0,T;X)} := \left(\int_0^T \|u(\tau)\|_X^2 d\tau \right)^{\frac{1}{p}}$$

for $1 \leq p < \infty$ and

$$\|u\|_{L^\infty(0,T;X)} := \operatorname{ess\,sup}_{0 \leq t \leq T} \|u(\tau)\|_X$$

are finite. Let further $C([0,T];X)$ be the space of all continuous real vector-valued functions $u : (0,T) \to X$ with finite

$$\|u\|_{C([0,T];X)} := \max_{0 \leq t \leq T} \|u(\tau)\|_X.$$

Due to the definition of Bochner spaces, the elements $u \in L^p(0,T;X)$ are functions defined on $(0,T)$ with values in X. The next theorems for Bochner spaces will be essential for further regularity results on parabolic variational inequalities.

Theorem 3.8. *Suppose $u \in L^2(0,T;\mathcal{V}^0)$ and $u_t \in L^2(0,T,\mathcal{V}^*)$ then*

$$u \in C([0,T];H).$$

Proof. See [Eva98, Chapter 5, pp. 287–288]. □

In particular, Theorem 3.8 ensures that point evaluations on the interval $[0,T]$ exists for functions $u \in L^2(0,T;\mathcal{V}^0)$ and $u_t \in L^2(0,T,\mathcal{V}^*)$.

Theorem 3.9. *Suppose $u \in L^2(0,T;H^{\mathfrak{m}+2}(\Omega))$ with $u_t \in L^2(0,T;H^{\mathfrak{m}}(\Omega))$, then*

$$u \in C([0,T];H^{\mathfrak{m}+1}(\Omega)). \tag{3.24}$$

Parabolic variational inequalities in the context of option pricing problems have often the special structure that the initial condition and the obstacle function are given by the same non-smooth function $\psi \in H^1(\Omega)$ and can be formulated as follows:

Problem 3.10. *(Parabolic variational inequality)*
Let $\mathcal{K} := \{\varphi \in \mathcal{V} : \varphi \geq \psi\} \subseteq \mathcal{V}$ a closed convex non-empty set, $\psi \in \mathcal{V}$ a given obstacle function and $f(\tau) : (0,T) \to \mathcal{V}^*$ a given source term. Find $y : (0,T) \to \mathcal{V}$ such that $y(\tau) \in \mathcal{K}$ for a.e. $\tau \in (0,T)$, $y(0) = \psi$ and

$$\left\langle \frac{\partial y(\tau)}{\partial \tau}, \varphi - y(\tau) \right\rangle_{\mathcal{V}^*,\mathcal{V}} + a(y(\tau), \varphi - y(\tau)) \geq \langle f(\tau), \varphi - y(\tau) \rangle_{\mathcal{V}^*,\mathcal{V}} \quad \text{for all } \varphi \in \mathcal{K}. \tag{3.25}$$

First, we rewrite Problem 3.10 to a problem with homogeneous boundary condition with the time independent function u_ψ as defined in Section 3.1.1.

Problem 3.11. *(Parabolic variational inequality with homogeneous boundary condition)*
Let $\mathcal{K}^0 := \{\varphi \in \mathcal{V}^0 : \varphi \geq \tilde{\psi}\} \subseteq \mathcal{V}$ a closed convex non-empty set, where $\tilde{\psi} := \psi - u_\psi$ denotes the transformed obstacle function. Find $u : (0,T) \to \mathcal{V}^0$ such that $u(\tau) \in \mathcal{K}^0$ for a.e. $\tau \in (0,T)$,

$$\left\langle \frac{\partial u(\tau)}{\partial \tau}, \varphi - u(\tau) \right\rangle_{\mathcal{V}^*,\mathcal{V}} + a(u(\tau) + u_\psi, \varphi - u(\tau)) \geq \langle f(\tau), \varphi - u(\tau)\rangle_{\mathcal{V}^*,\mathcal{V}} \quad \text{for all } \varphi \in \mathcal{K}^0 \tag{3.26}$$

and $u(0) = \psi - u_\psi =: \tilde{\psi}$.

There are many articles discussing existence and uniqueness results for parabolic variational inequalities. The authors in [TLG81] show the existence and uniqueness of a solution u, $\frac{\partial u}{\partial \tau} \in L^2(0,T;\mathcal{V}) \cap L^\infty(0,T;H)$ for parabolic variational inequalities based on an approximation method, if the operator is bounded and coercive and f, $\frac{\partial f}{\partial \tau} \in L^2(0,T;\mathcal{V}^*)$ and $f(0) - \mathcal{A}u(0) \in H$ is given. In the context of option pricing problems the last-named result is not applicable, because due to the non-smooth initial condition usually $f(0) - \mathcal{A}u(0) \in H$ is not satisfied. Further Ito and Kunisch [IK06] introduced a Lagrange multiplier approach to prove the existence and uniqueness of solutions $u \in H^1(0,T;H) \cap C(0,T;\mathcal{V})$ for parabolic variational inequalities, if the operator is symmetric, bounded and satisfies a Gårding inequality and $f \in L^2(0,T;H)$ is given.
Existence and uniqueness of the solution and regularity results for parabolic variational inequalities using a penalty approach are discussed in the literature [BL82, Chapter 3.2]. In particular [Ach05, DL19b, KTK17, Mem12] discussed it in the context of American options. Following the authors mentioned, we introduce the penalized problem corresponding to Problem 3.11 with homogeneous boundary conditions as follows:
For $\varepsilon > 0$, find $u^\varepsilon : (0,T] \to \mathcal{V}^0$ such that

$$\left\langle \frac{\partial u^\varepsilon(\tau)}{\partial \tau}, \varphi \right\rangle + a(u^\varepsilon(\tau), \varphi) + \varepsilon^{-1}(\mathfrak{Q}u^\varepsilon, \varphi) = \langle f(\tau), \varphi\rangle - a(u_\psi, \varphi) \text{ for all } \varphi \in \mathcal{V}^0, \text{ a.e. } \tau \in (0,T], \tag{3.27}$$

and $u^\varepsilon(0) = \widetilde{\psi}$, where $\mathfrak{Q}u^\varepsilon := \min(u^\varepsilon - \widetilde{\psi}, 0)$ denotes the penalty function.
We shall see that the regularity results presented in Lemma 3.13 or Theorem 3.14 differ depending on which of the following cases we are considering:

Assumption 3.12. *Let $a(\cdot,\cdot)$ be a bilinear form. In Lemma 3.13 and Theorem 3.14 we will distinguish between the following cases:*

i) The bilinear form $a(\cdot,\cdot)$ is bounded, i.e. there exists $C_1 > 0$ such that

$$|a(\varphi, w)| \leq C_1 \|\varphi\|_{\mathcal{V}} \|w\|_{\mathcal{V}} \quad \text{for all } \varphi, w \in \mathcal{V}. \tag{3.28}$$

Moreover, there exists real numbers, $\alpha > 0$ and β, such that the following Gårding-inequality

$$a(\varphi, \varphi) \geq \alpha \|\varphi\|^2_{\mathcal{V}} - \beta \|\varphi\|^2_H \quad \text{for all } \varphi \in \mathcal{V}^0 \tag{3.29}$$

holds.

ii) Let $a(\cdot,\cdot) := a_0(\cdot,\cdot) - a_1(\cdot,\cdot)$ be a bilinear form, where $a_0(\cdot,\cdot)$ is assumed to be symmetric. There exists constants $C_2 > 0$ and $C_3 > 0$ such that

$$|a_0(\varphi, w)| \leq C_2 \|\varphi\|_{\mathcal{V}} \|w\|_{\mathcal{V}} \quad \text{for all } \varphi, w \in \mathcal{V} \tag{3.30}$$

and

$$|a_1(\varphi, w)| \leq C_3 \|\varphi\|_{\mathcal{V}} \|w\|_H \quad \text{for all } \varphi, w \in \mathcal{V}. \tag{3.31}$$

Moreover the bilinear form $a_0(\cdot,\cdot)$ is coercive in $\mathcal{V}^0$, i.e. there exists $\alpha_0 > 0$ such that

$$a_0(\varphi, \varphi) \geq \alpha_0 \|\varphi\|_{\mathcal{V}} \|\varphi\|_{\mathcal{V}} \text{ for all } \varphi \in \mathcal{V}^0. \tag{3.32}$$

Let $a(\cdot,\cdot)$ satisfy i) in Assumption 3.12 and let $\psi \in H$ be given, then it is well known that the penalized problem has a unique solution $u^\varepsilon \in L^2(0,T;\mathcal{V}^0) \cap H^1(0,T;\mathcal{V}^*)$ (see [BL82, Chapter 3.2, Theorem 2.3]). The following Lemma summarizes some regularity results for the penalized problem. A very similar proof of Lemma 3.13 in the context of the transformed Black-Scholes and the integro-differential variational inequality with $x := \log(S/K)$ can be found in [Mem12] and [KTK17]. Note that statements i) and ii) in Lemma 3.13 fit into the context of many option pricing problems, since it is only assumed that the obstacle function lies in the space $\mathcal{V}$. In particular, (3.35) is satisfied for a convex obstacle function $\widetilde{\psi}$ and ensures that the solution is given by the obstacle in a subdomain of Ω, where the obstacle function is smooth enough.

Lemma 3.13. *(Regularity for the penalty problem)*

(i) Suppose that i) in Assumption 3.12 is satisfied. Assume also $f \in L^2(0,T;\mathcal{V}^*)$, $\widetilde{\psi} \in \mathcal{V}^0$ and $u_\psi \in \mathcal{V}$. Then the solution satisfies $u^\varepsilon \in L^\infty(0,T;H) \cap L^2(0,T;\mathcal{V}^0)$ and we have the estimate

$$\|u^\varepsilon\|_{L^\infty(0,T;H)} + \|u^\varepsilon\|_{L^2(0,T;\mathcal{V})} \lesssim \|\widetilde{\psi}\|_{\mathcal{V}} + \|u_\psi\|_{\mathcal{V}} + \|f\|_{L^2(0,T;\mathcal{V}^*)} \tag{3.33}$$

with constant independent of the penalty parameter ε.

ii) Suppose that ii) in Assumption 3.12 is satisfied. Assume also $f \in H^1(0,T;\mathcal{V}^*)$, $\widetilde{\psi} \in \mathcal{V}^0$ and $u_\psi \in \mathcal{V}$. Then the solution satisfies $u^\varepsilon \in H^1(0,T;H) \cap L^\infty(0,T;\mathcal{V}^0)$ and we have the estimate

$$\|\partial_\tau u^\varepsilon\|_{L^2(0,T;H)} + \|u^\varepsilon\|_{L^\infty(0,T;\mathcal{V})} \lesssim \|\widetilde{\psi}\|_{\mathcal{V}} + \|u_\psi\|_{\mathcal{V}} + \|f\|_{H^1(0,T;\mathcal{V}^*)} \tag{3.34}$$

with constant independent of the penalty parameter ε.

(iii) Suppose that ii) in Assumption 3.12 is satisfied. Assume also $f \in L^2(0,T,H)$, $f \in H^1(0,T;\mathcal{V}^*)$, $\widetilde{\psi} \in \mathcal{V}^0$ and $u_\psi \in H^2(\Omega)$. Moreover let $\widetilde{\psi}$ such that

$$-a_0(\widetilde{\psi}, \mathfrak{Q}u^\varepsilon(\tau)) \lesssim \|\widetilde{\psi}\|_{\mathcal{V}} \|\mathfrak{Q}u^\varepsilon(\tau)\|_H \quad \text{for a.e. } \tau \in (0,T]. \tag{3.35}$$

Then the solution satisfies $u^\varepsilon \in L^2(0,T;H^2(\Omega))$ and we have the estimate

$$\|u^\varepsilon\|_{L^2(0,T;H^2(\Omega))} \lesssim \|\widetilde{\psi}\|_{\mathcal{V}} + \|u_\psi\|_{H^2(\Omega)} + \|f\|_{H^1(0,T;\mathcal{V}^*)} + \|f\|_{L^2(0,T;H)} \tag{3.36}$$

with constant independent of the penalty parameter ε.

Proof. (i) Since $u^\varepsilon(\tau) \in \mathcal{V}^0$ for a.e. $\tau \in (0,T]$ and taking $\varphi = u^\varepsilon(\tau) - \widetilde{\psi}$ in (3.27) we have

$$\langle u_\tau^\varepsilon, u^\varepsilon\rangle + a(u^\varepsilon, u^\varepsilon) + \varepsilon^{-1}\left(\mathfrak{Q}u^\varepsilon, u^\varepsilon\right) = \langle u_\tau^\varepsilon, \widetilde{\psi}\rangle + a(u^\varepsilon, \widetilde{\psi}) + \langle f, u^\varepsilon - \widetilde{\psi}\rangle - a(u_\psi, u^\varepsilon).$$

Since $\langle u_\tau^\varepsilon, u^\varepsilon\rangle = \frac{1}{2}\frac{\partial}{\partial\tau}\|u^\varepsilon\|_H^2$ and $(\mathfrak{Q}u^\varepsilon, u^\varepsilon) = \|\mathfrak{Q}u^\varepsilon\|_H^2$ for a.e. $\tau \in (0,T]$ one has

$$\frac{1}{2}\frac{\partial}{\partial\tau}\|u^\varepsilon\|_H^2 + a(u^\varepsilon, u^\varepsilon) + \varepsilon^{-1}\|\mathfrak{Q}u^\varepsilon\|_H^2 = \langle u_\tau^\varepsilon, \widetilde{\psi}\rangle + a(u^\varepsilon, \widetilde{\psi}) + \langle f, u^\varepsilon - \widetilde{\psi}\rangle - a(u_\psi, u^\varepsilon).$$

Integrating between 0 and τ and observing $u^\varepsilon(0) = \widetilde{\psi}(S)$ yields

$$\begin{aligned}\frac{1}{2}\|u^\varepsilon(\tau)\|_H^2 + \int_0^\tau \left(a(u^\varepsilon(\mathfrak{r}), u^\varepsilon(\mathfrak{r}))\right)\mathrm{d}\mathfrak{r} + \varepsilon^{-1}\int_0^\tau \left(\|\mathfrak{Q}u^\varepsilon(\mathfrak{r})\|_H^2\right)\mathrm{d}\mathfrak{r}\\ = \langle u^\varepsilon, \widetilde{\psi}\rangle - \frac{1}{2}\|\widetilde{\psi}\|_H^2 + \int_0^\tau \left(a(u^\varepsilon(\mathfrak{r}), \widetilde{\psi}) + \langle f(\mathfrak{r}), u^\varepsilon(\mathfrak{r}) - \widetilde{\psi}\rangle - a(u_\psi, u^\varepsilon(\mathfrak{r}))\right)\mathrm{d}\mathfrak{r}\end{aligned}$$

for all $\tau \in (0,T]$. Then Cauchy-Schwarz inequality, the boundedness in (3.28) and the Gårding inequality in (3.29) implies

$$\begin{aligned}&\frac{1}{2}\|u^\varepsilon(\tau)\|_H^2 + \alpha\int_0^\tau \left(\|u^\varepsilon(\mathfrak{r})\|_{\mathcal{V}}^2\right)\mathrm{d}\mathfrak{r} + \varepsilon^{-1}\int_0^\tau \left(\|\mathfrak{Q}u^\varepsilon(\mathfrak{r})\|_H^2\right)\mathrm{d}\mathfrak{r}\\ \le&\|u^\varepsilon\|_H\|\widetilde{\psi}\|_H + \int_0^\tau \left(C(\|u_\psi\|_{\mathcal{V}} + \|\widetilde{\psi}\|_{\mathcal{V}})\|u^\varepsilon(\mathfrak{r})\|_{\mathcal{V}} + \|f(\mathfrak{r})\|_{\mathcal{V}^*}\|u(\mathfrak{r})\|_H\right)\mathrm{d}\mathfrak{r}\\ &+ \beta\int_0^\tau \left(\|u^\varepsilon(\mathfrak{r})\|_H^2\right)\mathrm{d}\mathfrak{r} \quad \text{for all } \tau \in (0,T].\end{aligned}$$

Applying Young's inequality and $\int_0^\tau \|u_\psi\|_{\mathcal{V}}^2\mathrm{d}\mathfrak{r} \le \sqrt{T}\|u_\psi\|_{\mathcal{V}}^2$ leads to

$$\begin{aligned}&\frac{1}{2}\|u^\varepsilon(\tau)\|_H^2 + \alpha\int_0^\tau \left(\|u^\varepsilon(\mathfrak{r})\|_{\mathcal{V}}^2\right)\mathrm{d}\mathfrak{r} + \varepsilon^{-1}\int_0^\tau \left(\|\mathfrak{Q}u^\varepsilon(\mathfrak{r})\|_H^2\right)\mathrm{d}\mathfrak{r} \le \frac{1}{4}\|u^\varepsilon\|_H^2\\ &+ \frac{\alpha}{2}\int_0^\tau \left(\|u^\varepsilon(\mathfrak{r})\|_{\mathcal{V}}^2\right)\mathrm{d}\mathfrak{r} + \tilde{C}\left(\|u_\psi\|_{\mathcal{V}}^2 + \|\widetilde{\psi}\|_{\mathcal{V}}^2 + \int_0^\tau \left(\|f(\mathfrak{r})\|_{\mathcal{V}^*}^2 + \|u^\varepsilon(\mathfrak{r})\|_H^2\right)\mathrm{d}\mathfrak{r}\right)\end{aligned}$$

with $\tilde{C} := \tilde{C}(T)$ and for all $\tau \in (0,T]$. Rearranging yields

$$\begin{aligned}&\frac{1}{4}\|u^\varepsilon(\tau)\|_H^2 + \frac{\alpha}{2}\int_0^\tau \left(\|u^\varepsilon(\mathfrak{r})\|_{\mathcal{V}}^2\right)\mathrm{d}\mathfrak{r} + \varepsilon^{-1}\int_0^\tau \left(\|\mathfrak{Q}u^\varepsilon(\mathfrak{r})\|_H^2\right)\mathrm{d}\mathfrak{r} \le\\ &\qquad \tilde{C}\left(\|u_\psi\|_{\mathcal{V}}^2 + \|\widetilde{\psi}\|_{\mathcal{V}}^2 + \int_0^\tau \left(\|f(\mathfrak{r})\|_{\mathcal{V}^*}^2 + \|u^\varepsilon(\mathfrak{r})\|_H^2\right)\mathrm{d}\mathfrak{r}\right) \quad \text{for all } \tau \in (0,T].\end{aligned}$$

Applying Gronwall's inequality in integral form and $\int_0^\tau \|u_\psi\|_{\mathcal{V}}^2\mathrm{d}\mathfrak{r} \le \sqrt{T}\|u_\psi\|_{\mathcal{V}}^2$ leads to

$$\begin{aligned}&\|u^\varepsilon(\tau)\|_H^2 + \int_0^\tau \left(\|u^\varepsilon(\mathfrak{r})\|_{\mathcal{V}}^2\right)\mathrm{d}\mathfrak{r} + \varepsilon^{-1}\int_0^\tau \left(\|\mathfrak{Q}u^\varepsilon(\mathfrak{r})\|_H^2\right)\mathrm{d}\mathfrak{r}\\ &\qquad \le \tilde{C}\left(\|u_\psi\|_{\mathcal{V}}^2 + \|\widetilde{\psi}\|_{\mathcal{V}}^2 + \int_0^\tau \|f(\mathfrak{r})\|_{\mathcal{V}^*}^2\mathrm{d}\mathfrak{r}\right) \qquad (3.37)\end{aligned}$$

for all $\tau \in (0,T]$ and

$$\|u^\varepsilon\|_{L^\infty(0,T;H)} + \|u^\varepsilon\|_{L^2(0,T;\mathcal{V})} \leq \tilde{C}(\|u_\psi\|_{\mathcal{V}} + \|\widetilde{\psi}\|_{\mathcal{V}} + \|f\|_{L^2(0,T;\mathcal{V}^*)}) \tag{3.38}$$

completes the proof of (i).
To simplify the notation for the proof of ii), we will frequently use $\widehat{u}(\tau) := \int_0^\tau u(\mathfrak{r})\mathrm{d}\mathfrak{r}$ and $\widehat{\mathfrak{Q}u}(\tau) := \int_0^\tau \mathfrak{Q}u(\mathfrak{r})\mathrm{d}\mathfrak{r}$. First, we rewrite the homogenized penalty problem in (3.27) as

$$\left\langle \frac{\partial u^\varepsilon}{\partial \tau}, \varphi \right\rangle + a_0(u^\varepsilon(\tau),\varphi) + \varepsilon^{-1}(\mathfrak{Q}u^\varepsilon,\varphi) = a_1(u^\varepsilon(\tau),\varphi) + \langle f,\varphi\rangle - a(u_\psi,\varphi) \text{ for a.e. } \tau \in (0,T].$$

Then integration between 0 and τ and using $\int_0^\tau a_0(u^\varepsilon(\mathfrak{r}),\varphi))\mathrm{d}\mathfrak{r} = a_0(\widehat{u^\varepsilon}(\tau),\varphi))$ and $u^\varepsilon(0) = \widetilde{\psi}$ yields

$$\begin{aligned}\langle u^\varepsilon(\tau),\varphi\rangle + a_0(\widehat{u^\varepsilon}(\tau),\varphi)) + \varepsilon^{-1}(\widehat{\mathfrak{Q}u^\varepsilon}(\tau),\varphi) =& \langle\widetilde{\psi},\varphi\rangle + \int_0^\tau a_1(u^\varepsilon(\mathfrak{r}),\varphi)\mathrm{d}\mathfrak{r} \\ &+ \langle\widehat{f}(\tau),\varphi\rangle - a(\widehat{u_\psi}(\tau),\varphi)\end{aligned} \tag{3.39}$$

for all $\tau \in (0,T]$. Then the relation above also implies

$$\begin{aligned}\langle u^\varepsilon(\tau-s),\varphi\rangle + a_0(\widehat{u^\varepsilon}(\tau-s),\varphi)) + \varepsilon^{-1}(\widehat{\mathfrak{Q}u^\varepsilon}(\tau-s),\varphi) =& \langle\widetilde{\psi},\varphi\rangle + \int_0^{\tau-s} a_1(u^\varepsilon(\mathfrak{r}),\varphi)\mathrm{d}\mathfrak{r} \\ &+ \langle\widehat{f}(\tau-s),\varphi\rangle - a(\widehat{u_\psi}(\tau-s),\varphi)\end{aligned} \tag{3.40}$$

for all $\tau \in (s,T]$. Subtracting (3.40) from (3.39) and dividing by s leads to

$$\begin{aligned}\langle \delta_s u^\varepsilon(\tau),\varphi\rangle + a_0(\delta_s\widehat{u^\varepsilon}(\tau),\varphi)) + \varepsilon^{-1}(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau),\varphi) =& \frac{1}{s}\int_\tau^{\tau-s} a_1(u^\varepsilon(\mathfrak{r}),\varphi)\mathrm{d}\mathfrak{r} \\ &+ \langle\delta_s\widehat{f}(\tau),\varphi\rangle - a(\delta_s\widehat{u_\psi}(\tau),\varphi),\end{aligned} \tag{3.41}$$

where $\delta_s u^\varepsilon(\tau) := \frac{u^\varepsilon(\tau)-u^\varepsilon(\tau-s)}{s}$ denotes the difference quotient.
Now we take $\varphi = \delta_s u^\varepsilon(\tau) = \partial_\tau(\delta_s\widehat{u^\varepsilon}(\tau))$ in the above relation, which results in

$$\begin{aligned}&\|\delta_s u^\varepsilon(\tau)\|_H^2 + a_0\left(\delta_s\widehat{u^\varepsilon}(\tau), \partial_\tau(\delta_s\widehat{u^\varepsilon}(\tau))\right) + \varepsilon^{-1}(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau), \partial_\tau(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau))) \\ &= \varepsilon^{-1}(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau), \partial_\tau(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau)) - \delta_s u^\varepsilon(\tau)) + \frac{1}{s}\int_\tau^{\tau-s} a_1(u^\varepsilon(\mathfrak{r}),\delta_s u^\varepsilon(\tau))\mathrm{d}\mathfrak{r} \\ &\quad + \langle\delta_s\widehat{f}(\tau),\delta_s u^\varepsilon(\tau)\rangle - a(\delta_s\widehat{u_\psi}(\tau),\delta_s u^\varepsilon(\tau)).\end{aligned} \tag{3.42}$$

Since $a_0(\cdot,\cdot)$ is symmetric one has

$$a_0\left(\delta_s\widehat{u^\varepsilon}(\tau), \partial_\tau(\delta_s\widehat{u^\varepsilon}(\tau))\right) = \frac{1}{2}\frac{\partial}{\partial\tau}a_0\left(\delta_s\widehat{u^\varepsilon}(\tau), \delta_s\widehat{u^\varepsilon}(\tau)\right) \tag{3.43}$$

and for the penalty term one observes

$$\varepsilon^{-1}(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau), \partial_\tau(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau))) = \frac{\varepsilon^{-1}}{2}\frac{\partial}{\partial\tau}\|\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau)\|_H^2. \tag{3.44}$$

Then relation (3.42) together with (3.43) and (3.44) yields

$$\begin{aligned}
&\|\delta_s u^\varepsilon(\tau)\|_H^2 + \frac{1}{2}\frac{\partial}{\partial\tau} a_0\left(\delta_s\widehat{u^\varepsilon}(\tau), \delta_s\widehat{u^\varepsilon}(\tau)\right) + \frac{\varepsilon^{-1}}{2}\frac{\partial}{\partial\tau}\|\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau)\|_H^2 \\
&= \varepsilon^{-1}(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau), \partial_\tau(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau)) - \delta_s u^\varepsilon(\tau)) + \frac{1}{s}\int_\tau^{\tau-s} a_1(u^\varepsilon(\mathfrak{r}), \delta_s u^\varepsilon(\tau))\mathrm{d}\mathfrak{r} \\
&\quad + \langle\delta_s\widehat{f}(\tau), \delta_s u^\varepsilon(\tau)\rangle - a(\delta_s\widehat{u_\psi}(\tau), \delta_s u^\varepsilon(\tau)).
\end{aligned} \tag{3.45}$$

Further applying the boundedness in (3.31) and Young's inequality yields

$$\frac{1}{s}\int_\tau^{\tau-s} a_1(u^\varepsilon(\mathfrak{r}), \delta_s u^\varepsilon(\tau))\mathrm{d}\mathfrak{r} \leq C\frac{1}{s}\int_\tau^{\tau-s}\|u^\varepsilon(\mathfrak{r})\|_{\mathcal{V}}^2\mathrm{d}\mathfrak{r} + \frac{1}{2}\|\delta_s u^\varepsilon(\tau))\|_H^2. \tag{3.46}$$

Integration between s and τ in (3.45) and applying (3.46) implies

$$\begin{aligned}
&\int_s^\tau \|\delta_s u^\varepsilon(\mathfrak{s})\|_H^2\mathrm{d}\mathfrak{s} + a_0\left(\delta_s\widehat{u^\varepsilon}(\tau), \delta_s\widehat{u^\varepsilon}(\tau)\right) + \varepsilon^{-1}\|\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau)\|_H^2 \\
&\leq a_0\left(\delta_s\widehat{u^\varepsilon}(s), \delta_s\widehat{u^\varepsilon}(s)\right) + 2\varepsilon^{-1}\int_s^\tau(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\mathfrak{s}), \partial_\tau(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\mathfrak{s})) - \delta_s u^\varepsilon(\mathfrak{s}))\mathrm{d}\mathfrak{s} \\
&\quad + \int_s^\tau C\frac{1}{s}\int_\tau^{\tau-s}\|u^\varepsilon(\mathfrak{r})\|_{\mathcal{V}}^2\mathrm{d}\mathfrak{r} + 2\langle\delta_s\widehat{f}(\mathfrak{s}), \delta_s u^\varepsilon(\mathfrak{s})\rangle - 2a(\delta_s\widehat{u_\psi}(\mathfrak{s}), \delta_s u^\varepsilon(\mathfrak{s}))\mathrm{d}\mathfrak{s}.
\end{aligned} \tag{3.47}$$

Applying the boundedness in (3.30), the coercivity in (3.32) of $a_0(\cdot,\cdot)$ and Young's inequality yields

$$\begin{aligned}
&\int_s^\tau \|\delta_s u^\varepsilon(\mathfrak{s})\|_H^2\mathrm{d}\mathfrak{s} + \alpha_0\|\delta_s\widehat{u^\varepsilon}(\tau)\|_{\mathcal{V}}^2 + \varepsilon^{-1}\|\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau)\|_H^2 \\
&\leq C\|\delta_s\widehat{u^\varepsilon}(s)\|_{\mathcal{V}} + 2\varepsilon^{-1}\int_s^\tau(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\mathfrak{s}), \partial_\tau(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\mathfrak{s})) - \delta_s u^\varepsilon(\mathfrak{s}))\mathrm{d}\mathfrak{s} \\
&\quad + \int_s^\tau C\frac{1}{s}\int_\tau^{\tau-s}\|u^\varepsilon(\mathfrak{r})\|_{\mathcal{V}}^2\mathrm{d}\mathfrak{r} + 2\langle\delta_s\widehat{f}(\mathfrak{s}), \delta_s u^\varepsilon(\mathfrak{s})\rangle - 2a(\delta_s\widehat{u_\psi}(\mathfrak{s}), \delta_s u^\varepsilon(\mathfrak{s}))\mathrm{d}\mathfrak{s}.
\end{aligned} \tag{3.48}$$

Then integration by parts and $\partial_{\mathfrak{s}}(\delta_s\widehat{u_\psi}(\mathfrak{s})) = \delta_s u_\psi = 0$ implies

$$\begin{aligned}
-\int_s^\tau 2a(\delta_s\widehat{u_\psi}(\mathfrak{s}), \delta_s u^\varepsilon(\mathfrak{s}))\mathrm{d}\mathfrak{s} &= -2a(\delta_s\widehat{u_\psi}(\tau), \delta_s\widehat{u^\varepsilon}(\tau)) + 2a(\delta_s\widehat{u_\psi}(s), \delta_s\widehat{u^\varepsilon}(s)) \\
&\quad + \int_s^\tau 2a(\partial_{\mathfrak{s}}(\delta_s\widehat{u_\psi}(\mathfrak{s})), \delta_s\widehat{u^\varepsilon}(\mathfrak{s}))\mathrm{d}\mathfrak{s} \\
&= -2a(\delta_s\widehat{u_\psi}(\tau), \delta_s\widehat{u^\varepsilon}(\tau)) + 2a(\delta_s\widehat{u_\psi}(s), \delta_s\widehat{u^\varepsilon}(s)) \\
&\leq C(\|\delta_s\widehat{u_\psi}(\tau)\|_{\mathcal{V}}^2 + \|\delta_s\widehat{u_\psi}(s)\|_{\mathcal{V}}^2 + \|\delta_s\widehat{u^\varepsilon}(s)\|_{\mathcal{V}}^2) \\
&\quad + \frac{\alpha_0}{4}\|\delta_s\widehat{u^\varepsilon}(\tau)\|_{\mathcal{V}}^2
\end{aligned} \tag{3.49}$$

and

$$\begin{aligned}\int_s^\tau 2\langle\delta_s\widehat{f}(\mathfrak{s}),\delta_s u^\varepsilon(\mathfrak{s})\rangle\mathrm{d}\mathfrak{s} =&2\langle\delta_s\widehat{f}(\tau),\delta_s\widehat{u^\varepsilon}(\tau)\rangle - 2\langle\delta_s\widehat{f}(s),\delta_s\widehat{u^\varepsilon}(s)\rangle\\ &-\int_s^\tau 2\langle\partial_{\mathfrak{s}}\delta_s\widehat{f}(\mathfrak{s}),\delta_s\widehat{u^\varepsilon}(\mathfrak{s})\rangle\\ \leq& C\left(\|\delta_s\widehat{f}(\tau)\|_{\mathcal{V}^*}^2+\|\delta_s\widehat{f}(s)\|_{\mathcal{V}^*}^2+\int_s^\tau\left(\|\delta_s f(\mathfrak{s})\|_{\mathcal{V}^*}^2+\|\delta_s\widehat{u}^\varepsilon(\mathfrak{s})\|_{\mathcal{V}}^2\right)\mathrm{d}\mathfrak{s}\right)\\ &+\frac{\alpha_0}{4}\|\delta_s\widehat{u}^\varepsilon\|_{\mathcal{V}^*}^2. \end{aligned}\tag{3.50}$$

Relation (3.48) together with (3.49) and (3.50) implies

$$\begin{aligned}&\int_s^\tau\|\delta_s u^\varepsilon(\mathfrak{s})\|_H^2\mathrm{d}\mathfrak{s}+\frac{\alpha_0}{2}\|\delta_s\widehat{u^\varepsilon}(\tau)\|_{\mathcal{V}}^2+\varepsilon^{-1}\|\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau)\|_H^2\\ &\leq C\|\delta_s\widehat{u^\varepsilon}(s)\|_{\mathcal{V}}+2\varepsilon^{-1}\int_s^\tau(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\mathfrak{s}),\partial_\tau(\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\mathfrak{s}))-\delta_s u^\varepsilon(\mathfrak{s}))\mathrm{d}\mathfrak{s}\\ &\quad+\int_s^\tau C\frac{1}{s}\int_\tau^{\tau-s}\|u^\varepsilon(\mathfrak{r})\|_{\mathcal{V}}^2\mathrm{d}\mathfrak{r}\mathrm{d}\mathfrak{s}+C(\|\delta_s\widehat{u_\psi}(\tau)\|_{\mathcal{V}}^2+\|\delta_s\widehat{u_\psi}(s)\|_{\mathcal{V}}^2)\\ &\quad+C\left(\|\delta_s\widehat{f}(\tau)\|_{\mathcal{V}^*}^2+\|\delta_s\widehat{f}(s)\|_{\mathcal{V}^*}^2+\int_s^\tau\left(\|\delta_s f(\mathfrak{s})\|_{\mathcal{V}^*}^2+\|\delta_s\widehat{u}^\varepsilon(\mathfrak{s})\|_{\mathcal{V}}^2\right)\mathrm{d}\mathfrak{s}\right).\end{aligned}\tag{3.51}$$

By passing to the limit as $s\to 0$ on both sides yields

$$\begin{aligned}&\lim_{s\to0}\left(\int_s^\tau\|\delta_s u^\varepsilon(\mathfrak{s})\|_H^2\mathrm{d}\mathfrak{s}+\frac{\alpha_0}{2}\|\delta_s\widehat{u^\varepsilon}(\tau)\|_{\mathcal{V}}^2+\varepsilon^{-1}\|\delta_s\widehat{\mathfrak{Q}u^\varepsilon}(\tau)\|_H^2\right)\\ &\leq C\|\widetilde{\psi}\|_{\mathcal{V}}^2+2\varepsilon^{-1}\int_0^\tau(\mathfrak{Q}u^\varepsilon(\mathfrak{s}),\partial_\tau(\mathfrak{Q}u^\varepsilon(\mathfrak{s}))-\partial_{\mathfrak{s}}u^\varepsilon(\mathfrak{s}))\mathrm{d}\mathfrak{s}+C(\|u^\varepsilon\|_{L^2(0,T;\mathcal{V})}^2+\|u_\psi\|_{\mathcal{V}}^2)\\ &\quad+C\|f\|_{H^1(0,T;\mathcal{V}^*)}^2\\ &=C(\|\widetilde{\psi}\|_{\mathcal{V}}^2+\|u^\varepsilon\|_{L^2(0,T;\mathcal{V})}^2+\|u_\psi\|_{\mathcal{V}}^2+\|f\|_{H^1(0,T;\mathcal{V}^*)}^2).\end{aligned}\tag{3.52}$$

Taking the limit and using relation (3.38) leads to

$$\|\partial_t u^\varepsilon\|_{L^2(0,T;H)}^2+\|u^\varepsilon\|_{L^\infty(0,T;\mathcal{V})}^2+\varepsilon^{-1}\|\mathfrak{Q}u^\varepsilon\|_H^2\leq C(\|\widetilde{\psi}\|_{\mathcal{V}}^2+\|u_\psi\|_{\mathcal{V}}^2+\|f\|_{H^1(0,T;\mathcal{V}^*)}^2),\tag{3.53}$$

which gives the desired result (ii).
To prove (iii) we set $\varphi=\mathfrak{Q}u^\varepsilon$ in (3.27). This yields

$$\left\langle\frac{\partial u^\varepsilon}{\partial\tau},\mathfrak{Q}u^\varepsilon\right\rangle+a_0(u^\varepsilon(\tau),\mathfrak{Q}u^\varepsilon)+\varepsilon^{-1}\|\mathfrak{Q}u^\varepsilon\|_H^2=a_1(u^\varepsilon(\tau),\mathfrak{Q}u^\varepsilon)-a(u_\psi,\mathfrak{Q}u^\varepsilon)+\langle f(\tau),\varphi\rangle\tag{3.54}$$

for a.e. $\tau\in(0,T]$.
Then, as $\left\langle\frac{\partial u^\varepsilon}{\partial\tau},\mathfrak{Q}u^\varepsilon\right\rangle=\frac{1}{2}\frac{\partial}{\partial\tau}\|\mathfrak{Q}u^\varepsilon\|_H^2$ and $a_0(\mathfrak{Q}u^\varepsilon,\mathfrak{Q}u^\varepsilon)=a_0(u^\varepsilon-\widetilde{\psi},\mathfrak{Q}u^\varepsilon)$ for a.e. $\tau\in(0,T]$, the relation above leads to

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial\tau}\|\mathfrak{Q}u^\varepsilon\|_H^2+a_0(\mathfrak{Q}u^\varepsilon,\mathfrak{Q}u^\varepsilon)+\varepsilon^{-1}\|\mathfrak{Q}u^\varepsilon\|_H^2=&-a_0(\widetilde{\psi},\mathfrak{Q}u^\varepsilon)+a_1(u^\varepsilon(\tau),\mathfrak{Q}u^\varepsilon)-a(u_\psi,\mathfrak{Q}u^\varepsilon)\\ &+\langle f(\tau),\varphi\rangle.\end{aligned}\tag{3.55}$$

The coercivity of $a_0(\cdot,\cdot)$ in (3.32), the boundedness of $a_1(\cdot,\cdot)$ in (3.31), $f(\tau) \in H$ and integrating between 0 and τ implies

$$\frac{1}{2}\|\mathfrak{Q}u^\varepsilon\|_H^2 + \alpha_0 \int_0^\tau \|\mathfrak{Q}u^\varepsilon\|_{\mathcal{V}}^2 \mathrm{d}\mathfrak{r} + \varepsilon^{-1} \int_0^\tau \|\mathfrak{Q}u^\varepsilon(\mathfrak{r})\|_H^2 \mathrm{d}\mathfrak{r} \leq -\int_0^\tau a_0(\widetilde{\psi}, \mathfrak{Q}u^\varepsilon(\mathfrak{r}))\mathrm{d}\mathfrak{r}$$
$$+C\left(\int_0^\tau (\|u^\varepsilon(\mathfrak{r})\|_{\mathcal{V}} + \|u_\psi\|_{H^2(\Omega)} + \|f(\mathfrak{r})\|_H)^2 \mathrm{d}\mathfrak{r}\right)^{\frac{1}{2}} \left(\int_0^\tau \|\mathfrak{Q}u^\varepsilon\|_H^2 \mathrm{d}\mathfrak{r}\right)^{\frac{1}{2}}. \quad (3.56)$$

Using the assumption in (3.35) and applying Cauchy-Schwarz inequality results in

$$-\int_0^\tau a_0(\widetilde{\psi}, \mathfrak{Q}u^\varepsilon(\mathfrak{r}))\mathrm{d}\mathfrak{r} \lesssim \int_0^\tau \|\widetilde{\psi}\|_{\mathcal{V}} \, \|\mathfrak{Q}u^\varepsilon(\mathfrak{r})\|_H \mathrm{d}\mathfrak{r}$$
$$\lesssim \sqrt{T}\|\widetilde{\psi}\|_{\mathcal{V}} \left(\int_0^\tau \|\mathfrak{Q}u^\varepsilon(\mathfrak{r})\|_H \mathrm{d}\mathfrak{r}\right)^{\frac{1}{2}}. \quad (3.57)$$

Next, relation (3.56) together with (3.57) and (3.38) implies

$$\|\mathfrak{Q}u^\varepsilon\|_{L^2(0,T;H)} \lesssim \varepsilon \left(\|u^\varepsilon\|_{L^2(0,T;\mathcal{V})} + \|u_\psi\|_{H^2(\Omega)} + \|\widetilde{\psi}\|_{\mathcal{V}} + \|f\|_{L^2(0,T;H)}\right)$$
$$\lesssim \varepsilon \left(\|u_\psi\|_{H^2(\Omega)} + \|\widetilde{\psi}\|_{\mathcal{V}} + \|f\|_{L^2(0,T;H)}\right). \quad (3.58)$$

It follows from the elliptic regularity theory that for $\Omega \subset \mathbb{R}$

$$a_0(u^\varepsilon, \varphi) = a_1(u^\varepsilon, \varphi)) - a(u_\psi, \varphi) - \varepsilon^{-1}(\mathfrak{Q}u^\varepsilon, \varphi) - \langle \partial_\tau u^\varepsilon, \varphi\rangle \quad \text{for all } \varphi \in \mathcal{V}^0 \quad (3.59)$$

with $f(\tau) \in H$, $u^\varepsilon(\tau) \in \mathcal{V}$, $u_\psi \in H^2(\Omega)$ and $\mathfrak{Q}u^\varepsilon(\tau)$, $\partial_\tau u^\varepsilon(\tau) \in H$ that there exists a unique solution $u^\varepsilon(\tau) \in H^2(\Omega)$ and the following estimates holds

$$\|u^\varepsilon(\tau)\|_{H^2(\Omega)} \lesssim \|u(\tau)\|_{\mathcal{V}} + \|u_\psi\|_{H^2(\Omega)} + \varepsilon^{-1}\|\mathfrak{Q}u^\varepsilon(\tau)\|_H + \|\partial_\tau u^\varepsilon\|_H + \|f(\tau)\|_H. \quad (3.60)$$

Finally, applying (3.53) and (3.58) to (3.60) gives the desired result in (iii). □

By passing to the limit in (3.27) as $\varepsilon \to 0$, it can be shown that the solution of the variational inequality satisfies the following regularity results.

Theorem 3.14. *(Regularity for parabolic variational inequalities)*

(i) Suppose that i) in Assumption 3.12 is satisfied. Assume also $f \in L^2(0,T;\mathcal{V}^*)$, $\widetilde{\psi} \in \mathcal{V}^0$ and $u_\psi \in \mathcal{V}$ then the solution satisfies $u \in L^\infty(0,T;H) \cap L^2(0,T;\mathcal{V}^0)$.

(ii) Suppose that ii) in Assumption 3.12 is satisfied. Assume also $f \in H^1(0,T;\mathcal{V}^*)$, $\widetilde{\psi} \in \mathcal{V}^0$ and $u_\psi \in \mathcal{V}$ then the solution satisfies $u \in H^1(0,T;H) \cap L^\infty(0,T;\mathcal{V}^0)$.

(iii) Suppose that ii) in Assumption 3.12 is satisfied. Moreover, let the obstacle $\widetilde{\psi}$ satisfy (3.35). Assume also $f \in L^2(0,T;H)$, $f \in H^1(0,T;\mathcal{V}^*)$, $\widetilde{\psi} \in \mathcal{V}^0$ and $u_\psi \in H^2(\Omega)$. Then the solution satisfies $u \in L^2(0,T;H^2(\Omega))$.

Proof. The proof consists of passing to the limit as $\varepsilon \to 0$ and proving that the solution is unique. It can be proved analogously to [BL82, Chapter 3.2]. □

Remark 3.15. *(Smoothness)*
In the variational equation framework ($\mathcal{K} = \mathcal{V}$) under suitable compatibility assumption (e.g. $u_0 = 0$, $f(\tau)$ vanishes near $\tau = 0$) the smoothness of the solution u is directly related to the smoothness of f, i.e. $f \in H^q(0,T;\mathcal{V}) \Rightarrow u \in H^q(0,T;\mathcal{V}) \cap H^{q+1}(0,T;\mathcal{V}^*)$. This is not true for variational inequalities ($\mathcal{K} \subset \mathcal{V}$). Already in the simple scalar case $\mathcal{V} = H = \mathbb{R}$ the derivative of the solution $u(\tau)$ with respect to time τ can have jump discontinuities (see e.g. [Bai89] on p.61). $\frac{\partial u}{\partial \tau} \in L^2(0,T;\mathcal{V}) \cap L^\infty(0,T;H^*)$ is the maximal smoothness, as $\frac{\partial u}{\partial \tau}$ cannot be continuous.

3.2. B-Splines

In this section we present an introduction to B-spline spaces and their basis functions. The goal of this section is to introduce the notation and to summarize relevant properties of B-splines for this work. For more detailed information on B-splines, we refer to [dB01, PT97, Sch07, KLM+17]. A short introduction with the most relevant properties of B-splines can also be found in [Mol16].
Let $\Theta := \{\theta_i\}_{i=1}^{n+k}$ be the ordered sequence of knots for fixed $k \in \mathbb{N}$, where we allow repetition of knots, that is

$$x_{\min} := \theta_1 = \ldots = \theta_k < \theta_{k+1} \leq \ldots \leq \ldots \leq \theta_n < \theta_{n+1} = \ldots = \theta_{n+k} := x_{\max} \tag{3.61}$$

where $\theta_1 = a$ and $\theta_{n+1} = b$. We say that a knot has multiplicity q_i, if it occurs exactly q_i times. The maximum multiplicity in the interior of $[x_{\min}, x_{\max}]$ we allow in this thesis is $k-1$. Then we define the B-splines $N_{i,k}(x)$ of order k for $i = 1, \ldots, n$ recursively as

$$\begin{aligned} N_{i,1}(x) &= \begin{cases} 1, & \text{if } x \in [\theta_i, \theta_{i+1}) \\ 0, & \text{else} \end{cases} \\ N_{i,k}(x) &= \frac{x - \theta_i}{\theta_{i+k-1} - \theta_i} N_{i,k-1}(x) + \frac{\theta_{i+k} - x}{\theta_{i+k} - \theta_{i+1}} N_{i+1,k-1}(x) \end{aligned} \tag{3.62}$$

for $\theta_{i+k-1} \neq \theta_i$ and $\theta_{i+k} \neq \theta_{i+1}$. Since $N_{i,1}(x) = 0$ if $\theta_i = \theta_{i+1}$, the quotient for the coinciding knots is set to zero.

Properties 3.16. *B-splines defined in (3.62) have the following properties*

- *Local support:* $\operatorname{supp}(N_{i,k}) \subset [\theta_i, \theta_{i+k}]$
- *Nonnegativity:* $N_{i,k} \geq 0$ *for all* $x \in [x_{\min}, x_{\max}]$.
- *Piecewise structure:* $N_{i,k}$ *are piecewise polynomials of order* k *on* $[\theta_j, \theta_{j+1}]$.
- *Smoothness: If* θ_j *is a knot of* $N_{i,k}$ *with multiplicity* q_j *then they have* $k - q_j - 1$ *continuous derivatives, i.e.* $N_{i,k} \in C^{k-q_j-1}(\theta_j)$.
- *Local partition of unity: B-spline functions constitute a partition of unity, that is* $\sum_{i=1}^{n} N_{i,k}(x) = 1$ *for all* $x \in (a, b)$.

A proof for these properties can be found for example in [KLM+17]. The space of splines spanned by the basis function $N_{i,k}$ is denoted by

$$\mathbb{S}_{k,\Theta} := \operatorname{span}\{N_{1,k}, \ldots, N_{n,k}\}, \tag{3.63}$$

where Θ is defined as in (3.61). In the following, we present some basic properties which are important for an efficient implementation of B-splines. These formulas are used for the implementation of the code developed for this thesis. Since B-Splines

are piecewise polynomials fulfilling the recurrence relation the derivatives can also be expressed recursively.

Corollary 3.17. *Let $N_{i,k}$ for $k \in \mathbb{N}$ and $i \in 1, \dots, n$ regarding the extended sequence of knots as defined in (3.61). Then the derivatives of B-splines are given by*

$$N'_{i,k}(x) = (k-1)\left(\frac{N_{i,k-1}(x)}{\theta_{i+k-1} - \theta_i} - \frac{N_{i+1,k-1}(x)}{\theta_{i+k} - \theta_{i+1}}\right). \tag{3.64}$$

Due to the recurrence relation and the local support of B-splines the point evaluation of splines can efficiently be implemented by the following Neville-like scheme (see e.g. [Sch07], Theorem 5.7).

Theorem 3.18. *Let*

$$\mathcal{S}(x) = \sum_{i=1}^{n} c_i N_{i,k}(x) \tag{3.65}$$

a spline regarding the extended sequence of knots defined in (3.61). Then the spline can expressed equivalently as

$$\mathcal{S}(x) = \sum_{i=\mathfrak{q}+1}^{n} c_i^{[\mathfrak{q}]}(x) N_{i,k-\mathfrak{q}}(x) \tag{3.66}$$

for $0 \leq \mathfrak{q} \leq k-1$, where

$$c_i^{[\mathfrak{q}]}(x) := \begin{cases} c_i & \text{if } \mathfrak{q} = 0 \\ \frac{x-\theta_i}{\theta_{i+k-\mathfrak{q}}-\theta_i} c_i^{[\mathfrak{q}-1]}(x) + \frac{\theta_{i+k-\mathfrak{q}}-x}{\theta_{i+k-\mathfrak{q}}-\theta_i} c_{i-1}^{[\mathfrak{q}-1]}(x) & \text{if } \mathfrak{q} > 0 \\ 0 & \text{if } \theta_{i+k-\mathfrak{q}} = \theta_i. \end{cases} \tag{3.67}$$

In particular, if $\tilde{x} \in [\theta_i, \theta_{i+1})$ the spline can be evaluated by

$$\mathcal{S}(\tilde{x}) = c_i^{[k-1]}(\tilde{x}). \tag{3.68}$$

At this point it should be mentioned that the Matlab implementation of the Neville-scheme as defined in Theorem 3.18 and the recurrence formula for the derivatives of B-splines as defined in Corollary 3.17 originally stems from [Mol16].
Due to the recurrence relation for the derivatives of B-splines one can also implement the evaluation of the derivatives of splines by the following Neville-like scheme. The scheme can also be used for repeating knots (c.f. [DR08]).

Corollary 3.19. *Let $\mathcal{S}^{(\mathfrak{q})}(x)$ the $\mathfrak{q}$'th derivative of the spline $\mathcal{S}$ then the derivative can be expressed equivalently by*

$$\mathcal{S}^{(\mathfrak{q})}(x) = (k-1)\dots(k-\mathfrak{q}) \sum_{i=\mathfrak{q}+1}^{n} c_i^{(\mathfrak{q})} N_{i,k-\mathfrak{q}}(x) \tag{3.69}$$

with

$$c_i^{(\mathfrak{q})} := \begin{cases} c_i & \text{if } \mathfrak{q} = 0 \\ \frac{c_i^{(\mathfrak{q}-1)} - c_{i-1}^{(\mathfrak{q}-1)}}{\theta_{i+k-\mathfrak{q}} - \theta_i} & \text{if } 0 < \mathfrak{q} \leq k-2. \end{cases} \tag{3.70}$$

After the recursive evaluation of $c_i^{(\mathfrak{q})}$ *one can compute* $\mathcal{S}^{(\mathfrak{q})}(x)$ *by Theorem 3.18.*

Since we also consider a variational inequality on a higher dimensional spatial rectangular domain, we introduce multidimensional tensor-product splines. Therefore, let d knot vectors $\Theta^{(j)}$ with $1 \leq j \leq d$ be given. Associated with the knot sequence $\Theta^{(j)}$ similar defined as in (3.61) with uniform grid size $h_j := \max\{\theta_{i+1}^{(j)} - \theta_i^{(j)}\}$ there is a mesh Θ, that is a partition of $\Omega \subset \mathbb{R}^d$ into d-dimensional open knot sequences $\Theta = \{\otimes_{j=1}^d \Theta^{(j)}\}$. For the multidimensional case, it is assumed in this thesis that the grid sizes h_j are uniform but repetitions of knots are possible. We define by $\mathbb{S}_{k,\Theta^{(j)}}$ the space of splines with respect to the knot sequence $\Theta^{(j)}$. Then the d-dimensional tensor-product spline space of order k is given by

$$\mathbb{S}_{k,\Theta}^d := \otimes_{j=1}^d \mathbb{S}_{k,\Theta^{(j)}}. \tag{3.71}$$

To this end we consider tensor-product splines of order k with given expansion coefficients $c_{i_1,\ldots,i_d}$

$$\mathcal{S}(x) := \mathcal{S}(x_1, \ldots, x_d) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} c_{i_1,\ldots,i_d} N_{i_1,k,\Theta_1}(x_1) \cdots N_{i_d,k,\Theta_d}(x_d) \tag{3.72}$$

with respect to the knot sequence $\Theta^{(j)}$ with grid size h_j for $j = 1, \ldots, d$. Properties for one dimensional splines can be extended straightforwardly to tensor products in most cases by introducing a Kronecker product.
We conclude the section by some global (optimal) approximation properties for splines in Sobolev spaces. Considering integer Sobolev spaces, the following local error estimate is proved for $d-$dimensional splines in [Sch07, Theorem 13.20]. A detailed proof for one dimensional splines can also be found in [KLM$^+$17].

Theorem 3.20. *Let* $\Omega := \otimes_{j=1}^d (a_i, b_i)$ *a rectangular domain and let* $f \in H^{\mathfrak{m}}(\Omega)$ *with* $\mathfrak{m} \in \mathbb{N}$. *Then there exists a tensor-product spline* $\mathcal{S} \in \mathbb{S}_{k,\Theta}^d$ *with* $\mathfrak{q} \leq \mathfrak{m} \leq k$ *such that*

$$\|f - \mathcal{S}\|_{H^{\mathfrak{q}}(\Omega)} \lesssim h^{\mathfrak{m}-\mathfrak{q}} \left\| \frac{\partial^{\mathfrak{m}} f}{\partial x_j^{\mathfrak{m}}} \right\|_{L^2(\Omega)}, \tag{3.73}$$

where $h = \max\limits_{1 \leq j \leq d} \{h_j\}$ *with* $h_j := \max\limits_{k \leq i \leq n^{(j)}} \{\theta_{i+1}^{(j)} - \theta_i^{(j)}\}$ *and constants only depending on* $\mathfrak{m}, k$ *and* d.

In the context of variational inequalities we also need some global approximation properties for tensor-product splines in the Sobolev-Slobodeckij space $W^{\mathfrak{s},p}(\Omega)$ with non integer $\mathfrak{s} > 0$. These estimate can be obtained from [Sch07, Theorem 6.31 and (6.60)], where an approximation estimate for splines in Besov spaces is established. Due to some

known embedding theorems for Besov spaces (for more details of the embeddings see [Tri78]) one can derive the following estimates in Sobolev norms from the corresponding estimate in Besov spaces.

Theorem 3.21. *Let $\Omega := \otimes_{j=1}^{d}(a_i, b_i)$ a rectangular domain and $1 \leq q \leq \infty$. Suppose $0 \leq \mathfrak{r} < \lfloor \mathfrak{s} - 1 \rfloor$ and $f \in W^{q,\mathfrak{s}}$ with non-integer $\mathfrak{s} > 0$. Then there exists a tensor-product spline approximation $\mathcal{S} \in \mathbb{S}_{k,\Theta}^{d}$ such that*

$$\|f - \mathcal{S}\|_{W^{q,\mathfrak{r}}(\Omega)} \lesssim h^{(\mathfrak{s}-\mathfrak{r})} \|f\|_{W^{q,\mathfrak{s}}(\Omega)}, \tag{3.74}$$

where $h = \max\limits_{1\leq j\leq d}\{h_j\}$ with $h_j := \max\limits_{k\leq i\leq n^{(j)}}\{\theta_{i+1}^{(j)} - \theta_i^{(j)}\}$.

3.3. A priori estimates for elliptic variational inequalities

In this section, we consider a priori estimates for elliptic variational inequalities. Since one cannot solve the problem in an infinite space $\mathcal{V}$ we introduce a finite dimensional space $\mathcal{V}_h \subset \mathcal{V}$ with (tensor product) B-spline basis functions. The corresponding approximation of the convex set $\mathcal{K}$ is denoted by $\mathcal{K}_h$. If all assumptions for the a priori estimates are satisfied, the following results can be applied to the semi-discrete formulation of an parabolic variational inequality. A priori estimates for elliptic variational inequalities were often discussed for finite elements. For an introduction to finite elements, see [JL13]. The first a priori estimate for elliptic variational inequalities with homogeneous boundary condition, where an approximate convex set $\mathcal{K}_h$ needs not to be contained in $\mathcal{K}$ has been established by Falk [Fal74]. He proves an optimal convergence rate in the $H^1(\Omega)$-norm for piecewise linear elements. Later Brezzi, Raviart and Hager [BHR77] have extended the result for variational inequalities with nonhomogeneous boundary condition for linear and quadratic elements. We concentrate on the case $\mathcal{K}_h \subset \mathcal{K}$ on an interval $\Omega \subset \mathbb{R}$ or a rectangular domain $\Omega \subset \mathbb{R}^2$. Let $\mathcal{V}^0$ the space of once weak differentiable function with homogeneous boundary condition, $\mathcal{V}^*$ the associated dual space and $\mathcal{K}$ be a non-empty closed subset of $\mathcal{V}$. Let $a(\cdot,\cdot)$ be a continuous bilinear form on $\mathcal{V}$ and coercive on $\mathcal{V}^0$ with $\mathcal{V}^0 \subset \mathcal{V}$. Then we consider the following elliptic variational inequality:
Find $y \in \mathcal{K} := \{\varphi \in \mathcal{V} : \varphi \geq \psi \text{ in } \Omega,\ \varphi = \psi \text{ on } \Upsilon\}$ such that

$$a(y, \varphi - y) \geq \langle f, \varphi - y\rangle \quad \text{for all } \varphi \in \mathcal{K}, \tag{3.75}$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $\mathcal{V}$ and $\mathcal{V}^*$. We also use the notation for the linear mapping $\mathcal{A} : \mathcal{V} \to \mathcal{V}^*$ with $\langle \mathcal{A}y, \varphi\rangle = a(y, \varphi)$. We denote by $\Omega_f \subset \Omega$ and $\Omega \setminus \Omega_f$ the parts of Ω, where the constraints are binding $y = \psi$ and nonbinding $y > \psi$ respectively. In the B-spline Galerkin discretization, we replace the solution space by $\mathcal{V}_h$ with B-splines or tensor product B-splines of order $k = 2, 3, 4$. For an obstacle function and boundary conditions given by the same function ψ one can construct a discrete solution of the form $y_h := u_h + \psi$, which lies in the following discrete convex set

$$\mathcal{K}_h := \{\varphi_h \in \mathcal{V}_h : \varphi_h \geq \psi \text{ in } \Omega,\ \varphi = \psi \text{ on } \Upsilon\} \subset \mathcal{K}. \tag{3.76}$$

In the B-spline Galerkin method we then search the discrete solution $y_h \in \mathcal{K}_h$ such that

$$a(y_h, \varphi_h - y_h) \geq \langle f, \varphi_h - y_h \rangle \quad \text{for all } \varphi_h \in \mathcal{K}_h. \tag{3.77}$$

The next lemma can be seen as an generalised formulation of the well known Céa Lemma for variational equations ($\mathcal{K} = \mathcal{V}$). Note, that in comparison to [BHR77, Fal74] we consider the simpler case $\mathcal{K}_h \subset \mathcal{K}$ (conforming approximation).

Lemma 3.22. *Let $a(\cdot,\cdot)$ be a bounded bilinear form on $\mathcal{V} \times \mathcal{V}$ which is coercive on $\mathcal{V}^0$. Let y and y_h be the solutions of (3.75) and (3.77), respectively. Then one can estimate*

$$\|y - y_h\|_{\mathcal{V}}^2 \lesssim \inf_{\varphi_h \in \mathcal{K}_h} \Big(\|\varphi_h - y\|_{\mathcal{V}}^2 + \langle \mathcal{A}y - f, \varphi_h - y \rangle \Big) \quad \textit{for all } \varphi_h \in \mathcal{K}_h \subset \mathcal{K} \tag{3.78}$$

and if $\mathcal{A}y - f \in H$, then

$$\|y - y_h\|_{\mathcal{V}}^2 \lesssim \inf_{\varphi_h \in \mathcal{K}_h} \Big(\|\varphi_h - y\|_{\mathcal{V}}^2 + \|\mathcal{A}y - f\|_H \|\varphi_h - y\|_H \Big) \quad \textit{for all } \varphi_h \in \mathcal{K}_h \subset \mathcal{K}. \tag{3.79}$$

Proof. Since $\mathcal{K}_h \subset \mathcal{K}$ (conforming approximation) we have the following variational inequalities for problem (3.75) and (3.77),

$$\langle \mathcal{A}y, \varphi_h - y \rangle \geq \langle f, \varphi_h - y \rangle \quad \text{for all } \varphi_h \in \mathcal{K}_h, \tag{3.80}$$

$$\langle \mathcal{A}y_h, \varphi_h - y_h \rangle \geq (f, \varphi_h - y_h) \quad \text{for all } \varphi_h \in \mathcal{K}_h. \tag{3.81}$$

Since $\mathcal{K}_h \subset \mathcal{K}$ this yields

$$\langle \mathcal{A}y - f, y - y_h \rangle \leq 0 \quad \text{and} \quad \langle \mathcal{A}y_h - f, y_h - \varphi_h \rangle \leq 0. \tag{3.82}$$

Employing the estimates in (3.82) one has

$$\begin{aligned} \mathcal{E} &:= \langle \mathcal{A}(y - y_h), y - y_h \rangle = \langle \mathcal{A}y - f, y - y_h \rangle - \langle \mathcal{A}y_h - f, y - y_h \rangle \leq -\langle \mathcal{A}y_h - f, y - y_h \rangle \\ &= \langle \mathcal{A}y_h - f, y_h - \varphi_h \rangle + \langle \mathcal{A}y_h - f, \varphi_h - y \rangle \leq \langle \mathcal{A}y_h - f, \varphi_h - y \rangle \\ &= -\langle \mathcal{A}(y - y_h), \varphi_h - y \rangle + \langle \mathcal{A}y - f, \varphi_h - y \rangle. \end{aligned}$$

Exploiting the coercivity and the boundedness of the operator results in

$$\alpha \|y - y_h\|_{\mathcal{V}}^2 \leq \mathcal{E} \leq C \|y - y_h\|_{\mathcal{V}} \|\varphi_h - y\|_{\mathcal{V}} + \langle \mathcal{A}y - f, \varphi_h - y \rangle. \tag{3.83}$$

Finally, Young's inequality leads to

$$C \|y - y_h\|_{\mathcal{V}} \|\varphi_h - y\|_{\mathcal{V}} \leq \frac{\alpha}{2} \|y - y_h\|_{\mathcal{V}}^2 + \frac{C^2}{2\alpha} \|\varphi_h - y\|_{\mathcal{V}}^2.$$

This relation together with (3.83) results in

$$\frac{\alpha}{2} \|y - y_h\|_{\mathcal{V}}^2 \leq \frac{C^2}{2\alpha} \|\varphi_h - y\|_{\mathcal{V}}^2 + \langle \mathcal{A}y - f, \varphi_h - y \rangle, \tag{3.84}$$

whence (3.78) is satisfied. Finally, applying Cauchy-Schwarz and $\mathcal{A}y - f \in H$

gives us the desired result in (3.79). □

Remark 3.23. *If $\mathcal{A}y - f = 0$ ($\mathcal{K} = \mathcal{V}$) is satisfied the second term in (3.79) vanishes. Then Theorem 3.22 leads to the well known Céa Lemma for an elliptic variational equation.*

The next theorem gives us a priori estimates for linear B-splines. Note that there is no need to assume that the obstacle lies in $H^2(\Omega)$ on the whole domain Ω as in [BHR77]. We only assume that the solution satisfies $y \in H^2(\Omega)$. This implies that in Ω_f, where the solution is given by the obstacle, the obstacle satisfies $\psi \in H^2(\Omega_f)$. An example, where the obstacle is not in $H^2(\Omega)$ on the whole domain but the solution lies in $H^2(\Omega)$ is the semi-discrete Black-Scholes variational inequality (c.f. Theorem 4.5) or the semi-discrete Heston variational inequality (c.f. Theorem 5.11).

Theorem 3.24. *Let $y \in H^2(\Omega)$ be the solution of the variational inequality in (3.75). If $f \in L^2(\Omega)$, then the continuous piecewise linear approximation $y_h \in \mathcal{V}_h$ ($k = 2$) satisfies*

$$\|y - y_h\|_{\mathcal{V}} = \mathcal{O}(h). \tag{3.85}$$

Proof. Recall the result from Lemma 3.22. Since $\mathcal{A}y - f \in L_2(\Omega)$ for $y \in H^2(\Omega)$ and $f \in L_2(\Omega)$ it yields

$$\|y - y_h\|_{\mathcal{V}}^2 \lesssim \|\varphi_h - y\|_{\mathcal{V}}^2 + \|\mathcal{A}y - f\|_{L_2(\Omega)}\|\varphi_h - y\|_{L^2(\Omega)} \quad \text{for all } \varphi_h \in \mathcal{K}_h \subset \mathcal{K}. \tag{3.86}$$

The approximation theorem for piecewise linear splines in Theorem 3.20 implies

$$\|I_h y - y\|_{\mathcal{V}} \lesssim h\|y\|_{H^2(\Omega)} \quad \text{and} \quad \|I_h y - y\|_{L_2(\Omega)} \lesssim h^2\|y\|_{H^2(\Omega)}. \tag{3.87}$$

By choosing $\varphi_h = I_h y$ in (3.86) it is estimated

$$\|y - y_h\|_{\mathcal{V}}^2 \lesssim \|I_h y - y\|_{\mathcal{V}}^2 + \|I_h y - y\|_{L^2(\Omega)} \lesssim h^2.$$

Extracting the root gives us the result. □

In the next theorem, we consider estimates for piecewise quadratic B-splines. A very similar result for the more general case of nonconforming approximation with piecewise quadratic finite elements has been established by Brezzi, Hager and Raviart in [BHR77]. The proof of [BHR77] is modified such that $\psi \in H^1(\Omega) \cap W^{2,\infty}(\Omega_f)$ instead of $\psi \in W^{2,\infty}(\Omega) \cap H^3(\Omega)$ is assumed.

Theorem 3.25. *Let $y \in W^{l,p}(\Omega)$ for all $1 < p < \infty$ and $l < 2 + \frac{1}{p}$ the solution regarding the variational inequality in (3.75). If $f \in L^\infty(\Omega)$ and $\psi \in H^1(\Omega) \cap W^{2,\infty}(\Omega_f)$ then the continuous piecewise quadratic B-spline approximation $y_h \in \mathcal{V}_h$ ($k = 3$) satisfies*

$$\|y - y_h\|_{\mathcal{V}} = \mathcal{O}(h^{\frac{3}{2}-\epsilon}). \tag{3.88}$$

Proof. Since $\mathcal{A}y - f = 0$ on $\Omega \setminus \Omega_f$ and $\mathcal{A}y - f = \mathcal{A}\psi - f \in L^\infty$ in Ω_f, we have $\mathcal{A}y - f \in L^\infty(\Omega)$. Applying the Hölder inequality leads to

$$\begin{aligned}\|y - y_h\|_{\mathcal{V}}^2 &\lesssim \|\varphi_h - y\|_{\mathcal{V}}^2 + \langle \mathcal{A}y - f, \varphi_h - y\rangle \\ &\leq \|y - \varphi_h\|_{\mathcal{V}}^2 + \|\mathcal{A}y - f\|_{L^\infty(\Omega)}\|\varphi_h - y\|_{L^1(\Omega)}. \end{aligned} \tag{3.89}$$

Due to the regularity of the solution $y \in H^{5/2-\epsilon}(\Omega)$ and $y \in W^{3-\epsilon_1,1-\epsilon_2}(\Omega)$ the approximation result for quadratic splines in Theorem 3.21 with $\mathfrak{r} = 1, \mathfrak{s} = 2.5 - \varepsilon, q = 2$ and $\mathfrak{r} = 0, \mathfrak{s} = 3 - \epsilon_1, q = 1 + \epsilon_1$ implies

$$\|I_h y - y\|_{\mathcal{V}} \lesssim h^{\frac{3}{2}-\epsilon}\|y\|_{H^{5/2-\epsilon}(\Omega)} \quad \text{and} \quad \|y - I_h y\|_{L^1(\Omega)} \lesssim h^{3-\epsilon}\|y\|_{W^{3-\epsilon_1,1+\epsilon_2}(\Omega)}. \tag{3.90}$$

By choosing $\varphi_h = I_h y$ in (3.89) one estimates

$$\|y - y_h\|_{\mathcal{V}}^2 \lesssim \|I_h y - y\|_{\mathcal{V}}^2 + \|I_h y - y\|_{L^1(\Omega)} \lesssim h^{3-\epsilon}.$$

Extracting the root gives us the result. □

Since one cannot expected a higher regularity than $y \in W^{\mathfrak{s},p}$ for all $\mathfrak{s} < 2 + \frac{1}{p}$ for variational inequalities, increasing the order of the B-spline basis function for $k > 3$ does not improve the global error estimate in the $H^1(\Omega)$-norm for uniform grid sizes in general.

4. Pricing American Put and Call option with Black-Scholes model

4.1. Well-posedness and regularity

In this section, we present some well-posedness and global regularity results for the Black-Scholes variational inequality formulated in Problem 2.9. The results of the next section can also be found in [AP05, Ach05, DL19a, DL19b]. However, we give a slightly modified proof based on the results for parabolic variational inequalities from Section 3.1.2.

Assumption 4.1. *We make the following assumption on the data σ, D_0 and r: The volatility σ and interest rate r are positive and bounded, i.e.*

$$0 < \sigma < \infty \quad 0 < r < \infty. \tag{4.1}$$

The dividend yield D_0 is assumed to be non-negative in the case of a put option, strictly positive in the case of a call option and bounded in both cases, i.e.

$$0 \leq D_0 < \infty \text{ (put) and } 0 < D_0 < \infty \text{ (call).} \tag{4.2}$$

Let us consider the weighted Sobolev space $\mathcal{V}$ from (2.70). From [AP05, p. 30] it is known that the space of infinitely differentiable functions with compact support is dense in $\mathcal{V}$. With these results it can be shown that $\mathcal{V}^0 \subseteq H \subseteq \mathcal{V}^*$ forms a Gelfand triple.

Corollary 4.2. *$\mathcal{V}^0 \subseteq H \subseteq \mathcal{V}^*$ forms a Gelfand triple.*

Proof. Since $\mathcal{D}(I)$ is dense in $\mathcal{V}^0$ and $L^2(I)$, the spaces $L^2(I)$ and $\mathcal{V}^0$ are separable and $\mathcal{V}^0$ is also dense in $L^2(I)$. It is clear that the embedding $\mathcal{V}^0 \subseteq L^2(I)$ is continuous because the $\mathcal{V}$-norm is by definition stronger than the L^2-norm. □

In the next lemma, we show that the bilinear form is bounded and fulfills a Gårding inequality under Assumption 4.1. A proof of Lemma 4.3 can also be found in [AP05, p. 32].

Lemma 4.3. *(Properties of $a^B(\cdot,\cdot)$ defined in (2.78))*
Under Assumption 4.1 the following properties hold:

i) The bilinear form $a^B(\cdot,\cdot)$ is continuous on $\mathcal{V}$, i.e. there exists a constant $C_1 > 0$ such that

$$|a^B(w,\varphi)| \leq C_1\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} \quad \text{for all } w,\varphi \in \mathcal{V}. \tag{4.3}$$

Moreover there exists a constant $C_2 > 0, C_3 > 0$ such that

$$|a_1^B(w,\varphi)| \leq C_2\|w\|_{\mathcal{V}}\|\varphi\|_H \quad \text{for all } w,\varphi \in \mathcal{V} \tag{4.4}$$

and

$$|a_0^B(w,\varphi)| \leq C_2\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} \quad \text{for all } w,\varphi \in \mathcal{V}. \tag{4.5}$$

ii) There exist positive α, β such that the Gårding inequality

$$a^B(\varphi,\varphi) \geq \alpha\|\varphi\|_{\mathcal{V}}^2 - \beta\|\varphi\|_{L^2(I)}^2 \quad \text{for all } \varphi \in \mathcal{V}^0 \tag{4.6}$$

is satisfied. Moreover, there exists a positive constant α_0 such that

$$a_0^B(\varphi,\varphi) \geq \alpha_0\|\varphi\|_{\mathcal{V}}^2 \quad \text{for all } \varphi \in \mathcal{V}^0. \tag{4.7}$$

Proof. To prove i) we have under Assumption 4.1 for all $w,\varphi \in \mathcal{V}$

$$\begin{aligned}
&\left|\int_I \left(\frac{\sigma^2}{2}S^2\frac{\partial w}{\partial S}\frac{\partial \varphi}{\partial S}\right)\mathrm{d}S\right| \leq \frac{\sigma^2}{2}\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}},\\
&\left|\int_I (rw\varphi)\mathrm{d}S\right| \leq r\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}},\\
&\left|\int_I \left((\sigma^2 + D_0 - r)S\frac{\partial w}{\partial S}\varphi\right)\mathrm{d}S\right| \leq |(\sigma^2 + D_0 - r)|\ \|w\|_{\mathcal{V}}\|\varphi\|_{L^2(I)}\\
&\qquad\qquad \leq (\sigma^2 + D_0 + r)\|w\|_{\mathcal{V}}\|\varphi\|_H,
\end{aligned}$$

which leads to

$$|a^B(w,\varphi)| \leq \left(\frac{3}{2}\sigma^2 + 2r + D_0\right)\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} =: C_1\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} \tag{4.8}$$

and

$$|a_1^B(w,\varphi)| \leq \left(\sigma^2 + 2r + D_0\right)\|w\|_{\mathcal{V}}\|\varphi\|_H =: C_2\|w\|_{\mathcal{V}}\|\varphi\|_H. \tag{4.9}$$

For the proof of ii) we have under Assumption 4.1 for all $\varphi \in \mathcal{V}^0$

$$\begin{aligned}
\int_I \frac{\sigma^2}{2}S^2\left(\frac{\partial\varphi}{\partial S}\right)^2 + r\varphi^2\mathrm{d}S &= \frac{\sigma^2}{2}\left\|S\left(\frac{\partial\varphi}{\partial S}\right)\right\|_{L^2(I)}^2 + r\|\varphi\|_{L_2(I)}^2\\
&\geq \min\left\{\frac{\sigma^2}{2}, r\right\}\|\varphi\|_{\mathcal{V}}^2.
\end{aligned} \tag{4.10}$$

Since $\varphi \in \mathcal{V}^0$ integration by parts leads to

$$\begin{aligned}\int_I S\frac{\partial\varphi}{\partial S}\varphi \mathrm{d}S &= \frac{1}{2}\int_I S\frac{\partial\varphi}{\partial S}\varphi \mathrm{d}S + \frac{1}{2}\int_I S\frac{\partial\varphi}{\partial S}\varphi \mathrm{d}S \\ &= -\frac{1}{2}\int_I S\frac{\partial\varphi}{\partial S}\varphi + \varphi^2 \mathrm{d}S + \frac{1}{2}\int_I S\frac{\partial\varphi}{\partial S}\varphi \mathrm{d}S \\ &= -\frac{1}{2}\int_I \varphi^2 \mathrm{d}S = -\frac{1}{2}\|\varphi\|^2_{L^2(I)}. \end{aligned} \tag{4.11}$$

With $c := |\sigma^2 + D_0 - r| > 0$ and (4.11) we have the following estimates for the convection term

$$\begin{aligned}\int_I (\sigma^2 + D_0 - r) S\frac{\partial\varphi}{\partial S}\varphi \mathrm{d}S &\geq -|(\sigma^2 + D_0 - r)| \left| \int_I S\frac{\partial\varphi}{\partial S}\varphi \mathrm{d}S \right| \\ &= -\frac{c}{2}\|\varphi\|^2_{L^2(I)}. \end{aligned} \tag{4.12}$$

Finally, together with (4.10) and (4.12), one estimates

$$a^B(\varphi,\varphi) \geq \min\left\{\frac{\sigma^2}{2}, r\right\}\|\varphi\|^2_{\mathcal{V}} - \frac{c}{2}\|\varphi\|^2_H.$$

Together with $\alpha := \min\left\{\frac{\sigma^2}{2}, r\right\} > 0$ and $\beta := \frac{c}{2} > 0$ this provides the proof of the Gårding inequality. Using the Cauchy-Schwarz inequality in (4.11) yields

$$2\left\|S\frac{\partial\varphi}{\partial S}\right\|_H \|\varphi\|_H \geq \|\varphi\|^2_H, \tag{4.13}$$

which leads to the following Poincare-type-inequality

$$2\left\|S\frac{\partial\varphi}{\partial S}\right\|_H \geq \|\varphi\|_H. \tag{4.14}$$

Finally,

$$\begin{aligned} a^0(\varphi,\varphi) := \int_I \frac{\sigma^2}{2}S^2\left(\frac{\partial\varphi}{\partial S}\right)^2 \mathrm{d}S &= \frac{\sigma^2}{4}\left\|S\frac{\partial\varphi}{\partial S}\right\|^2_H + \frac{\sigma^2}{4}\left\|S\frac{\partial\varphi}{\partial S}\right\|^2_H \\ &\geq \frac{\sigma^2}{4}\left\|S\frac{\partial\varphi}{\partial S}\right\|^2_H + \frac{\sigma^2}{16}\|\varphi\|^2_H \\ &\geq \frac{\sigma^2}{16}\|\varphi\|^2_{\mathcal{V}} =: \alpha_0\|\varphi\|^2_{\mathcal{V}} \end{aligned} \tag{4.15}$$

achieves the proof of (4.7). □

To prove some regularity results for the Black-Scholes variational inequality, we introduce the penalized problem corresponding to Problem 2.9 with inhomogeneous boundary condition as follows:
For $\varepsilon > 0$, find $y^\varepsilon(\tau) : (0,T] \to \mathcal{V}$ such that

$$\left\langle \frac{\partial y^\varepsilon(\tau)}{\partial \tau}, \varphi \right\rangle + a(y^\varepsilon(\tau), \varphi) + \varepsilon^{-1}(\mathfrak{Q} y^\varepsilon(\tau), \varphi) = 0 \ \text{ for all } \varphi \in \mathcal{V}^0, \text{ a.e. } \tau \in (0,T], \tag{4.16}$$

with $y^\varepsilon(\tau) = \mathcal{H}$ on ∂I and $y^\varepsilon(0) = \mathcal{H}$, where $\mathfrak{Q} y^\varepsilon(\tau) := \min(y^\varepsilon(\tau) - \mathcal{H}, 0)$ denotes the penalty function.
In the following lemma, we summarized some regularity results for the penalized problem. The authors in [Ach05] for the case of a put option and in [DL19b] for the case of a call option give a short proof for the regularity in weighted Sobolev spaces. We give a more detailed proof by applying Lemma 3.13 based on the ideas of [BL82, Mem12, KTK17].

Lemma 4.4. *Let Assumption 4.1 hold and let further $\mathcal{H}(S)$ be the payoff function for a put or call option, then the penalty problem in (4.16) has a unique solution $y^\varepsilon \in L^\infty(0,T;H) \cap L^2(0,T;\mathcal{V})$, $y^\varepsilon \in H^1(0,T;H) \cap L^\infty(0,T;\mathcal{V})$ and $y^\varepsilon \in L^2(0,T;H^2(I))$, and we have the following estimates*

(i) $\|y^\varepsilon\|_{L^\infty(0,T;H)} + \|y^\varepsilon\|_{L^2(0,T;\mathcal{V})} \lesssim \|\mathcal{H}\|_{\mathcal{V}} + \|\mathcal{H}\|_{H^{1/2}(\partial I)}$

ii) $\|\partial_\tau y^\varepsilon\|_{L^2(0,T;H)} + \|y^\varepsilon\|_{L^\infty(0,T;\mathcal{V})} \lesssim \|\mathcal{H}\|_{\mathcal{V}} + \|\mathcal{H}\|_{H^{1/2}(\partial I)}$

(iii) $\|y^\varepsilon\|_{L^2(0,T;H^2(I))} \lesssim \|\mathcal{H}\|_{\mathcal{V}} + \|\mathcal{H}\|_{H^{3/2}(\partial I)}$

with constants independent of the penalty parameter $\varepsilon > 0$.

Proof. For the proof, it is convenient to transform the penalty problem into a problem with homogeneous boundary condition. We set $\tilde{u}^\varepsilon(\tau) := y^\varepsilon(\tau) - \tilde{u}_{\mathcal{H}}$, where the function $\tilde{u}_{\mathcal{H}}$ on $I := (0, S_{\max})$ is defined as

$$\tilde{u}_{\mathcal{H}} := \frac{S_{\max} - S}{S_{\max}} \mathcal{H}(0) + \frac{S}{S_{\max}} \mathcal{H}(S_{\max}). \tag{4.17}$$

Then the penalty problem in (4.16) is equivalent to the following problem: For $\varepsilon > 0$, find $\tilde{u}^\varepsilon : (0,T] \to \mathcal{V}^0$ such that

$$\left\langle \frac{\partial \tilde{u}^\varepsilon}{\partial \tau}, \varphi \right\rangle + a^B(\tilde{u}^\varepsilon(\tau), \varphi) + \varepsilon^{-1}(\mathfrak{Q}\tilde{u}^\varepsilon, \varphi) = -a^B(\tilde{u}_{\mathcal{H}}, \varphi) \ \text{ for all } \varphi \in \mathcal{V}^0, \text{ a.e. } \tau \in (0,T], \tag{4.18}$$

and $\tilde{u}^\varepsilon(0) = \mathcal{H}(S) - \tilde{u}_{\mathcal{H}} =: \tilde{\mathcal{H}}(S)$, where $(\mathfrak{Q}\tilde{u}^\varepsilon, \varphi) := \min(\tilde{u}^\varepsilon - \tilde{\mathcal{H}}(S), 0)$.
Since $\tilde{\mathcal{H}} \in \mathcal{V}$, $\tilde{u}_{\mathcal{H}} \in \mathcal{V}$ are satisfied and due to Lemma 4.3, the use of Lemma 3.13 implies

$$\|\tilde{u}^\varepsilon\|_{L^\infty(0,T;H)} + \|\tilde{u}^\varepsilon\|_{L^2(0,T;\mathcal{V})} \lesssim \|\widetilde{\mathcal{H}}\|_{\mathcal{V}} + \|\tilde{u}_{\mathcal{H}}\|_{\mathcal{V}} \tag{4.19}$$

and

$$\|\partial_\tau \tilde{u}^\varepsilon\|_{L^2(0,T;H)} + \|\tilde{u}^\varepsilon\|_{L^\infty(0,T;\mathcal{V})} \lesssim \|\widetilde{\mathcal{H}}\|_{\mathcal{V}} + \|\tilde{u}_{\mathcal{H}}\|_{\mathcal{V}} \tag{4.20}$$

with constants independent of the penalty parameter ε. Substituting $\tilde{u}^\varepsilon := y^\varepsilon - \tilde{u}_\mathcal{H}$ and applying the triangle inequality in (4.19) leads to

$$\begin{aligned}\|y^\varepsilon\|_{L^\infty(0,T;H)} - \|\tilde{u}_\mathcal{H}\|_H + \|y^\varepsilon\|_{L^2(0,T;\mathcal{V})} - \|\tilde{u}_\mathcal{H}\|_\mathcal{V} &\lesssim \|\tilde{u}^\varepsilon\|_{L^\infty(0,T;H)} + \|\tilde{u}^\varepsilon\|_{L^2(0,T;\mathcal{V})} \\ &\lesssim \|\widetilde{\mathcal{H}}\|_\mathcal{V} + \|\tilde{u}_\mathcal{H}\|_\mathcal{V}. \end{aligned} \tag{4.21}$$

Since $\mathcal{H} \in H^{1/2}(\partial I)$ and $\tilde{u}_\mathcal{H} \in H^1(I)$ are satisfied it is estimated $\|\tilde{u}_\mathcal{H}\|_\mathcal{V} \lesssim \|\mathcal{H}\|_{H^{1/2}(\partial I)}$. This together with (4.19) and resubstituting $\tilde{\mathcal{H}} = \mathcal{H} - \tilde{u}_\mathcal{H}$ gives

$$\|y^\varepsilon\|_{L^\infty(0,T;H)} + \|y^\varepsilon\|_{L^2(0,T;\mathcal{V})} \lesssim \|\tilde{u}_\mathcal{H}\|_\mathcal{V} + \|\tilde{\mathcal{H}}\|_\mathcal{V} \lesssim \|\mathcal{H}\|_{H^{1/2}(\partial I)} + \|\mathcal{H}\|_\mathcal{V},$$

which completes the proof of (i). A similar approach gives the desired result in (ii).

For the proof of (iii), it is clear that $\tilde{\mathcal{H}} \in \mathcal{V}$, $\tilde{u}_\mathcal{H} \in H^2(\Omega)$ are satisfied. What remains to show is that (3.35) in Lemma 3.13 are satisfied. Therefore we have to consider the case of a put and call option separately. Due to the homogenization with the function $\tilde{u}_\mathcal{H}$ defined in (4.17) for a put option it follows

$$\begin{aligned}\widetilde{\mathcal{H}_\mathrm{P}}(S) := \mathcal{H}_\mathrm{P}(S) - u_{\mathcal{H}_\mathrm{P}}(S) &:= \mathcal{H}_\mathrm{P}(S) - \left(\frac{S_\mathrm{max} - S}{S_\mathrm{max}}\mathcal{H}_\mathrm{P}(0) + \frac{S}{S_\mathrm{max}}\mathcal{H}_\mathrm{P}(S_\mathrm{max})\right) \\ &= \max\{0, K - S\} - \frac{S_\mathrm{max} - S}{S_\mathrm{max}}K.\end{aligned}$$

This together with the definition of $a_0^B(\cdot,\cdot)$ in (2.79) leads to

$$\begin{aligned}-a_0^B(\widetilde{\mathcal{H}_\mathrm{P}}, \mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})) :&= -\int_I \frac{\sigma^2}{2}S^2\frac{\partial\widetilde{\mathcal{H}_\mathrm{P}}}{\partial S}\frac{\partial\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})}{\partial S}\mathrm{d}S \\ &= \int_0^K \frac{\sigma^2}{2}S^2\frac{\partial\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})}{\partial S}\mathrm{d}S + \int_I \frac{\sigma^2}{2}S^2\frac{\partial u_{\mathcal{H}_\mathrm{P}}}{\partial S}\frac{\partial\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})}{\partial S}\mathrm{d}S. \end{aligned} \tag{4.22}$$

Integration by parts and $\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r}) \le 0$ in I and $\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r}) = 0$ on ∂I for a.e. $\mathfrak{r} \in (0,T]$ and Cauchy-Schwarz implies

$$\begin{aligned}-a_0^B(\widetilde{\mathcal{H}_\mathrm{P}}, \mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})) =& \frac{\sigma^2}{2}S^2\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\Big|_0^K - \int_0^K \sigma^2 S\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\mathrm{d}S \\ &+ \frac{\sigma^2}{2}S^2\frac{\partial u_{\mathcal{H}_\mathrm{P}}}{\partial S}\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\Big|_{\partial I} - \int_I \sigma^2 S\frac{\partial u_{\mathcal{H}_\mathrm{P}}}{\partial S}\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\mathrm{d}S \\ \le& -\int_0^K \sigma^2 S\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\mathrm{d}S - \int_I \sigma^2 S\frac{\partial u_{\mathcal{H}_\mathrm{P}}}{\partial S}\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\mathrm{d}S \\ =& \sigma^2\int_I S\left(\frac{\partial\mathcal{H}_\mathrm{P}}{\partial S} - \frac{\partial u_{\mathcal{H}_\mathrm{P}}}{\partial S}\right)\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\mathrm{d}S \\ \le& \sigma^2\|\mathcal{H}_\mathrm{P} - u_{\mathcal{H}_\mathrm{P}}\|_\mathcal{V}\|\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\|_H \\ =:& \sigma^2\|\widetilde{\mathcal{H}_\mathrm{P}}\|_\mathcal{V}\|\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\|_H. \end{aligned} \tag{4.23}$$

For a call option the transformed payoff function is given as follows

$$\begin{aligned}\widetilde{\mathcal{H}_{\mathrm{C}}}(S) := \mathcal{H}_{\mathrm{C}}(S) - u_{\mathcal{H}_{\mathrm{C}}}(S) &:= \mathcal{H}_{\mathrm{C}}(S) - \left(\frac{S_{\max}-S}{S_{\max}}\mathcal{H}_{\mathrm{C}}(0) + \frac{S}{S_{\max}}\mathcal{H}_{\mathrm{C}}(S_{\max})\right)\\ &= \max\{0, S-K\} - \frac{S}{S_{\max}}(S_{\max}-K).\end{aligned}$$

This leads to

$$\begin{aligned}-a_0^B(\widetilde{\mathcal{H}_{\mathrm{C}}}, \mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})) :&= -\int_I \frac{\sigma^2}{2}S^2\frac{\partial\widetilde{\mathcal{H}_{\mathrm{C}}}}{\partial S}\frac{\partial\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})}{\partial S}\mathrm{d}S\\ &= -\int_K^{S_{\max}} \frac{\sigma^2}{2}S^2\frac{\partial\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})}{\partial S}\mathrm{d}S + \int_I \frac{\sigma^2}{2}S^2\frac{\partial u_{\mathcal{H}_{\mathrm{C}}}}{\partial S}\frac{\partial\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})}{\partial S}\mathrm{d}S.\end{aligned}$$

Further, similar considerations as for the put option, i.e. integration by parts and $\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r}) \le 0$ in I and $\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r}) = 0$ on ∂I for a.e. $\mathfrak{r} \in (0,T]$ and Cauchy-Schwarz inequality imply

$$\begin{aligned}-a_0^B(\widetilde{\mathcal{H}_{\mathrm{C}}}, \mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})) &= -\frac{\sigma^2}{2}S^2\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\Big|_K^{S_{\max}} + \int_K^{S_{\max}} \sigma^2 S\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\mathrm{d}S\\ &\quad + \frac{\sigma^2}{2}S^2\frac{\partial u_{\mathcal{H}_{\mathrm{C}}}}{\partial S}\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\Big|_{\partial I} - \int_I \sigma^2 S\frac{\partial u_{\mathcal{H}_{\mathrm{C}}}}{\partial S}\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\mathrm{d}S\\ &\le \int_K^{S_{\max}} \sigma^2 S\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\mathrm{d}S - \int_I \sigma^2 S\frac{\partial u_{\mathcal{H}_{\mathrm{C}}}}{\partial S}\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\mathrm{d}S\\ &= \sigma^2\int_I S\left(\frac{\partial\mathcal{H}_{\mathrm{C}}}{\partial S} - \frac{\partial u_{\mathcal{H}_{\mathrm{C}}}}{\partial S}\right)\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\mathrm{d}S\\ &\le \sigma^2\|\mathcal{H}_{\mathrm{C}} - u_{\mathcal{H}_{\mathrm{C}}}\|_{\mathcal{V}}\ \|\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\|_H =: \sigma^2\|\widetilde{\mathcal{H}_{\mathrm{C}}}\|_{\mathcal{V}}\ \|\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\|_H. \qquad (4.24)\end{aligned}$$

In summary, using (4.23) and (4.24) and applying Cauchy-Schwarz inequality we conclude

$$-\int_0^\tau a_0^B(\tilde{\mathcal{H}}, \mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r}))\mathrm{d}\mathfrak{r} \lesssim \int_0^\tau \|\tilde{\mathcal{H}}\|_{\mathcal{V}}\ \|\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\|_H\mathrm{d}\mathfrak{r} \lesssim \sqrt{T}\|\tilde{\mathcal{H}}\|_{\mathcal{V}}\left(\int_0^\tau \|\mathfrak{Q}\tilde{u}^\varepsilon(\mathfrak{r})\|_H\mathrm{d}\mathfrak{r}\right)^{\frac{1}{2}} \qquad (4.25)$$

for a put and call option. Then the use of Lemma 3.13 implies

$$\|\tilde{u}^\varepsilon\|_{L^2(0,T;H^2(I))} \lesssim \|\tilde{\mathcal{H}}\|_{\mathcal{V}} + \|\tilde{u}_{\mathcal{H}}\|_{H^2(I)}.$$

Finally, resubstituting $\tilde{u}^\varepsilon = y^\varepsilon - \tilde{u}_{\mathcal{H}}$, $\tilde{\mathcal{H}} = \mathcal{H} - \tilde{u}_{\mathcal{H}}$, applying the triangle inequality and using $\|\tilde{u}_{\mathcal{H}}\|_{H^2(I)} \lesssim \|\mathcal{H}\|_{H^{3/2}(\partial I)}$ leads to the desired result (iii). □

Finally, by passing to the limit in (4.16) as $\varepsilon \to 0$ it can be shown that the solution of the variational inequality satisfies the following regularity results.

Theorem 4.5. *Under Assumption 4.1 Problem 2.9 has a unique solution*

$$y \in L^\infty(0,T;H)\cap L^2(0,T;\mathcal{V}),\ y \in H^1(0,T;H)\cap L^\infty(0,T;\mathcal{V}) \text{ and } y \in L^2(0,T;H^2(\Omega)).$$

Proof. See Theorem 3.14 in Section 3.1.2. □

4.2. Discretization

4.2.1. Semi-discretization schemes in time

In this section, we introduce a semi-discrete scheme for Problem 2.10. Let $s > 0$ be the time step size and let $\tau_0 < \tau^{(1)} < \ldots < \tau^{(\#\mathbb{T})} := T$ be mesh points on $[0, T]$. We denote the semi-discrete solution in time step z by $u^{(z)} := u(\tau^{(z)}, S)$. We approximate the partial derivative in time step $z-1$ by the following forward difference quotient:

$$\frac{\partial u(\tau, S)}{\partial \tau}\Big|_{\tau=\tau^{(z-1)}} \approx \frac{u^{(z)} - u^{(z-1)}}{s} \tag{4.26}$$

In time step z the backward difference quotient is the same as above. Combining a weighted average $\varpi \in [0, 1]$ of the backward Euler scheme at $\tau^{(z)}$ and the forward Euler scheme at $\tau^{(z-1)}$ in Problem 2.10 yields

$$\begin{aligned}\left(\frac{u^{(z)} - u^{(z-1)}}{s}, \varphi - u^{(z)}\right) + \varpi a^B(u^{(z)}, \varphi - u^{(z)}) + (1-\varpi) a^B(u^{(z-1)}, \varphi - u^{(z)}) \\ \geq -a^B(\mathcal{H}, \varphi - u^{(z)}) \quad \text{for all } \varphi \in \mathcal{K}^0. \end{aligned} \tag{4.27}$$

Moving the known terms in time step $z-1$ to the right hand side leads to the following semi-discrete problem:

Problem 4.6. *(Semi-discrete problem)*
Find $u^{(z)} \in \mathcal{K}^0$ for $z = 1, \ldots, \mathbb{T}$ such that

$$\tilde{a}^B(u^{(z)}, \varphi - y^{(z)}) \geq (\tilde{f}^{(z-1)}, \varphi - u^{(z)}) \quad \textit{for all } \varphi \in \mathcal{K}, \tag{4.28}$$

$$u^{(z)} = u, \tag{4.29}$$

where the bilinear form $\tilde{a}^B(\cdot, \cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is defined as

$$\tilde{a}^B(u^{(z)}, \varphi - u^{(z)}) := s\varpi \; a^B(u^{(z)}, \varphi - u^{(z)}) + (u^{(z)}, \varphi - u^{(z)}) \tag{4.30}$$

and right hand side

$$(\tilde{f}^{(z-1)}, \varphi - y^{(z)}) := s(\varpi - 1) \; a^B(u^{(z-1)}, \varphi - u^{(z)}) + (u^{(z-1)}, \varphi - u^{(z)}) - s a^B(\mathcal{H}, \varphi - u^{(z)}). \tag{4.31}$$

For $\varpi = 0$ we have an *explicit Euler scheme*, for $\varpi = 0.5$ we have the *Crank-Nicolson scheme* and for $\varpi = 1$ an *implicit Euler scheme*.
In the literature [Glo84, chapter III, section 4], the use of an explicit Euler scheme for variational inequalities is not recommended, since it is conditionally stable and the numerical approximation of $u^{(z)} \in \mathcal{K}$ will generally require the use of an iterative method.

Remark 4.7. *For parabolic partial differential equations with non-smooth initial data, people often use the so called* Rannacher timestepping *instead of the well-known* Crank-Nicolson *scheme. This is due to the fact that applying a Crank-Nicolson method can result in oscillations in the numerical solution and its partial derivatives. That seems to be confusing since the Crank-Nicolson method is known to be unconditionally stable. But it is unconditionally stable in the L^2-norm for initial data which lies in L^2, see [GC06]. In particular, some initial errors are damped very slowly and can result in oscillations in the partial derivatives of the numerical solution, if the initial condition is only continuous. This is due to the fact that the Crank-Nicolson timestepping is only A-stable, not strongly A-stable. One often observes that the convergence in time is less than the second order in the L^2-norm as achieved for smooth initial data.*
To avoid such problems it has been proposed by Rannacher [Ran84], known as Rannacher timestepping, to replace the first two Crank-Nicolson time steps by four half-time steps of an implicit scheme. A priori estimates for parabolic partial differential equations with non-smooth initial data in [Ran84] show that a second order convergence for the Rannacher scheme in time can be achieved. In [GC06] the authors present some results for the Black-Scholes Problem in the case of European options. They show that the Rannacher timestepping together with a finite difference scheme leads to a second order convergence for the first and second derivative of the option price, while applying a Crank-Nicolson scheme can result in oscillations.
In the context of parabolic variational inequalities, higher order schemes like the Crank-Nicolson scheme or the Rannacher timestepping are not recommended, since the regularity of the solution over time is poor and such schemes do not improve the convergence in time ([Glo84], chapter III, section 4). In Section 8.2.3 we present some numerical results with a spatial cubic B-spline Galerkin-discretization, which show that applying the Rannacher timestepping scheme to the Black-Scholes variational inequality facilitates the numerical computation of the first and second derivative of the numerical solution ($\widehat{=}$ Greeks) without oscillations but does not lead to an optimal convergence rate in time. Similar observations are made in [FV02] for an American put option, where the authors present some numerical results for the partial derivatives computed with the Black-Scholes variational inequality together with a spatial finite volume discretization.

For the well-posedness of the semi-discrete schemes we make the following assumption on the time step size $0 < s < 1$.

Assumption 4.8. *We assume $s\varpi\beta \leq \Lambda$ for some $\Lambda \in (0,1)$, where $\beta := \frac{|\sigma^2 + D_0 - r|}{2}$ is the constant from Gårding's inequality and $\varpi \in \{0.5, 1\}$.*

Corollary 4.9 provides the well-posedness of the semi-discrete Problem 4.6.

Corollary 4.9. *Suppose that Assumptions 4.1 and 4.8 are satisfied, then for fixed time level $\tau^{(k)}$ there exists a unique solution to the semi-discrete problem 4.6.*

Proof. Due to the existence and uniqueness Theorem 3.3 for elliptic variational inequalities, it remains to show that the bilinear form $\tilde{a}^B(\cdot,\cdot)$ for $\varpi \in \{0.5, 1\}$ is bounded and coercive. Applying Lemma 4.3 i) to the bilinear form in (4.30) we have

$$\begin{aligned} |\tilde{a}^B(y,\varphi)| :&= |s\varpi\ a^B(y,\varphi) + (y,\varphi)| \\ &\leq (s\varpi C_1 + 1)\|y\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} \quad \text{for all } y,\varphi \in \mathcal{V} \end{aligned}$$

for $C_1 > 0$. Since $a^B(\cdot,\cdot)$ satisfies a Gårding inequality with $\alpha, \beta > 0$ by Lemma 4.3 ii) and under Assumption 4.8, $s\varpi\beta \leq \Lambda$ for $\Lambda \in (0,1)$, the estimation is

$$\tilde{a}^B(\varphi,\varphi) := s\varpi\ a^B(\varphi,\varphi) + (\varphi,\varphi) \geq s\varpi\alpha\ \|\varphi\|^2_{\mathcal{V}} + (1 - s\varpi\beta)\|\varphi\|^2_{L_2(I)} \geq s\varpi\alpha\ \|\varphi\|^2_{\mathcal{V}}$$

for all $\varphi \in \mathcal{V}^0$. Hence $\tilde{a}^B(\cdot,\cdot)$ is linear, bounded and coercive and there exists a unique solution to Problem 4.6 for fixed time level $\tau^{(k)}$. □

4.2.2. B-spline Galerkin discretization

In this section we derive a fully discrete B-spline Galerkin scheme for the parabolic variational inequality arising from the Black-Scholes model.
Recall that, in finance, to determine optimal risk strategies, the interest lies not only in the solution of the variational inequality, i.e. the option price, but also in its partial derivatives up to order two, the so-called Greeks. A cubic B-spline discretization facilitates a pointwise approximation of Gamma, the second derivative of the solution. One advantage of the B-spline discretization in the context of variational inequalities (for example in comparison to higher order finite elements) is that due to the non-negativity of B-splines there is a guarantee that the solution lies in the convex set by ensuring the non-negativity of the B-spline coefficients without evaluating the corresponding spline. This idea was first introduced by [Hol04, HK07].

Remark 4.10. *One particular difficulty for the discretization with higher order B-splines for the valuation of an option arises from the fact that the initial condition is typically not differentiable and an approximation with higher order B-splines results in oscillations. Consequently, we use a spatial discretization based on B-splines of order k with $k-1$ coinciding knots at the point where the initial condition is not differentiable, i.e. $S = K$ for an American put and call option. To the best of my knowledge, this application of B-splines with internal coinciding knots for the numerical solution of the Black-Scholes variational inequality appears for the first time in the literature.*

Let $\mathcal{V}_h^0 \subset \mathcal{V}$ be the finite dimensional spline space, a subspace of $\mathcal{V}$, with uniform grid size h. Let $\Theta := \{\theta_i\}_{i=1}^{n+k}$, $i \in \mathbb{N}$, be the extended sequence of knots with $\theta_1 := 0$, $\theta_\wp := K$ and $\theta_{n+1} := S_{\max}$, such that

$$\theta_1 = \ldots = \theta_k < \theta_{k+1} < \ldots < \theta_\wp = \ldots = \theta_{\wp+k-2} < \ldots < \ldots < \theta_{n+1} = \ldots = \theta_{n+k}. \tag{4.32}$$

The recursive definition of B-splines $N_{i,k}$ regarding the extended sequence of knots can be found in (3.62). Then the discrete solution space $\mathcal{V}_h^0$ is given by

$$\mathcal{V}_h^0 := \operatorname{span}\{N_{i,k}, i \in \mathcal{I} := \{i = 2, \ldots, n-1\}\}, \tag{4.33}$$

where $N_{1,k}$ and $N_{n,k}$ are omitted due to the zero boundary conditions on ∂I. The dimension of the discrete solution space is $\dim(\mathcal{V}_h^0) = \#\mathcal{I} = n - 2$. We define an approximation of the convex set $\mathcal{K}^0$ by

$$\mathcal{K}_h^0 := \{\varphi_h \in \mathcal{V}_h^0 : \varphi_h \geq 0\}. \tag{4.34}$$

Note that $\mathcal{K}_h^0 = \mathcal{K} \cap \mathcal{V}_h^0$ is satisfied because the obstacle is equal to zero and no approximation of the obstacle is needed. Then the discrete solution $u_h^{(z)} \in \mathcal{K}_h^0$ for the variational inequality is given by

$$u_h^{(z)} = \sum_{i \in \mathcal{I}} u_i^{(z)} N_{i,k}(S) \quad \text{for } z = 1, \ldots, \#\mathbb{T}. \tag{4.35}$$

This leads to the following discrete Galerkin scheme for problem 4.6:

Problem 4.11. *(Fully discrete variational inequality)*
Find $u_h^{(z)} \in \mathcal{K}_h^0$ for $z = 1, \ldots, \mathbb{T}$ such that $u_h^{(z)} = 0$

$$\tilde{a}^B(u_h^{(z)}, \varphi_h - u_h^{(z)}) \geq (\tilde{f}^{(z-1)}, \varphi_h - u_h^{(z)}) \quad \textit{for all } \varphi_h \in \mathcal{K}_h^0,$$

with bilinear form as defined in (4.30) and right hand side as defined in (4.31).

Inserting the elements $u_h^{(z)}, \varphi_h \in \mathcal{K}_h^0$ yields

$$\tilde{a}^B(u_h^{(z)}, \varphi_h - u_h^{(z)}) =: (\boldsymbol{\varphi} - \mathbf{u}^{(z)})^T C \mathbf{u}^{(z)} \tag{4.36}$$

where the discretization matrix $C \in \mathbb{R}^{\#\mathcal{I} \times \#\mathcal{I}}$ is given by

$$C := \varpi s(\mathfrak{A} + \mathfrak{B}) + (\varpi sr + 1)\mathfrak{G}. \tag{4.37}$$

Due to the support of B-splines $\mathcal{Q} := \operatorname{supp}(N_{i,k}) \cap \operatorname{supp}(N_{j,k})$ we have

$$\mathfrak{A}_{j,i} := \frac{\sigma^2}{2} \int_{\mathcal{Q}} \left(S^2 N'_{i,k}(S) N'_{j,k}(S) \right) \mathrm{d}S,$$

$$\mathfrak{B}_{j,i} := \left(\sigma^2 + D_0 - r \right) \int_{\mathcal{Q}} S N'_{i,k}(S) N_{j,k}(S) \mathrm{d}S,$$

$$\mathfrak{G}_{j,i} := \int_{\mathcal{Q}} N_{i,k}(S) N_{j,k}(S) \mathrm{d}S.$$

For the right hand side we obtain

$$(\tilde{f}^{(z-1)}, \varphi_h - u_h^{(z)}) - \tilde{a}^B(\mathcal{H}(S), \varphi_h - u_h^{(z)}) = (\boldsymbol{\varphi} - \mathbf{u}^{(z)})^T \mathbf{f}^{(z-1)},$$

where $\mathbf{f}^{(z-1)} \in \mathbb{R}^{\#\mathcal{I}}$ is defined as

$$\mathbf{f}^{(z-1)} := ((\varpi - 1) s (\mathfrak{A} + \mathfrak{B}) + ((\varpi - 1) sr + 1) \mathfrak{G}) \, \mathbf{u}^{(z-1)} - s \mathfrak{d} \tag{4.38}$$

with $\mathfrak{d} \in \mathbb{R}^{\#\mathcal{I}}$ defined as follows

$$\mathfrak{d}_i := \int \left(\frac{\sigma^2}{2} S^2 \mathcal{H}'(S) N'_{i,k}(S) \right) + r \mathcal{H}(S) N_{i,k}(S) \mathrm{d}S + \int \left(\sigma^2 + D_0 - r \right) S \mathcal{H}'(S) N_{i,k}(S) \mathrm{d}S.$$

Since B-splines are piecewise polynomials and $\mathcal{H}(S)$ is piecewise linear, employing an appropriate quadrature rule such as Gauss quadrature the matrices and the right hand side above can be computed exactly. Due to the non-negativity of B-splines the discrete convex set $\mathcal{K}_h^0$ can be determined by

$$\mathbf{K} := \{ \boldsymbol{\varphi} \in \mathbb{R}^{\#\mathcal{I}} : \boldsymbol{\varphi} \geq 0 \}. \tag{4.39}$$

This formulation has the advantage that it can be ensured more efficiently that the discrete solution $u_h^{(z)}$ lies in the convex set $\mathcal{K}_h^0$ as it is not necessary to evaluate the corresponding spline. In summary, we can reformulate Problem 4.11 to the following problem in matrix vector notation:

Problem 4.12. *(Discrete variational inequality in matrix vector notation)*
Find $\mathbf{u}^{(z)} \in \mathbf{K}$ *for all* $z = 1, \ldots, \mathbb{T}$ *such that* $\mathbf{u}^{(0)} = 0$ *and*

$$\left(\boldsymbol{\varphi} - \mathbf{u}^{(z)} \right)^T \left(C \mathbf{u}^{(z)} - \mathbf{f}^{(z-1)} \right) \geq 0 \quad \text{for all } \boldsymbol{\varphi} \in \mathbf{K}$$

is satisfied.

For further numerical computations, it is convenient to reformulate Problem 4.12 as a discrete complementarity problem because there exist a number of iterative schemes which are devoted to solve such problems.

Problem 4.13. *(Discrete linear complementarity problem)*
Find $\mathbf{u}^{(z)} \in \mathbb{R}^{\#\mathcal{I}}$ *for all* $z = 1, \ldots, \mathbb{T}$ *such that* $\mathbf{u}^{(0)} = 0$ *and*

$$\begin{aligned} (\mathbf{u}^{(z)})^T \left(C\mathbf{u}^{(z)} - \mathbf{f}^{(z-1)} \right) &= 0 \\ C\mathbf{u}^{(z)} - \mathbf{f}^{(z-1)} &\geq 0 \\ \mathbf{u}^{(z)} &\geq 0 \end{aligned}$$

are satisfied.

5. Pricing American Put option with Heston's model

5.1. Well-posedness and regularity

In this section, we prove the existence and uniqueness of a solution for the Heston variational inequality as stated in Problem 2.17. Therefore, we are going to apply some results from Section 3.1.2. Unless stated otherwise, the results for the Heston variational inequality in this section are of my own founding.

Assumption 5.1. *For the analysis we make the following assumption on the data* $v_{\min}, v_{\max}$, γ, κ, ρ *and* r. *We assume that the variance attains a minimum and is strictly positive and bounded i.e.*

$$0 < v_{\min} < v < v_{\max} < \infty. \tag{5.1}$$

It is also assumed that the following parameters are positive and bounded

$$0 < r < \infty,\ 0 < \xi < \infty\ ,\ 0 < \kappa < \infty, 0 \leq \lambda < \infty \text{ and } 0 \leq \gamma < \infty \tag{5.2}$$

and the correlation satisfies

$$|\rho| < 1. \tag{5.3}$$

For the set of boundary conditions in (2.110) it is assumed that the maximal volatility $v_{\max}$ *satisfies*

$$\infty > v_{\max} \geq \frac{\gamma\kappa - 0.5\xi^2}{\kappa + \lambda}. \tag{5.4}$$

For the set of boundary conditions in (2.111) we assume that

$$0 < v_{\min} \leq \frac{\gamma\kappa - 0.5\xi^2}{\kappa + \lambda} \quad \text{and} \quad \infty > v_{\max} \geq \frac{\gamma\kappa - 0.5\xi^2}{\kappa + \lambda}. \tag{5.5}$$

Remark 5.2. *Note that Assumptions (5.2) and (5.3) are also assumed in the derivation of the Heston equation due to their financial meaning and that there is no problem to choose* $v_{\max}$ *large enough such that (5.4) is satisfied. The only restrictive assumptions are* $v_{\min} > 0$ *in (5.1) and* $v_{\min} \leq \frac{\gamma\kappa - 0.5\xi^2}{\kappa+\lambda}$ *in (5.5), which are only satisfied if the Feller condition discussed in Section 2.1.3 is imposed.*

Lemma 5.3. *(Properties of $a^H(\cdot,\cdot)$ in (2.128))*
Under Assumption 5.1 the following properties are satisfied:

i) $a^H(\cdot,\cdot)$ is a bounded bilinear form on $\mathcal{V}\times\mathcal{V}$, i.e. there exists $C>0$ such that

$$|a^H(w,\varphi)| \le C\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} \quad \text{for all } w,\varphi\in\mathcal{V}. \tag{5.6}$$

ii) There exists $\alpha>0$ and β such that

$$a^H(\varphi,\varphi) \ge \alpha\|\varphi\|^2_{\mathcal{V}} - \beta\|\varphi\|^2_H \quad \text{for all } \varphi\in\mathcal{V}^0. \tag{5.7}$$

Proof. For the proof of i) we have

$$\begin{aligned}
|a^H(w,\varphi)| &= \left|\int_\Omega \left(\tilde{A}\nabla w\cdot\nabla\varphi\right) + \left((\tilde{\mathbf{b}}\cdot\nabla w + rw)\varphi\right)\mathrm{d}\Omega\right| \\
&= \left|\int_\Omega \frac{1}{2}v\left(w_x\varphi_x + 2\rho\xi w_v\varphi_x + \xi^2 w_v\varphi_v\right)\right. \\
&\quad \left. + \left(\left(\frac{1}{2}v - r\right)w_x + \left(\kappa(v-\gamma)+\lambda v + \frac{1}{2}\xi^2\right)w_v\right)\varphi + rw\varphi\mathrm{d}\Omega\right|
\end{aligned} \tag{5.8}$$

Now we consider each term separately. By taking the maximum, applying the Cauchy-Schwarz inequality and the definition of the $\mathcal{V}$-norm, we have for all $w,\varphi\in\mathcal{V}$

$$\begin{aligned}
&\left|\int_\Omega \left(\frac{1}{2}vw_x\varphi_x \;+ \frac{1}{2}v\xi^2 w_v\varphi_v + rw\varphi\right)\mathrm{d}\Omega\right| \\
&\qquad \le \max\left\{\frac{1}{2}v_{\max}, \frac{1}{2}v_{\max}\xi^2, r\right\}\left|\int_\Omega \nabla w\cdot\nabla\varphi + w\varphi\mathrm{d}\Omega\right| \\
&\qquad \le \max\left\{\frac{1}{2}v_{\max}, \frac{1}{2}v_{\max}\xi^2, r\right\}\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} =: c_1\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}}.
\end{aligned} \tag{5.9}$$

Applying the Cauchy-Schwarz inequality for the mixed term, we estimate

$$\begin{aligned}
\left|\int_\Omega v\rho\xi w_v\varphi_x\mathrm{d}\Omega\right| &\le v_{\max}\rho\xi\|w_v\|_{L^2(\Omega)}\|\varphi_x\|_{L^2(\Omega)} \\
&\le v_{\max}\rho\xi\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} =: c_2\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}}.
\end{aligned} \tag{5.10}$$

For the convection term we have

$$\begin{aligned}
\left|\int_\Omega \left(\left(\frac{1}{2}v - r\right)w_x\varphi\right)\mathrm{d}\Omega\right| &\le \left|\left(\frac{1}{2}v_{\max} - r\right)\right|\|w_x\|_{L^2(\Omega)}\|\varphi\|_{L^2(\Omega)} \\
&\le \left|\left(\frac{1}{2}v_{\max} - r\right)\right|\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} =: c_3\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}}
\end{aligned} \tag{5.11}$$

and, finally,

$$\begin{aligned}
\left|\int_\Omega \left(\left(\kappa(v-\gamma)+\lambda v + \frac{1}{2}\xi^2\right)w_v\varphi\right)\mathrm{d}\Omega\right| &\le \left|\left(\kappa(v_{\max}-\gamma)+\lambda v_{\max} + \frac{1}{2}\xi^2\right)\right|\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} \\
&=: c_4\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}}.
\end{aligned} \tag{5.12}$$

By applying the triangle inequality to (5.8) together with (5.9),(5.10),(5.11) and (5.12) one has

$$\left|a^H(w,\varphi)\right| \leq \sum_{i=1}^{4} c_i \|w\|_{\mathcal{V}} \|\varphi\|_{\mathcal{V}} =: C\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}}. \tag{5.13}$$

To prove that the bilinear form satisfies the Gårding inequality in (ii), we need the following formulation of the convection term. We have to distinguish between the two sets of boundary conditions. We start by applying integration by parts to the convection term imposing the boundary conditions in (2.110), which leads to

$$\frac{1}{2}\int_\Omega (\tilde{\mathbf{b}}\cdot\nabla\varphi)\varphi \; d\Omega = -\frac{1}{2}\int_\Omega \varphi(\tilde{\mathbf{b}}\cdot\nabla\varphi) \; + (\nabla\cdot\tilde{\mathbf{b}})\varphi^2 \; d\Omega + \frac{1}{2}\tilde{b}_2(v_{\max})\int_{\Upsilon_4}\varphi^2 d\Upsilon_4.$$

The boundary integrals on Υ vanishes because of the choice of $\varphi \in \mathcal{V}^0$. Then the convection term becomes

$$\int_\Omega (\tilde{\mathbf{b}}\cdot\nabla\varphi)\varphi \; \mathrm{d}\Omega = -\frac{1}{2}\int_\Omega (\nabla\cdot\tilde{\mathbf{b}})\varphi^2 \; \mathrm{d}\Omega + \frac{1}{2}\tilde{b}_2(v_{\max})\int_{\Upsilon_4}\varphi^2 \mathrm{d}\Upsilon_4. \tag{5.14}$$

Since the term $\nabla\cdot\tilde{\mathbf{b}} = (\kappa+\lambda)$ is independent of the variables $(x,v)\in\Omega$ and in (5.4) it is assumed that $v_{\max} \geq \frac{\gamma\kappa - 0.5\xi^2}{\kappa+\lambda}$, i.e. $\tilde{b}_2(v_{\max}) \geq 0$ is satisfied, one has

$$\begin{aligned}\int_\Omega (\tilde{\mathbf{b}}\cdot\nabla\varphi)\varphi \; \mathrm{d}\Omega &= -\frac{1}{2}(\kappa+\lambda)\int_\Omega \varphi^2 \; d\Omega + \frac{1}{2}\tilde{b}_2(v_{\max})\int_{\Upsilon_4}\varphi^2 \mathrm{d}\Upsilon_4 \\ &\geq -\frac{1}{2}(\kappa+\lambda)\int_\Omega \varphi^2 \; \mathrm{d}\Omega. \end{aligned} \tag{5.15}$$

Imposing the boundary condition in (2.111), an analogous approach yields

$$\begin{aligned}\int_\Omega (\tilde{\mathbf{b}}\cdot\nabla\varphi)\varphi \; \mathrm{d}\Omega &= -\frac{1}{2}(\kappa+\lambda)\int_\Omega \varphi^2 \; d\Omega + \frac{1}{2}\tilde{b}_2(v_{\max})\int_{\Upsilon_4}\varphi^2 \mathrm{d}\Upsilon_4 - \frac{1}{2}\tilde{b}_2(v_{\min})\int_{\Upsilon_3}\varphi^2 \mathrm{d}\Upsilon_3 \\ &\geq -\frac{1}{2}(\kappa+\lambda)\int_\Omega \varphi^2 \; \mathrm{d}\Omega, \end{aligned} \tag{5.16}$$

where $\tilde{b}_2(v_{\max}) \geq 0$ and $\tilde{b}_2(v_{\min}) \leq 0$ due to the assumptions in (5.5).
Using the definition of the bilinear form in (2.128) and the formulation (5.15) or (5.16) above results in

$$a^H(\varphi,\varphi) \geq \int_\Omega \left(\tilde{A}\nabla\varphi\right)\cdot\nabla\varphi - \frac{1}{2}(\kappa+\lambda)\varphi^2 + r\varphi^2\varphi \mathrm{d}\Omega$$

for $\varphi \in \mathcal{V}^0$. Rearranging leads to

$$\begin{aligned} a^H(\varphi,\varphi) + \frac{1}{2}(\kappa+\lambda)\int_\Omega \varphi^2 \mathrm{d}\Omega &\geq \int_\Omega \left(\tilde{A}\nabla\varphi\right)\cdot\nabla\varphi + r\varphi^2 \mathrm{d}\Omega \\ &= \int_\Omega \frac{1}{2}v\varphi_x^2 + \varrho\xi v\varphi_x\varphi_v + \frac{1}{2}\xi^2 v\varphi_v^2 + r\varphi^2 \mathrm{d}\Omega. \end{aligned}$$

Applying Young's inequality for $v \geq v_{\min} > 0$ yields

$$\int_\Omega \xi v \varphi_x \varphi_v \mathrm{d}\Omega \geq \int_\Omega -\left|(\sqrt{v}\varphi_x)\,(\xi\sqrt{v}\varphi_v)\right| \mathrm{d}\Omega \geq \int_\Omega -\frac{1}{2} v\varphi_x^2 - \frac{1}{2} v\xi^2 \varphi_v^2 \mathrm{d}\Omega.$$

Assuming $|\rho| < 1$, $v_{\min} > 0$ and defining $\tilde{C} := \min_{v\in\Omega}\left\{\frac{1}{2}v(1-\varrho), \frac{1}{2}v\xi^2(1-\varrho)\right\}$ we obtain

$$\begin{aligned} a^H(\varphi,\varphi) + \frac{1}{2}(\kappa+\lambda)\int_\Omega \varphi^2 \mathrm{d}\Omega &\geq \int_\Omega \frac{1}{2} v(1-\varrho)\varphi_x^2 + \frac{1}{2} v\xi^2(1-\varrho)\varphi_v^2 + r\varphi^2 \mathrm{d}\Omega \\ &\geq \min\{\tilde{C}, r\} \int_\Omega \varphi_x^2 + \varphi_v^2 + \varphi^2 \mathrm{d}\Omega. \end{aligned}$$

Since we assume $|\rho| < 1$ and $v_{\min} > 0$ the constant $\tilde{C}$ is strict positive. By choosing $\beta := \frac{1}{2}(\kappa+\lambda)$ and $\alpha := \min\{\tilde{C}, r\}$ this leads to the following Gårding inequality

$$a^H(\varphi,\varphi) + \beta\|\varphi\|_H^2 \geq \alpha\|\varphi\|_{\mathcal{V}}^2 \qquad \text{for all } \varphi \in \mathcal{V}^0.$$

□

Remark 5.4. *It is also possible to show that $a^H(\cdot,\cdot)$ is continuous and satisfies a Gårding inequality in the weighted Sobolev space $\mathcal{V}_v^0 := \left\{\varphi \in H : \sqrt{v}\frac{\partial\varphi}{\partial x}, \sqrt{v}\frac{\partial\varphi}{\partial v} \in H\right\}$. Then the boundedness constant also depends on $\frac{1}{\sqrt{v}}$ and it is assumed $v_{\min} > 0$.*
In [AWW01] the authors show that the different bilinear form, where the boundary integral in (2.120) is equal to zero due to a Robin type boundary condition for $x = x_{\max}$ and Dirichlet boundary conditions for $x = x_{\min}, v = v_{\max}$ and $v = v_{\min}$, is bounded and coercive in the space $\mathcal{V}_A^0 := \left\{(\int_\Omega A\nabla\varphi\cdot\nabla\varphi + \tilde{c}\varphi^2 \mathrm{d}\Omega)^{1/2} < \infty, \varphi = 0 \text{ on } \partial\Omega \setminus (x_{\max}, v)\right\}$. There, the authors assume $v_{\min} > 0$ since the boundedness constant also depends on $\frac{1}{\sqrt{v}}$. In [HRSW13, pp. 108-113] the authors use a different transformation and assume a Dirichlet boundary condition on $\partial\Omega$, which leads to a different problem formulation as in this thesis. Using this formulation, the authors can show that the derived bilinear form is bounded and satisfies a Gårding-inequality, where the constant does not depend on v.

In order to prove some regularity results for the solution of Problem 2.17, we apply the results from Section 3.1.2. Therefore, we introduce the penalized problem corresponding to Problem 2.17 as follows: For $\varepsilon > 0$ find y^ε such that

$$\left\langle \frac{\partial y^\varepsilon(\tau)}{\partial\tau}, \varphi \right\rangle + a^H(y^\varepsilon(\tau),\varphi) + \varepsilon^{-1}(\mathfrak{Q}y^\varepsilon(\tau),\varphi) = \langle f(\tau),\varphi\rangle \quad \text{for all } \varphi \in \mathcal{V}^0, \tag{5.17}$$

with penalty operator $\mathfrak{Q}y^\varepsilon(\tau) := \min(y^\varepsilon(\tau) - y_0^\xi, 0)$ and $y^\varepsilon(\tau) = g$ on Υ, $y^\varepsilon(0) = g(x)$. It is known that the penalized problem, where the bilinear form is bounded and satisfies a Gårding inequality and $g \in \mathcal{V}$, has a unique solution $y^\varepsilon \in L^2(0,T;\mathcal{V}) \cap H^1(0,T;\mathcal{V}^*)$ (see Chapter 3.2 in [BL82]). To prove higher regularity for the penalized problem, Lemma 3.13 is applied, which is strongly based on the ideas of [BL82, Mem12, KTK17].

Lemma 5.5. *Let Assumption 5.1 hold and let further g be the transformed payoff function for a put option. Then the penalty problem in (5.17) has a unique solution $y^\varepsilon \in L^\infty(0,T;H) \cap L^2(0,T;\mathcal{V})$, and we have the following estimate*

$$\|y^\varepsilon\|_{L^\infty(0,T;H)} + \|y^\varepsilon\|_{L^2(0,T;\mathcal{V})} \lesssim \|g\|_{\mathcal{V}} \tag{5.18}$$

with constant independent of the penalty parameter ε.

Proof. First, we introduce a homogenized problem for the penalty problem in (5.17) and we set $u^\varepsilon := y^\varepsilon - g$: For $\varepsilon > 0$, find $u^\varepsilon : (0,T] \to \mathcal{V}^0$ such that

$$\left\langle \frac{\partial u^\varepsilon}{\partial \tau}, \varphi \right\rangle + a^H(u^\varepsilon, \varphi) + \varepsilon^{-1}(\mathfrak{Q}u^\varepsilon, \varphi) = -a^H(g, \varphi) \quad \text{for all } \varphi \in \mathcal{V}^0 \tag{5.19}$$

and $u^\varepsilon(0) = 0$, where $\mathfrak{Q}u^\varepsilon := \min(u^\varepsilon, 0)$ is a nondecreasing function. Since $g \in \mathcal{V}$ and due to Lemma the bilinear form $a^H(\cdot,\cdot)$ is bounded and satisfies a Gårding inequality, the use of Lemma 3.13 (i) implies

$$\|u^\varepsilon\|_{L^\infty(0,T;H)} + \|u^\varepsilon\|_{L^2(0,T;\mathcal{V})} \lesssim \|g\|_{\mathcal{V}} \tag{5.20}$$

with a constant independent of the penalty parameter ε. Resubstituting $u^\varepsilon := y^\varepsilon - g$ and applying the triangle inequality in (5.20) leads to

$$\|y^\varepsilon\|_{L^\infty(0,T;H)} - \|g\|_H + \|y^\varepsilon\|_{L^2(0,T;\mathcal{V})} - \|g\|_{\mathcal{V}} \lesssim \|u^\varepsilon\|_{L^\infty(0,T;H)} + \|u^\varepsilon\|_{L^2(0,T;\mathcal{V})} \lesssim \|g\|_{\mathcal{V}},$$

which achieves the proof. □

Finally, by passing to the limit $\varepsilon \to 0$ in (5.17) under Assumption 5.1 there exists a unique solution $y \in L^2(0,T;\mathcal{V}) \cap H^1(0,T;\mathcal{V}^*)$ to Problem 2.17, which satisfies the following regularity result.

Theorem 5.6. *Let the Assumption 5.1 hold. Then the solution of Problem 2.17 satisfies $y \in L^\infty(0,T;H) \cap L^2(0,T;\mathcal{V})$ and we have the estimate*

$$\|y\|_{L^\infty(0,T;H)} + \|y\|_{L^2(0,T;\mathcal{V})} \lesssim \|g\|_{\mathcal{V}}. \tag{5.21}$$

Proof. See Theorem 3.14 in Section 3.1.2. □

Remark 5.7. *To prove the existence of a strong solution, i.e. $y \in H^1(0,T;H) \cap L^\infty(0,T,\mathcal{V})$, one important step is exploiting the symmetry of the diffusion term. This result is also needed to show that the solution satisfies $y \in L^2(0,T,H^2(\Omega))$ and to conclude $y \in C(0,T;\mathcal{V})$. Due to the asymmetric diffusion operator in the Heston inequality in (2.128), $a_0^H(y,\varphi) := \int \tilde{A}\nabla y \cdot \nabla\varphi \mathrm{d}\Omega$, it is not possible to proceed as in Lemma 3.13 (ii)-(iii). However, imposing the set of boundary conditions in (2.110) with the non-smooth Dirichlet boundary condition $g(x)$ for $v = v_{\min}$ and $x \in (x_{\min}, x_{\max})$ it is clear that the solution cannot satisfy $y(\tau) \in H^2(\Omega)$ for some $\tau \in (0,T]$ on the whole spatial domain Ω. Imposing the boundary conditions in (2.111) we can expect that the solution satisfies the above mentioned regularity.*

5.2. Discretization

5.2.1. Semi-discretization schemes in time

In this section we establish a finite-difference semi-discretization of Problem 2.17 and discuss its well-posedness. Therefore, we decompose the time interval into equidistant grid points $0 = \tau^{(0)} < \tau^{(1)} < \cdots < \tau^{(\#\mathbb{T})} = T$ with time step size $s := \tau^{(z+1)} - \tau^{(z)}$. We define the semi-discrete solution in time step z by $y(\tau^{(z)}, x, v) =: y^{(z)}$. By an analogous approach as in Section 4.2.1 we derive the semi-discrete formulation of Problem 2.17:

$$\left(\frac{y^{(z)} - y^{(z-1)}}{s}, \varphi - y^{(z)}\right) + (1-\varpi)\ a^H(y^{(z-1)}, \varphi - y^{(z)}) + \varpi\ a^H(y^{(z)}, \varphi - y^{(z)}) \geq 0,$$

where $\varpi \in [0,1]$ and $(\cdot,\cdot)$ denotes the L_2-inner product. Sorting the known terms in time step $z-1$ to the right hand side, leads to the following semi-discrete problem:

Problem 5.8. *(Semi-discrete variational inequality with inhomogeneous Dirichlet boundary condition)*
Find $y^{(z)} \in \mathcal{K}$ for all $z \in \{1, \cdots, \#\mathbb{T}\}$, such that

$$\tilde{a}^H(y^{(z)}, \varphi - y^{(z)}) \geq f_{inh}^{(z-1)}(\varphi - y^{(z)}),$$

for all $\varphi \in \mathcal{K}^0$ with start condition $y^{(0)} = g(x)$, where the bilinear form is defined as

$$\begin{aligned}\tilde{a}^H(y^{(z)}, \varphi - y^{(z)}) :=& \varpi s\ a^H(y^{(z)}, \varphi - y^{(z)}) + (y^{(z)}, \varphi - y^{(z)}) \\ =& \int_\Omega \varpi s \tilde{A} \nabla y^{(z)} \cdot \nabla(\varphi - y^{(z)}) \mathrm{d}\Omega \\ &+ \int_\Omega (\varpi s(\tilde{\mathbf{b}} \cdot \nabla y^{(z)} + r y^{(z)}) + y^{(z)})(\varphi - y^{(z)}) \mathrm{d}\Omega \end{aligned} \tag{5.22}$$

and the right hand side as

$$\begin{aligned} f_{inh}^{(z-1)}(\varphi - y^{(z)}) :=& -(1-\varpi)s\ a^H(y^{(z-1)}, \varphi - y^{(z)}) + (y^{(z-1)}, \varphi - y^{(z)}) \\ =& -\int_\Omega (1-\varpi) s \tilde{A} \nabla y^{(z-1)} \cdot \nabla(\varphi - y^{(z)}) \mathrm{d}\Omega \\ &- \int_\Omega ((1-\varpi)s(\tilde{\mathbf{b}} \cdot \nabla y^{(z-1)} + r y^{(z-1)}) - y^{(z-1)})(\varphi - y^{(z)}) \mathrm{d}\Omega. \end{aligned} \tag{5.23}$$

For $\varpi = 0$ we obtain an explicit Euler scheme, for $\varpi = 0.5$ we have a Crank-Nicolson scheme and for $\varpi = 1$ an implicit Euler scheme. For further discussion of the schemes and recommendations for numerical computation of the Greeks in the context of variational inequalities the reader is referred to Section 4.2.1.

To study the well-posedness we make the following assumption on the time step size $0 < s < 1$.

Assumption 5.9. *We assume $s\varpi\beta \leq \Lambda$ for some $\Lambda \in (0,1)$, where $\beta := \frac{\kappa+\lambda}{2}$ is the constant from Gårding's inequality and $\varpi = \{0.5, 1\}$.*

For a fixed time level $\tau^{(z)}$ the bilinear form in (5.22) under Assumption 5.9 and Assumption 5.1 is bounded and coercive.

Lemma 5.10. *Suppose that Assumption 5.9 is fulfilled. Then the following properties are satisfied for fixed time level $\tau^{(z)}$:*

i) $\tilde{a}^H(\cdot,\cdot)$ is a bounded bilinear form on $\mathcal{V} \times \mathcal{V}$, i.e. there exists $C_0 > 0$ such that

$$|\tilde{a}^H(w,\varphi)| \leq C_0 \|w\|_{\mathcal{V}} \|\varphi\|_{\mathcal{V}} \quad \text{for all } w, \varphi \in \mathcal{V}. \tag{5.24}$$

ii) $\tilde{a}^H(\cdot,\cdot)$ is coercive under Assumption 5.1, i.e. there exists $\alpha_0 > 0$ such that

$$\tilde{a}^H(\varphi,\varphi) \geq \alpha_0 \|\varphi\|_{\mathcal{V}}^2 \quad \text{for all } \varphi \in \mathcal{V}^0. \tag{5.25}$$

Proof. We start with the proof of i). By the definition of the bilinear form $\tilde{a}^H(\cdot,\cdot)$ in (5.22), exploiting the boundedness of $a^H(\cdot,\cdot)$ from Lemma 5.3 and applying Cauchy-Schwarz inequality it is estimated

$$|\tilde{a}^H(w,\varphi)| := |\varpi s\ a^H(w,\varphi) + (w,\varphi)| \leq (s\varpi C + 1)\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} =: C_0\|w\|_{\mathcal{V}}\|\varphi\|_{\mathcal{V}} \tag{5.26}$$

for $w, \varphi \in \mathcal{V}$. For the proof of ii) we know from Lemma 5.3 that $a^H(\cdot,\cdot)$ satisfies a Garding inequality with constants $\alpha, \beta > 0$. Under assumption 5.9 this yields

$$\begin{aligned} \tilde{a}^H(\varphi,\varphi) := \varpi s\ a^H(w,\varphi) + (w,\varphi) &\geq \varpi s\alpha\|\varphi\|_{\mathcal{V}}^2 + (1 - \varpi s\beta)\|\varphi\|_H^2 \\ &\geq \varpi s\alpha\|\varphi\|_{\mathcal{V}}^2 \ \text{ for all } \varphi \in \mathcal{V}^0, \end{aligned} \tag{5.27}$$

whence $\tilde{a}^H(\cdot,\cdot)$ is coercive with constant $\alpha_0 := \varpi s\alpha > 0$. □

Finally, one can show that the semi-discrete solution $y(\tau^{(z)})$ for $\tau^{(z)} > 0$ lies in the space $H^2(\Omega)$ for the set of boundary conditions in (2.111).

Theorem 5.11. *Suppose that Assumption 5.3 and Assumption 5.9 are fulfilled. Then the solution of the semi-discrete Problem 5.8 for an implicit Euler or Rannacher time stepping method with the set of boundary conditions in (2.111) satisfies $y(\tau^{(z)}, x, v) \in H^2(\Omega)$ for fixed $\tau^{(z)} > 0$.*

Proof. We start by formulating the semi-discrete Heston variational inequality in Problem 5.8 with the set of boundary condition in (2.111) as a penalty problem: For $\varepsilon > 0$ find $y_\varepsilon^{(z)} \in \mathcal{V}$ such that

$$\tilde{a}^H(y_\varepsilon^{(z)}, \varphi) + \varepsilon^{-1}(\mathfrak{Q}y^\varepsilon, \varphi) = f_{\text{inh}}^{(z-1)}(\varphi), \tag{5.28}$$

with penalty function $\mathfrak{Q}y_\varepsilon^{(z)} := \min(y_\varepsilon^{(z)} - g, 0)$ for all $\varphi \in \mathcal{V}^0$. In order to find an equivalent formulation with homogeneous Dirichlet boundary condition on Υ we set $\tilde{u}_\varepsilon^{(z)} = y_\varepsilon^{(z)} - \tilde{u}_g$ and $\tilde{g} := g - \tilde{u}_g$ with

$$\tilde{u}_g = \frac{x_{\max} - x}{x_{\max} - x_{\min}} g(x_{\min}) + \frac{x - x_{\min}}{x_{\max} - x_{\min}} g(x_{\max}). \tag{5.29}$$

Then the penalty problem is equivalent to the following problem with homogeneous boundary condition on Υ: For $\varepsilon > 0$ find $u_\varepsilon^{(z)} \in \mathcal{V}^0$ such that

$$\tilde{a}^H(\tilde{u}_\varepsilon^{(z)}, \varphi) + \varepsilon^{-1}(\mathfrak{Q}u^\varepsilon, \varphi) = f_{\text{inh}}^{(z-1)}(\varphi) - \tilde{a}^H(\tilde{u}_g, \varphi - \tilde{u}_\varepsilon^{(z)}), \tag{5.30}$$

with penalty function $\mathfrak{Q}\tilde{u}_\varepsilon^{(z)} := \min(\tilde{u}_\varepsilon^{(z)} - \tilde{g}(x), 0)$ for all $\varphi \in \mathcal{V}^0$.
Now we have to show that (3.15) in Lemma 3.5 is satisfied: Recall the definition of the bilinear form from (5.22)

$$-\tilde{a}^H(\tilde{g}, \mathfrak{Q}\tilde{u}_\varepsilon^{(z)}) := -\varpi s a^H(\tilde{g}, \mathfrak{Q}\tilde{u}_\varepsilon^{(z)}) - (\tilde{g}, \mathfrak{Q}\tilde{u}_\varepsilon^{(z)}) \tag{5.31}$$

with

$$-a^H(\tilde{g}, \mathfrak{Q}\tilde{u}_\varepsilon^{(z)}) := -\int_\Omega \tilde{A}\nabla\tilde{g} \cdot \nabla\mathfrak{Q}\tilde{u}_\varepsilon^{(z)} + (\tilde{b} \cdot \nabla g + rg)\mathfrak{Q}\tilde{u}_\varepsilon^{(z)} \mathrm{d}\Omega.$$

Now we consider the diffusion term separately. Since $\tilde{g}(x)$ is independent of v one has

$$\begin{aligned}
-\int_\Omega \tilde{A}\nabla\tilde{g} \cdot \nabla\mathfrak{Q}\tilde{u}_\varepsilon^{(z)} \mathrm{d}\Omega &= -\int_\Omega \frac{1}{2} v \frac{\partial\tilde{g}}{\partial x} \frac{\partial\mathfrak{Q}\tilde{u}_\varepsilon^{(z)}}{\partial x} \mathrm{d}\Omega \\
&= -\int_\Omega \frac{1}{2} v \frac{\partial g}{\partial x} \frac{\partial\mathfrak{Q}\tilde{u}_\varepsilon^{(z)}}{\partial x} \mathrm{d}\Omega + \int_\Omega \frac{1}{2} v \frac{\partial\tilde{u}_g}{\partial x} \frac{\partial\mathfrak{Q}\tilde{u}_\varepsilon^{(z)}}{\partial x} \mathrm{d}\Omega \\
&= \int \int_{x_{\min}}^{0} \frac{1}{2} v K \exp(x) \frac{\partial\mathfrak{Q}\tilde{u}_\varepsilon^{(z)}}{\partial x} \mathrm{d}x\mathrm{d}v + \int_\Omega \frac{1}{2} v \frac{\partial\tilde{u}_g}{\partial x} \frac{\partial\mathfrak{Q}\tilde{u}_\varepsilon^{(z)}}{\partial x} \mathrm{d}\Omega.
\end{aligned}$$

Integration by parts with respect to x and $\mathfrak{Q}\tilde{u}_\varepsilon^{(z)} \leq 0$ in Ω, $\mathfrak{Q}\tilde{u}_\varepsilon^{(z)} = 0$ on Υ, $v > 0$, $\frac{\partial^2\tilde{u}_g}{\partial x^2} = 0$ and $\frac{\partial\tilde{u}_g}{\partial x} \geq 0$ results in

$$\begin{aligned}
-\int_\Omega \tilde{A}\nabla\tilde{g} \cdot \nabla\mathfrak{Q}\tilde{u}_\varepsilon^{(z)} \mathrm{d}\Omega &= -\int \int_{x_{\min}}^{0} \frac{1}{2} v K \exp(x) \mathfrak{Q}\tilde{u}_\varepsilon^{(z)} \mathrm{d}x\mathrm{d}v + \int \left[\frac{1}{2} v K \exp(x) \mathfrak{Q}\tilde{u}_\varepsilon^{(z)}\right]_{x_{\min}}^{0} \mathrm{d}v \\
&\leq -\int \int_{x_{\min}}^{0} \frac{1}{2} v K \exp(x) \mathfrak{Q}\tilde{u}_\varepsilon^{(z)} \mathrm{d}x\mathrm{d}v \\
&= \int_\Omega \frac{1}{2} v \frac{\partial g}{\partial x} \mathfrak{Q}\tilde{u}_\varepsilon^{(z)} \mathrm{d}\Omega \\
&\leq \int_\Omega \frac{1}{2} v \frac{\partial\tilde{g}}{\partial x} \mathfrak{Q}\tilde{u}_\varepsilon^{(z)} \mathrm{d}\Omega.
\end{aligned} \tag{5.32}$$

Finally, applying the Cauchy-Schwarz inequality to (5.31) and the use of (5.32) yields

$$\begin{aligned}-\tilde{a}^H(\tilde{g}, \mathfrak{Q}\tilde{u}_\varepsilon^{(z)}) &\leq \varpi s \left(\frac{1}{2} v_{\max} \|\tilde{g}\|_{\mathcal{V}} + |\tilde{b}_1(v_{\max})| \|\tilde{g}\|_{\mathcal{V}} + (r+1)\|g\|_H\right) \|\mathfrak{Q}\tilde{u}_\varepsilon^{(z)}\|_H \\ &\leq \left(\frac{1}{2} v_{\max} + |\tilde{b}_1(v_{\max})| + r + 1\right) \|\tilde{g}\|_{\mathcal{V}} \|\mathfrak{Q}\tilde{u}_\varepsilon^{(z)}\|_H,\end{aligned}$$

whence (3.15) in Lemma 3.5 is satisfied. It remains to show that the right hand side lies in $L^2(\Omega)$. Since we employ an implicit Euler-method for the initial time steps one has for $z = 1$ with $\tilde{u}_\varepsilon^{(0)} = g$ that the right hand side is given by $f_{\text{inh}}^{(0)}(\varphi) := (g, \varphi)$. Since $g \in L^2(\Omega)$ it follows from Lemma 3.5 that $\tilde{u}_\varepsilon^{(1)}$ lies in $H^2(\Omega)$. For the remaining time steps $2 \leq z \leq \#\mathbb{T}$ one has

$$f_{\text{inh}}^{(z-1)}(\varphi) := -(1-\varpi)s\ a^H(\tilde{u}_\varepsilon^{(z-1)}, \varphi) + (\tilde{u}_\varepsilon^{(z-1)}, \varphi).$$

Thus, it is clear that for $\tilde{u}_\varepsilon^{(z-1)} \in H^2(\Omega)$ for $2 \leq z \leq \#\mathbb{T}$ the right hand side lies in $L^2(\Omega)$. By applying Lemma 3.5 we can conclude that the following estimate holds

$$\|\tilde{u}_\varepsilon^{(z)}\|_{H^2(\Omega)} \lesssim \|\tilde{g}\|_{\mathcal{V}} + \|f_{\text{inh}}^{(z-1)}\|_H + \|\tilde{u}_g\|_{H^2(\Omega)}$$

with constant independent of the penalty parameter $\varepsilon > 0$. Resubstituting gives

$$\|y_\varepsilon^{(z)}\| \lesssim \|g\|_{\mathcal{V}} + \|f_{\text{inh}}^{(z-1)}\|_H + \|g\|_{H^{3/2}(\Upsilon)}.$$

Finally by passing to the limit as $\varepsilon \to 0$ it can be shown that the solution of the semi-discrete Heston variational inequality in Problem 5.8 satisfies $y^{(z)} \in H^2(\Omega)$ for a fixed time step z. □

At this point it should be pointed out that the solution of Problem 5.8 for a Crank-Nicolson scheme at the initial time step $z = 1$ does not lie in $H^2(\Omega)$. This is due to the fact that $y^{(0)} = g(x) \notin H^2(\Omega)$ so the right hand side as defined in (5.23) does not lie in $L^2(\Omega)$ for $z = 1$ and $\varpi = 0.5$.

As mentioned earlier for further numerical computation it is convenient to reformulate Problem 5.8 into a problem with zero Dirichlet boundary condition on Υ and an obstacle function, which is equal to zero.

Problem 5.12 (Semi-discrete variational inequality)**.** *Find $u^{(z)} \in \mathcal{K}^0$ for all $z \in \{1, \cdots, \#\mathbb{T}\}$, such that*

$$\tilde{a}^H(u^{(z)}, \varphi - u^{(z)}) \geq f^{(z-1)}(\varphi - u^{(z)}),$$

for all $\varphi \in \mathcal{K}^0$ with start condition $u^{(0)} = 0$, where the bilinear form is defined as in (5.22) and the right hand side as

$$\begin{aligned} f^{(z-1)}(\varphi - u^{(z)}) := & -(1-\varpi)s\ a^H(u^{(z-1)}, \varphi - u^{(z)}) + (u^{(z-1)}, \varphi - u^{(z)}) \\ & - s a^H(g, \varphi - u^{(z)}) \\ = & -\int_\Omega (1-\varpi) s \tilde{A} \nabla u^{(z-1)} \cdot \nabla(\varphi - u^{(z)}) \mathrm{d}\Omega \\ & -\int_\Omega ((1-\varpi)s(\tilde{\mathbf{b}} \cdot \nabla u^{(z-1)} + r u^{(z-1)}) - u^{(z-1)})(\varphi - u^{(z)}) \mathrm{d}\Omega \\ & -\int_\Omega s\tilde{A}\nabla g \cdot \nabla(\varphi - u^{(z)}) + s(\tilde{\mathbf{b}} \cdot \nabla g + rg)(\varphi - u^{(z)}) \mathrm{d}\Omega. \end{aligned} \tag{5.33}$$

5.2.2. Tensor product B-spline Galerkin discretization

In this section, we will introduce the fully discrete scheme of the semi-discrete Heston variational inequality corresponding to Problem 5.12 based on a spatial tensor product B-spline discretization. To facilitate a pointwise approximation of the Greeks up to order two, we use a cubic tensor product B-spline discretization. As in the case for the Black-Scholes variational inequality, one particular difficulty arises from the fact that the initial condition $g(x, v) = g(x)$ is not differentiable at $x = 0$. In order to derive a stable approximation of the partial derivatives without oscillations, we use a spatial tensor product B-spline discretization with $k-1$ coinciding knots at $x = 0$. Results for a discretization of the Heston variational inequality without internal coinciding knots and a bilinear form as in (2.120), where the boundary integral is set to zero, can be found in [Bos15, Wei14].
Let $\mathcal{V}_h^0 \subset \mathcal{V}^0$ be the finite dimensional tensor-product spline space of fixed order $k \in \mathbb{N}$ with homogeneous boundary conditions on Υ. Let further $\Theta_x := \{\theta_i^{(x)}\}_{i=1}^{n+k}$ be the extended sequence of knots with equidistant grid size $h_{\ell_x} := (x_{\max} - x_{\min})/2^{\ell_x}$ in the x-direction

$$\theta_1^{(x)} = \ldots = \theta_k^{(x)} < \theta_{k+1}^{(x)} < \ldots < \theta_{\wp}^{(x)} = \ldots = \theta_{\wp+k-2}^{(x)} < \ldots < \theta_{n+1}^{(x)} = \ldots = \theta_{n+k}^{(x)} \tag{5.34}$$

where $\theta_1^{(x)} = x_{\min}$, $\theta_{\wp}^{(x)} = \ldots = \theta_{\wp+k-2}^{(x)} = 0$ and $\theta_{n+1}^{(x)} = x_{\max}$. Then we define the B-splines $N_{i,k}(x)$ of order k for $i_x = 1, \ldots, n$ recursively as defined in (3.62).

Let further $\Theta_v := \{\theta_{i_v}^{(v)}\}_{i_v=1}^{m+k}$ be the extended sequence of knots with equidistant grid size $h_{\ell_v} := (v_{\max} - v_{\min})/2^{\ell_v}$ $\ell_v \in \mathbb{N}$ in the v-direction such that

$$\theta_1^{(v)} = \ldots = \theta_{\mathsf{p}}^{(v)} = v_{\min} < \theta_{k+1}^{(v)} < \ldots < \theta_m^{(v)} < v_{\max} = \theta_{m+1}^{(v)} = \ldots = \theta_{m+k}^{(v)}. \tag{5.35}$$

Then we can define the B-Splines $N_{i_v,k}(v)$ analogously to (3.62). Let $\mathbf{i} := (i_x, i_v) \in \mathbb{N}^2$ a multi-index, then the $\mathbf{i}$'th tensor product B-spline for the rectangular domain $\Omega \subset \mathbb{R}^2$ is given by $N_{\mathbf{i},k}(x,v) := N_{i_x,k}(x)N_{i_v,k}(v)$. Finally, the finite dimensional tensor product B-spline space is given by

$$\mathcal{V}_h^0 := \operatorname{span}\{N_{\mathbf{i},k}(x,v) = N_{i_x,k}(x)N_{i_v,k}(v) : \mathbf{i} = (i_x, i_v) \in \mathcal{I}\} \subset \mathcal{V}, \tag{5.36}$$

with index set $\mathcal{I}$. For the set of boundary conditions according to (2.110) the index set $\mathcal{I}$ is defined as $\mathcal{I}_D := \{i_x = 2, \ldots, n-1, i_v = 2, \ldots, m\}$ or for the boundary conditions in (2.111) as $\mathcal{I}_N := \{i_x = 1, \ldots, n-1, i_v = 2, \ldots, m\}$. Consequently, one has $n := 2^{\ell_x} + 2k - 3$ B-Spline basis functions in the x-direction due to the $k-2$ additional B-Splines in $x = 0$ and $m := 2^{\ell_v} + k - 1$ B-spline basis functions in the v-direction. Then the dimension of the discrete solution space for the set of boundary conditions in (2.110) is

$$\dim(\mathcal{V}_h^0) := \#\mathcal{I}_D := NM = (n-2)(m-1) = (2^{\ell_x} + 2k - 5)(2^{\ell_v} + k - 2) \tag{5.37}$$

and for the boundary conditions in (2.111) one has

$$\dim(\mathcal{V}_h^0) := \#\mathcal{I}_N := NM = (n-1)(m-1) = (2^{\ell_x} + 2k - 4)(2^{\ell_v} + k - 2). \tag{5.38}$$

Furthermore, the discrete convex set is defined as

$$\mathcal{K}_h^0 := \{\varphi_h \in H : \varphi_h \geq 0\} \cap \mathcal{V}_h^0. \tag{5.39}$$

Note that $\mathcal{K}_h^0 = \mathcal{K}^0 \cap \mathcal{V}_h^0$ is satisfied because the obstacle is equal to zero and no approximation of the obstacle is needed.
For the solution $u_h^{(z)} \in \mathcal{K}_h^0$ and test function $\varphi_h \in \mathcal{K}_h^0$ we obtain

$$\begin{aligned} u_h^{(z)} := u_h^{(z)}(x,v) &:= \sum_{\mathbf{i}\in\mathcal{I}} u_{\mathbf{i}}^{(z)} N_{\mathbf{i},k}(x,v) =: (\mathbf{u}^{(z)})^T \mathbf{N}, \\ \varphi_h := \varphi_h(x,v) &:= \sum_{\mathbf{i}\in\mathcal{I}} \varphi_{\mathbf{i}} N_{\mathbf{i},k}(x,v) =: (\boldsymbol{\varphi})^T \mathbf{N}. \end{aligned} \tag{5.40}$$

Due to the non-negativity of B-splines it can be ensured that $u_h^{(z)} \in \mathcal{K}_h^0$ by claiming that the coefficients satisfy $\mathbf{u}^{(z)} \in \mathbf{K}$, where the discrete convex set $\mathbf{K}$ is given by

$$\mathbf{K} := \{\boldsymbol{\varphi} \in \mathbb{R}^{NM} : \boldsymbol{\varphi} \geq \mathbf{0}\}. \tag{5.41}$$

For the componentwise positivity of the coefficients $\varphi_{\mathbf{i}} \geq 0$ for all $\mathbf{i} \in \mathcal{I}$ we write $\boldsymbol{\varphi} \geq \mathbf{0}$. Inserting the elements $u_h, \varphi_h \in \mathcal{K}_h$ into the bilinear form of Problem 5.12 leads to

$$\tilde{a}^H(u_h^{(z)}, \varphi_h - u_h^{(z)}) =: (\boldsymbol{\varphi} - \mathbf{u}^{(z)})^T C \mathbf{u}^{(z)}, \tag{5.42}$$

where the matrix $C \in \mathbb{R}^{NM \times NM}$ is defined as

$$C := \varpi s \mathfrak{Y} + \varpi s \mathfrak{X} + (\varpi s r + 1)\Xi \tag{5.43}$$

with $\mathfrak{Y},\ \mathfrak{X}, \Xi \in \mathbb{R}^{NM \times NM}$ given by

$$\mathfrak{Y}_{\mathbf{j},\mathbf{i}} := \int_\Omega (\tilde{A} \nabla N_{\mathbf{i},k}) \cdot \nabla N_{\mathbf{j},k} d\Omega, \quad \mathfrak{X}_{\mathbf{j},\mathbf{i}} := \int_\Omega \tilde{\mathbf{b}} \cdot \nabla N_{\mathbf{i},k} N_{\mathbf{j},k} d\Omega, \quad \Xi_{\mathbf{j},\mathbf{i}} := \int_\Omega N_{\mathbf{i},k} N_{\mathbf{j},k} d\Omega.$$

An analogous approach leads to the following right hand side

$$f^{(z-1)}(\varphi_h - u_h^{(z)}) =: (\boldsymbol{\varphi} - \mathbf{u}^{(z)})^T \mathbf{f}^{(z-1)}, \tag{5.44}$$

where $\mathbf{f}^{(z-1)}$ is defined as

$$\mathbf{f}^{(z-1)} := -(1-\varpi)s\,(\mathfrak{Y} + \mathfrak{X})\,\mathbf{u}^{(z-1)} - ((1-\varpi)sr - 1)\Xi \mathbf{u}^{(z-1)} - s\boldsymbol{\delta} - s\boldsymbol{\zeta}, \tag{5.45}$$

with $\boldsymbol{\delta}$ and $\boldsymbol{\zeta} \in \mathbb{R}^{NM}$ given by

$$\boldsymbol{\delta}_{\mathbf{i}} := \int_\Omega (\tilde{A} \nabla g) \cdot \nabla N_{\mathbf{i},k} d\Omega \qquad \boldsymbol{\zeta}_{\mathbf{i}} := \int_\Omega (\tilde{\mathbf{b}} \cdot \nabla g + rg) N_{\mathbf{i},k} d\Omega.$$

Finally, the fully discrete variational inequality can be formulated as follows:

Problem 5.13 (Discrete variational inequality)**.** *Find* $\mathbf{u}^{(z+1)} \in \mathbf{K}$ *for all* $z \in \{0, \ldots, \#\mathbb{T} - 1\}$ *such that*

$$(\boldsymbol{\varphi} - \mathbf{u}^{(z)})^T (C\mathbf{u}^{(z)} - \mathbf{f}^{(z-1)}) \geq \mathbf{0} \tag{5.46}$$

is fulfilled for all $\boldsymbol{\varphi} \in \mathbf{K}$*.*

For further numerical computations, it is convenient to rewrite Problem 5.13 to the following linear complementarity problem:

Problem 5.14 (Discrete linear complementary problem)**.** *Find* $\mathbf{u}^{(z+1)} \in \mathbf{K}$ *for all* $z \in \{0, \ldots, \#\mathbb{T} - 1\}$ *such that*

$$(\mathbf{u}^{(z)})^T (C\mathbf{u}^{(z)} - \mathbf{f}^{(z-1)}) = 0, \quad C\mathbf{u}^{(z)} - \mathbf{f}^{(z-1)} \geq \mathbf{0}, \quad \mathbf{u}^{(z)} \geq \mathbf{0}. \tag{5.47}$$

Due to the tensor product structure, all matrices in (5.43) can be expressed as sums of Kronecker products of matrices with respect to one coordinate direction x or v as follows

$$\mathfrak{Y} := \sum_{\mathfrak{l}=1}^{3} \left(\mathfrak{Y}^{\mathfrak{l},x} \otimes \mathfrak{Y}^{\mathfrak{l},v} \right), \qquad \mathfrak{X} := \sum_{\mathfrak{l}=1}^{2} \left(\mathfrak{X}^{\mathfrak{l},x} \otimes \mathfrak{X}^{\mathfrak{l},v} \right) \quad \text{and} \quad \Xi := \Xi^{1,x} \otimes \Xi^{1,v}. \tag{5.48}$$

Due to the local support of B-splines, $\mathcal{Q}^x := \operatorname{supp}(N_{i_x}) \cap \operatorname{supp}(N_{j_x})$ and $\mathcal{Q}^v := \operatorname{supp}(N_{i_v}) \cap \operatorname{supp}(N_{j_v})$, the matrices with respect to one coordinate direction result in

$$\begin{aligned}
\mathfrak{Y}^{1,x}_{j_x,i_x} &:= \int_{\mathcal{Q}^x} N'_{i_x}(x)N'_{j_x}(x)\mathrm{d}x, \qquad & \mathfrak{Y}^{1,v}_{j_v,i_v} &:= \int_{\mathcal{Q}^v} \frac{1}{2} v N_{i_v}(v)N_{j_v}(v)\mathrm{d}v, \\
\mathfrak{Y}^{2,x}_{j_x,i_x} &:= \int_{\mathcal{Q}^x} N_{i_x}(x)N'_{j_x}(x)\mathrm{d}x, \qquad & \mathfrak{Y}^{2,v}_{j_v,i_v} &:= \int_{\mathcal{Q}^v} v\rho\xi N'_{i_v}(v)N_{j_v}(v)\mathrm{d}v, \\
\mathfrak{Y}^{3,x}_{j_x,i_x} &:= \int_{\mathcal{Q}^x} N_{i_x}(x)N_{j_x}(x)\mathrm{d}x, \qquad & \mathfrak{Y}^{3,v}_{j_v,i_v} &:= \int_{\mathcal{Q}^v} \frac{1}{2} v\xi^2 N'_{i_v}N'_{j_v}\mathrm{d}v,
\end{aligned}$$

$$\begin{aligned}
\mathfrak{X}^{1,x}_{j_x,i_x} &:= \int_{\mathcal{Q}^x} N'_{i_x}(x)N_{j_x}\mathrm{d}x, \qquad & \mathfrak{X}^{1,v}_{j_v,i_v} &:= \int_{\mathcal{Q}^v} \left(\frac{1}{2}v - r\right) N_{i_v}N_{j_v}\mathrm{d}v, \\
\mathfrak{X}^{2,x}_{j_x,i_x} &:= \int_{\mathcal{Q}^x} N_{i_x}(x)N_{j_x}\mathrm{d}x, \qquad & \mathfrak{X}^{2,v}_{j_v,i_v} &:= \int_{\mathcal{Q}^v} \left(\kappa(v-\gamma) + \lambda(v) + \frac{1}{2}\xi^2\right) N'_{i_v}(v)N_{j_v}(v)\mathrm{d}v
\end{aligned}$$

and

$$\Xi^{1,x}_{j_x,i_x} := \int_{\mathcal{Q}^x} N_{i_x}(x)N_{j_x}(x)\mathrm{d}x, \qquad \Xi^{1,v}_{j_v,i_v} := \int_{\mathcal{Q}^v} N_{i_v}(v)N_{j_v}(v)\mathrm{d}v \tag{5.49}$$

for $i_x, j_x \in \{2, \dots, n-1\}$ and $i_v, j_v \in \{2, \dots, m\}$ for the set of boundary conditions in (2.110) or $i_v, j_v \in \{1, \dots, m\}$ for the set of boundary conditions in (2.111). The matrices with one coordinate direction can be set up efficiently by applying the Neville-scheme for the evaluation of B-splines combined with an appropriate quadrature rule to compute the integrals. Since B-splines are piecewise polynomials, the matrices above can be computed exactly with the Gauss quadrature rule.
It remains to present the set up of the vectors $\boldsymbol{\delta}$ and $\boldsymbol{\zeta}$. Since $g := g(x)$ is independent of the variable v one has

$$\boldsymbol{\delta} := \boldsymbol{\delta}^x \otimes \boldsymbol{\delta}^v \quad \text{and} \quad \boldsymbol{\zeta} := \sum_{\mathfrak{l}=1}^{2} \left(\boldsymbol{\zeta}^{\mathfrak{l},x} \otimes \boldsymbol{\zeta}^{\mathfrak{l},v}\right) \tag{5.50}$$

with

$$\boldsymbol{\delta}^x_{i_x} := \int g'(x)N'_{i_x}(x)dx, \qquad \boldsymbol{\delta}^v_{i_v} := \int \frac{1}{2} v N_{i_v}(v)dv$$

and

$$\begin{aligned}
\boldsymbol{\zeta}^{1,x}_{i_x} &:= \int g'(x)N_{i_x}(x)dx, \qquad & \boldsymbol{\delta}^{1,v}_{i_v} &:= \int \left(\frac{1}{2}v - r\right) N_{i_v}(v)dv \\
\boldsymbol{\zeta}^{2,x}_{i_x} &:= \int rg(x)N_{i_x}(x)dx, \qquad & \boldsymbol{\delta}^{2,v}_{i_v} &:= \int N_{i_v}(v)dv
\end{aligned} \tag{5.51}$$

for $i_x \in \{2, \dots, n-1\}$ and $i_v \in \{2, \dots, m\}$ for the set of boundary conditions in (2.110) or $i_v \in \{1, \dots, m\}$ for the set of boundary conditions in (2.111).
Since $g(x)$ is an exponential function, employing a Gauss quadrature to the integrals, which includes $g(x)$, is not exact but gives a good approximation if enough Gauss points are chosen. A detailed derivation of the matrices and vectors corresponding to the parabolic variational inequality for the set of boundary conditions in (2.110), where the boundary integral in (2.121) are set to zero combined with B-splines without coinciding knots at $x = 0$ can be found in [Bos15, Wei14].
Due to Theorem 5.11 and Theorem 3.24 it is clear that for a linear B-spline discretization and an accurate approximation of the initial condition ($x = 0$ aligned with one knot) an optimal convergence rate in $H^1(\Omega)$ can be expected for the semi-discrete Heston Problem. It is still unclear if the solution of the semi-discrete Heston Problem satisfies the smoothness assumption of Theorem 3.25. Since the obstacle function $g(x)$ is smooth, where the constraints are binding (i.e. for $S \leq S_f(\tau, v) < K$ for a put option) and the right hand side is smooth, where the constraints are non-binding, it can be expected that the solution satisfies $y^{(z)} \in W^{l,p}$ for $z \geq 2$, $1 < p < \infty$ and $l < 1 + \frac{1}{p}$ as required for the a priori estimates in Theorem 3.25. Hence, for an accurate approximation of the initial condition ($x = 0$ aligned with $k - 1$ knots) one can expect a convergence rate of $\mathcal{O}(h^{3/2-\varepsilon})$ in the $H^1(\Omega)$ norm for quadratic and cubic B-splines, if the smoothness assumption in Theorem 3.25 is satisfied.

6. Approximation of the Greeks for American options

In finance related businesses, as described in Section 2.1.5, a particular focus of interest lies on the approximation of the Greeks, which measure the sensitivity of the option value to a small change in a given model parameter. Formally, the Greeks are defined as the partial derivatives with respect to the model parameters up to order two. In this chapter we present the approximation of the Greeks with higher order B-splines in the Black-Scholes and Heston model. It should be mentioned at this point that the authors in [Hol04] and [Wei14, Bos15] also proposed a B-spline discretization to approximate the Greeks for an American put option in the Black-Scholes and Heston model. The authors use a B-spline discretization without internal coinciding knots there. In particular, for the two dimensional Heston problem in [Wei14, Bos15] it can be observed that the approximation of the partial derivatives with respect to the underlying price exhibits oscillations in a neighborhood of the point where the underlying price is equal to the strike price. To overcome this problem the following approach is recommendable: For the valuation of an American option the initial condition is often given by a piecewise smooth payoff function, such as defined in (2.1), (2.2) or (2.3), so the following aspects are necessary to find a pointwise accurate approximation of the partial derivatives (without oscillations) up to order two:

i) An implicit Euler scheme for the initial time steps has to be employed, since this method smoothen the numerical solution caused by irregularities of the initial data.

ii) As the initial condition, given by a payoff function, is only continuous at this point, where the underlying price is equal to the strike price, a cubic B-spline discretization with $k-1$ repeating knots at this point should be used.

Moreover, it should be clear from the discussions in Section 3.1.1 that the semi-discrete solution of the Black-Scholes or Heston problem cannot be more than once classically differentiable on the whole spatial domain and the second derivative must have discontinuities at the free boundary. As for the Black-Scholes variational inequality for a put and call option it can easily checked with the results in Lemma 4.4 and Theorem 4.5 that the semi-discrete solution satisfies $y^{(z)}(S) \in H^2(I)$. Applying the Sobolev embedding theorem for one dimensional problems implies that the solution lies in $C^1(\bar{I})$. Regarding the two dimensional Heston variational inequality the existence of a classical first order partial derivative of the semi-discrete solution is still an open question. The obstacle function $g(x)$ is smooth enough, where the constraints are binding (i.e. for $S \leq S_f(\tau, v) < K$ for a put option) and the right hand side is smooth enough,

where the constraints are non-binding, so it can be expected that the solution satisfies $y^{(z)} \in C^1(\overline{\Omega})$.

6.1. Greeks in Black-Scholes Model

After the discrete Black-Scholes variational inequality defined in Problem 4.13 is solved by an iterative scheme and in consideration of the transformation $\tau^{(z)} := T - t^{(z)}$ the approximation of the *American option value* can be computed by

$$V_h(t^{(z)}, S) = u_h(\tau^{(z)}, S) + \mathcal{H}(S) \text{ for each } z \in \{1, \dots, \#\mathbb{T}\}. \tag{6.1}$$

As stated previously, the solution is twice differentiable almost everywhere except at the free boundary. Therefore one can expect that the discrete solution satisfies

$$V_h(t^{(z)}, S) \in C^2(I \setminus \mathfrak{S}^{(z)}_{S_f^h}) \text{ for each } z \in \{1, \dots, \#\mathbb{T}\}, \tag{6.2}$$

where $\mathfrak{S}^{(z)}_{S_f^h}$ denotes the union of the supports of B-spline basis functions whose support intersects $S_f^h(t^{(z)})$. The approximation of *Delta*, as defined in (2.39), results in

$$\Delta_h(t^{(z)}, S) := \frac{\partial V_h(t^{(z)}, S)}{\partial S} \approx \frac{\partial (u_h(\tau^{(z)}, S))}{\partial S} + \mathcal{H}'_+(S), \tag{6.3}$$

where $\mathcal{H}'_+(S)$ denotes the derivative of $\mathcal{H}$ with a right-sided derivative at the discontinuity points. Furthermore, the approximation of *Gamma*, as defined in (2.40), can be computed by

$$\Gamma_h(t^{(z)}, S) := \frac{\partial^2 V_h(t^{(z)}, S)}{\partial S^2} \approx \frac{\partial^2 (u_h(\tau^{(z)}, S))}{\partial S^2} \quad \text{for} \quad S \notin \mathfrak{S}^{(z)}_{S_f^h} \tag{6.4}$$

where the second derivative with a right-sided second derivative at the discontinuity points satisfies $\mathcal{H}''_+(S) = 0$ for payoff functions given in (2.1), (2.2) or (2.3).

6.2. Greeks in Heston's Model

After solving the discrete Heston variational inequality defined in Problem 5.14 by an iterative scheme and in consideration of the transformations $\tau^{(z)} := T - t^{(z)}$ and $x = \log\left(\frac{S}{K}\right)$ the *American put option value* is approximated by

$$V_h(t^{(z)}, S, v) = y_h(\tau^{(z)}, x, v) = u_h(\tau^{(z)}, x, v) + g(x) \text{ for each } z \in \{1, \dots, \#\mathbb{T}\}. \tag{6.5}$$

Furthermore, resubstituting the coordinates of the approximated free boundary yields

$$S_f^h(t^{(z)}, v) = K \exp\left(x_f^h(\tau^{(z)}, v)\right). \tag{6.6}$$

Moreover, the solution for the discrete Heston variational inequality with set of boundary conditions in (2.111) satisfies

$$V_h(t^{(z)}, S, v) \in C^2(\Omega_{\mathcal{L}} \setminus \mathfrak{S}^{(z)}_{S_f^h}) \text{ for each } z \in \{1, \dots, \#\mathbb{T}\}, \tag{6.7}$$

where $\mathfrak{S}^{(z)}_{S_f^h}$ denotes the union of the supports of B-spline basis functions whose support intersects $S_f^h(t^{(z)}, v)$. Taking into account the log-transformation of the underlying the approximation of *Delta* results in

$$\begin{aligned}\Delta_h(t^{(z)}, S, v) := \frac{\partial V_h(t^{(z)}, S, v)}{\partial S} &= \frac{1}{K\exp(x)}\left(\frac{\partial y_h(\tau^{(z)}, x, v)}{\partial x}\right) \\ &\approx \frac{1}{K\exp(x)}\left(\frac{\partial u_h(\tau^{(z)}, x, v)}{\partial x} + g'_+(x)\right),\end{aligned} \tag{6.8}$$

where $g'_+(x) := g'_+(x, v)$ denotes the partial derivative of $g(x, v)$ regarding x with a right-sided derivative at discontinuity points $(x, v) = (0, v)$.
Since $g'_+(x) = g''_+(x)$ is satisfied, where $g''_+(x) := g''_+(x, v)$ denotes the second partial derivative of $g(x, v)$ regarding x with a right-sided derivative at discontinuity points $(x, v) = (0, v)$, *Gamma* can be approximated by

$$\begin{aligned}\Gamma_h(t^{(z)}, S, v) :&= \frac{\partial^2 V_h(t^{(z)}, S, v)}{\partial S^2} = \frac{1}{(K\exp(x))^2}\left(\frac{\partial^2 y_h(\tau^{(z)}, x, v)}{\partial x^2} - \frac{\partial y_h(\tau^{(z)}, x, v)}{\partial x}\right) \\ &\approx \frac{1}{(K\exp(x))^2}\left(\frac{\partial^2 u_h(\tau^{(z)}, x, v)}{\partial x^2} + g''_+(x) - \frac{\partial y_h(\tau^{(z)}, x, v)}{\partial x} - g'_+(x)\right) \\ &= \frac{1}{(K\exp(x))^2}\left(\frac{\partial^2 u_h(\tau^{(z)}, x, v)}{\partial x^2} - \frac{\partial u_h(\tau^{(z)}, x, v)}{\partial x}\right).\end{aligned} \tag{6.9}$$

Further, having in mind that the payoff function $g(x)$ is independent of the variance v, the approximation of *Vega*, as stated in (2.41), results in

$$\nu_h(t^{(z)}, S, v) := \frac{\partial V_h(t^{(z)}, S, v)}{\partial v} = \frac{\partial u_h(\tau^{(z)}, x, v)}{\partial v}. \tag{6.10}$$

For the efficient evaluation of the solution and the corresponding derivatives we employ the Neville-scheme for splines as described in Theorem 3.18 and its derivatives as described in Corollary 3.19. Note that a pointwise approximation of the first or second derivative is possible, if $k \geq 3$ or $k \geq 4$, respectively. Otherwise, the numerical solution has to be interpolated with a smooth enough polynomial or spline.

7. Solution algorithms to solve discrete variational inequalities

In this chapter we study several methods to solve variational inequalities arising from a B-spline Galerkin discretization.
Let $\Omega \subset \mathbb{R}^d$, then we define by $\mathcal{V}_L \subset \mathcal{V}^0$ the finite dimensional tensor product spline space of piecewise polynomials of order k with dimension $\dim(\mathcal{V}_L) = \#\mathcal{I}_L$ on level L. Let further $\mathcal{K}_L^0 := \{u_L \in \mathcal{V}_L : u_L(x) \geq \widetilde{\psi}_L \text{ for all } x \in \Omega\}$ be the discrete convex set on level L, where $\widetilde{\psi}_L = \sum_{\mathbf{i} \in \mathcal{I}_L} \widetilde{\boldsymbol{\psi}}_{\mathbf{i}} N_{\mathbf{i},k}$ denotes the obstacle in spline representation. Then we consider the following problem: Find $u_L \in \mathcal{K}_L^0$ such that

$$a(u_L, \varphi_L - u_L) \geq f(\varphi_L - u_L) \text{ for all } \varphi_L \in \mathcal{K}_L^0. \tag{7.1}$$

Or equivalently, written as linear complementarity problem: Find $u_L \in \mathcal{V}_L$ such that

$$\begin{aligned} (u_L - \widetilde{\psi})(\mathcal{A}_L u_L - f_L) &= 0 \\ \mathcal{A}_L u_L - f_L &\geq 0 \\ u_L &\geq \widetilde{\psi}, \end{aligned} \tag{7.2}$$

where $\mathcal{A}_L$ is the discrete form of the Riesz operator $\mathcal{A} : \mathcal{V}^0 \to \mathcal{V}^*$ on level L defined by $\langle \mathcal{A}u, \varphi \rangle := a(u, \varphi)$ for all $\varphi \in \mathcal{V}^0$. To project the solution $u_L(x)$ in the discrete convex set $\mathcal{K}_L^0$ for all $x \in \Omega$ one essential construction for higher order B-splines is the comparison of the B-spline expansion coefficients instead of function values, which was first established by [Hol04, HK07]. More precisely, due to the non-negativity of B-splines $N_{\mathbf{i},k} \geq 0$ it can be ensured that the solution $u_L = \sum_{\mathbf{i} \in \mathcal{I}_L} u_{\mathbf{i}} N_{\mathbf{i},k}$ lies in the convex set $\mathcal{K}_h^0$ by requiring that the coefficients satisfy

$$\mathbf{u}_L \in \mathbf{K}_L := \{\mathbf{u}_L \in \mathbb{R}^{\#\mathcal{I}_L} : \mathbf{u}_L \geq \widetilde{\boldsymbol{\psi}}_L\}. \tag{7.3}$$

Then problem (7.1) can be written in matrix vector notation: Find $\mathbf{u}_L \in \mathbf{K}_L$ such that

$$(\boldsymbol{\varphi}_L - \mathbf{u}_L)^T (C_L \mathbf{u}_L - \mathbf{f}_L) \geq \mathbf{0} \text{ for all } \boldsymbol{\varphi}_L \in \mathbf{K}_L. \tag{7.4}$$

Or equivalently formulated as discrete linear complementarity problem: Find $\mathbf{u}_L \in \mathbb{R}^{\#\mathcal{I}_L}$ such that

$$\begin{aligned} (\mathbf{u}_L - \widetilde{\boldsymbol{\psi}}_L)^T (C_L \mathbf{u}_L - \mathbf{f}_L) &= 0, \\ C_L \mathbf{u}_L - \mathbf{f}_L &\geq \mathbf{0}, \\ \mathbf{u}_L &\geq \widetilde{\boldsymbol{\psi}}_L, \end{aligned} \tag{7.5}$$

where $C_L \in \mathbb{R}^{\#\mathcal{I}_L \times \#\mathcal{I}_L}$ denotes the discretization matrix and $\mathbf{f}_L \in \mathbb{R}^{\#\mathcal{I}_L}$ the discrete right hand side.

7.1. Projected iterative methods

In the next section we describe the *PJOR (Projected Jacobi overrelaxation method)* and *PSOR (Projected successive overrelaxation)* method to solve the linear complementary problem (7.5) in matrix vector notation. To simplify the notation we omit the level L in this section. Since $\mathbf{u} - \widetilde{\boldsymbol{\psi}} > C\mathbf{u} - \mathbf{f}$, if $C\mathbf{u} - \mathbf{f} = 0$, and $C\mathbf{u} - \mathbf{f} > \mathbf{u} - \widetilde{\boldsymbol{\psi}}$, if $\mathbf{u} = \widetilde{\boldsymbol{\psi}}$, the linear complementarity problem is equivalent to:
Find $\mathbf{u} \in \mathbb{R}^{\#\mathcal{I}}$ such that

$$\min\{\mathbf{u} - \widetilde{\boldsymbol{\psi}}, C\mathbf{u} - \mathbf{f}\} = \mathbf{0} \tag{7.6}$$

is satisfied. Let $u^{(\varsigma)}$ denote the ς'th iterate then (7.6) can be solved by the following Jacobi iterate

$$u_{\mathbf{i}}^{(\varsigma+1)} = \max\left\{\widetilde{\psi}_{\mathbf{i}}, (1-\omega)u_{\mathbf{i}}^{(\varsigma)} + \frac{\omega}{C_{\mathbf{i},\mathbf{i}}}\left(f_{\mathbf{i}} - \sum_{\mathbf{j}=1}^{\mathbf{i}-1} C_{\mathbf{i},\mathbf{j}} u_{\mathbf{j}}^{(\varsigma)} - \sum_{\mathbf{j}=\mathbf{i}+1}^{\#\mathcal{I}} C_{\mathbf{i},\mathbf{j}} u_{\mathbf{j}}^{(\varsigma)}\right)\right\}, \tag{7.7}$$

with the relaxation parameter $\omega \in (0, 2]$, if the diagonal matrix has only positive entries. It is easy to see that for the computation of $u_{\mathbf{i}+1}^{(\varsigma+1)}$ the iterate $u_{\mathbf{i}}^{(\varsigma+1)}$ can be used. This leads to the following PSOR method

$$\begin{aligned} u_{\mathbf{i}}^{(\varsigma+1)} &= \max\left\{\widetilde{\psi}_{\mathbf{i}}, (1-\omega)u_{\mathbf{i}}^{(\varsigma)} + \frac{\omega}{C_{\mathbf{i},\mathbf{i}}}\left(f_{\mathbf{i}} - \sum_{\mathbf{j}=1}^{\mathbf{i}-1} C_{\mathbf{i},\mathbf{j}} u_{\mathbf{j}}^{(\varsigma+1)} - \sum_{\mathbf{j}=\mathbf{i}+1}^{\#\mathcal{I}} C_{\mathbf{i},\mathbf{j}} u_{\mathbf{j}}^{(\varsigma)}\right)\right\} \\ &= \max\left\{\widetilde{\psi}_{\mathbf{i}}, u_{\mathbf{i}}^{(\varsigma)} + \frac{\omega}{C_{\mathbf{i},\mathbf{i}}}\left(f_{\mathbf{i}} - \sum_{\mathbf{j}=1}^{\mathbf{i}-1} C_{\mathbf{i},\mathbf{j}} u_{\mathbf{j}}^{(\varsigma+1)} - \sum_{\mathbf{j}=\mathbf{i}}^{\#\mathcal{I}} C_{\mathbf{i},\mathbf{j}} u_{\mathbf{j}}^{(\varsigma)}\right)\right\} \end{aligned} \tag{7.8}$$

with initial iterate $\mathbf{u}^{(0)} \geq \widetilde{\boldsymbol{\psi}}$ see e.g. [BC83, Cry71, Glo71]. For $\omega = 1$ the above scheme is also known as *projective Gauss-Seidel* (PGS) method.
It is known from [Pan84] that the optimal convergence rate of such iterative methods can be estimated by

$$\|\mathbf{u}^* - \mathbf{u}^{(k)}\| \leq \varrho(B)\|\mathbf{u}^* - \mathbf{u}^{(k-1)}\|, \tag{7.9}$$

where $\varrho(B)$ denotes the spectral radius of the iteration Matrix B. If $\varrho(B) < 1$ the method is said to be convergent. For the PJOR method the iteration matrix is given by $B := I - \omega D^{-1}C$ and for the PSOR method by $B := I - \omega(D - L)^{-1}C$. A decisive disadvantage of classical iterative methods applied to complementarity problems arising from discretization of variational inequalities is that the convergence rate depends on the spectral condition number. That means, if C is a symmetric positive matrix the optimal convergence rate for the PJOR methods is estimated by

$$\varrho^*(I - \omega D^{-1}C) = \frac{\kappa(D^{-1}C) - 1}{\kappa(D^{-1}C) + 1} \leq 1 - \frac{2}{\kappa(D^{-1})\kappa(C) + 1}. \tag{7.10}$$

The condition number regarding the spectral matrix norm $\|\cdot\|$ of the d-dimensional tensor product B-spline discretization matrix C increases with the grid size h, i.e.

$$\kappa(C) = \|C\|\|C^{-1}\| = \lambda_{\max}(C)\lambda_{\max}(C^{-1}) \lesssim h^{d-2}h^{-d} = h^{-2}, \tag{7.11}$$

with a constant depending on k but independent of h (see Lemma 3 in [GKT13]). Finally, exploiting $\kappa(D^{-1}) \lesssim 1$ the optimal convergence rate can be estimated by

$$\varrho^*(I - \omega D^{-1}C) \leq 1 - \mathcal{O}(h^2). \tag{7.12}$$

and as $h \to 0$ then the convergence rate tends to $\varrho^* \to 1$.
The advantage of the projected Jacobi like version is that the iterates can be computed in parallel. But on the other hand, the Jacobi like version might lead to worse convergence in comparison to the PSOR method. It can be shown that the convergence rate for the PJOR method is larger than the PSOR method, i.e. we have $\varrho^*(I - \omega(D - L)^{-1}C) \leq \varrho^*(I - \omega D^{-1}C)$.

7.2. Monotone multigrid method

7.2.1. The basic monotone multigrid algorithm

In order to accelerate basic iterative schemes as described in Section 7.1, we introduce the *monotone multigrid method* (MMG) for the fast and efficient numerical solution of the semi-discrete variational inequality arising in the valuation of an American option in the next section.
Monotone multigrid methods to solve elliptic variational inequalities for linear hat functions in a finite element setting have been developed in [Man84, Kor94]. In comparison to standard multigrid methods for the unrestricted case monotone multigrid methods ensure that the constraints are met over all grid levels. In particular the author in [Kor94] improves the approach of [Man84] by extending the set of search directions. This approach, also known as truncated monotone multigrid method, profits from the nodal structure of linear basis functions.
Previously, multigrid methods have been proposed for such problems using a standard finite difference or finite element approach [BC83, HM83, Hop87], where not all of them ensure that the constraints are satisfied.
A monotone multigrid method for the valuation of an American option for higher order B-splines without coinciding knots has been developed by [Hol04, HK07]. Since, as mentioned earlier, in finance the initial condition of the parabolic variational inequality is often given as a piecewise linear function and a discretization with coinciding knots at the critical points stabilizes the approximation of the partial derivatives, so we present a monotone multigrid method for B-splines with coinciding knots. Therefore, monotone coarse grid approximations for B-splines with coinciding knots need to be constructed.

First, the basic idea of multigrid methods is summarized. For the implementation of the multigrid method the following hierarchy of grids and approximation spaces

$$\mathcal{V}_{\ell_{\min}} \subset \mathcal{V}_{\ell_{\min}+1} \subset \ldots \subset \mathcal{V}_{\ell} \subset \ldots \subset \mathcal{V}_{\ell_{\max}} := \mathcal{V}_L \subset \mathcal{V}$$

is needed, where $\ell = (\ell_1, \ldots, \ell_d)$ denotes the level regarding the different dimensions. On the highest level L the following linear complementary problem has to be solved,

$$\begin{aligned} (u_L - \widetilde{\psi}_L)^T (\mathcal{A}_L u_L - f_L) &= 0 \\ \mathcal{A}_L u_L &\geq f_L \\ u_L &\geq \widetilde{\psi}_L, \end{aligned} \tag{7.13}$$

where $\mathcal{A}_L$ denotes the linear operator, u_L the solution and $\widetilde{\psi}_L$ the obstacle function on level L. Let $u_L^{(\varsigma,0)}$ be the approximation of the solution in the ς'th iteration of the MMG method. To eliminate high-frequency portions of the error, η_1 a priori smoothing steps are performed. For the smoothing steps we use the PGS method, see e.g. [BC83]. Denoting the iterate after η_1 iterations of a smoother by $u_L^{(\varsigma,1)} := (\mathsf{Sm}(u_L^{(\varsigma,0)}))^{\eta_1}$, the error by $\mathsf{v}_L := u_L - u_L^{(\varsigma,1)}$ and the defect by $d_L := f_L - \mathcal{A}_L u_L^{(\varsigma,1)}$, we get the following defect problem

$$\begin{aligned} (\mathsf{v}_L - \widetilde{\psi}_L + u_L^{(\varsigma,1)})^T (\mathcal{A}_L \mathsf{v}_L - d_L) &= 0 \\ \mathcal{A}_L \mathsf{v}_L &\geq d_L \\ \mathsf{v}_L &\geq \widetilde{\psi}_L - u_L^{(\varsigma,1)}. \end{aligned} \tag{7.14}$$

Classical iterative methods eliminate high-frequency portions of the error after few iteration steps. Knowing that smooth functions can be approximated without essential loss of information on a coarser grid $\mathcal{V}_{L-1}$, the defect problem is transported to a coarser grid by using different restriction operators $\mathcal{R}, \tilde{r} : \mathcal{V}_L \to \mathcal{V}_{L-1}$,

$$\begin{aligned} (\mathsf{v}_{L-1} - \widetilde{\psi}_{L-1})^T (\mathcal{A}_{L-1} \mathsf{v}_{L-1} - f_{L-1}) &= 0 \\ \mathcal{A}_{L-1} \mathsf{v}_{L-1} &\geq f_{L-1} \\ \mathsf{v}_{L-1} &\geq \widetilde{\psi}_{L-1}, \end{aligned} \tag{7.15}$$

where $f_{L-1} := \mathcal{R} d_L$ and $\widetilde{\psi}_{L-1} := \tilde{r}(\widetilde{\psi}_L - u_L^{(\varsigma,1)})$. After solving this defect problem on a coarser grid with less effort, the approximation of the error v_{L-1} is transported back to the fine grid by using the prolongation operator $\mathcal{P} : \mathcal{V}_{L-1} \to \mathcal{V}_L$ and is added to the approximation of the solution $u_L^{(\varsigma,1)}$. Applying this procedure recursively on several grids, one receives the following MMG method.

Algorithm 7.1 Monotone Multigrid method V-cycle ($\mathbf{MMG}_\ell$)

Let $u_\ell^{(\varsigma,0)}$ a given initial iterate.

(1) *A priori smoothing and projection:* $u_\ell^{(\varsigma,1)} := (\mathsf{Sm}(u_\ell^{(\varsigma,0)}, \widetilde{\psi}))^{\eta_1}$.

(2) *Coarse grid correction:*

$$\begin{aligned} d_\ell &:= (f_\ell - \mathcal{A}_\ell u_\ell^{(\varsigma,1)}), && \textit{(update of the defect)} \\ f_{\ell-1} &:= \mathcal{R} d_\ell, && \textit{(standard restriction)} \\ \mathcal{A}_{\ell-1} &:= \mathcal{R}\mathcal{A}_\ell\mathcal{P} \\ \widetilde{\psi}_\ell &:= \widetilde{\psi}_\ell - u_\ell^{(\varsigma,1)} && \textit{(update of the defect obstacle)} \\ \widetilde{\psi}_{\ell-1} &:= \tilde{r}(\widetilde{\psi}_\ell) && \textit{(monotone restriction)} \end{aligned}$$

If $\ell = \ell_{\min}$: Solve the linear complementary problem

$$\begin{aligned} (\mathsf{v} - \widetilde{\psi}_{\ell-1})^T(\mathcal{A}_{\ell-1}\mathsf{v} - f_{\ell-1}) &= 0 \\ \mathcal{A}_{\ell-1}\mathsf{v} &\geq f_{\ell-1} \\ \mathsf{v} &\geq \widetilde{\psi}_{\ell-1}, \end{aligned}$$

exactly and set $\mathsf{v}_{\ell-1} := \mathsf{v}$.

If $\ell > \ell_{\min}$: Perform one steps of $\mathbf{MMG}_{\ell-1}$ with initial value $u_{\ell-1}^{(\varsigma,0)} := 0$ and the solution $\mathsf{v}_{\ell-1}$.

(3) *Prolongation*: Set $u_\ell^{(\varsigma,2)} := u_\ell^{(\varsigma,1)} + \mathcal{P}\mathsf{v}_{\ell-1}$.

(4) *A posteriori smoothing and projection*: $u_\ell^{(\varsigma,3)} := (\mathsf{Sm}(u_\ell^{(\varsigma,2)}, \widetilde{\psi}))^{\eta_2}$.
Set $u_\ell^{(\varsigma+1,0)} := u_\ell^{(\varsigma,3)}$.

To ensure that the new iterate $u_\ell^{(\varsigma,2)}$ satisfies the constraint $u_\ell^{(\varsigma,2)} \geq \widetilde{\psi}_\ell$ on each level, it is important that the restriction operator $\tilde{r}$ is chosen such that

$$\begin{aligned} \mathcal{P}\widetilde{\psi}_{\ell-1} := \mathcal{P}\tilde{r}\left(\widetilde{\psi}_\ell - u_\ell^{(\varsigma,1)}\right) \geq \widetilde{\psi}_\ell - u_\ell^{(\varsigma,1)} \\ \Rightarrow u_\ell^{(\varsigma,2)} := u_\ell^{(\varsigma,1)} + \mathcal{P}\mathsf{v}_{\ell-1} \geq u_\ell^{(\varsigma,1)} + \mathcal{P}\widetilde{\psi}_{\ell-1} \geq \widetilde{\psi}_\ell. \end{aligned} \tag{7.16}$$

That means that $\widetilde{\psi}_{\ell-1}$ must be a *monotone upper coarse grid approximation* to the lower obstacle $\widetilde{\psi}_\ell - u_\ell^{(\varsigma,1)}$. The monotonicity ensures the global convergence of the scheme, see Section 7.2.3. This idea was also used in [GK09, Kor94, Man84, HK07].

7.2.2. Monotone coarse grid approximation for tensor-product splines

In this section, we first establish coarse grid approximations of one dimensional splines of order k with and without coinciding knots in the interior, which lead to suitable restriction operators $\tilde{r}$. Then the coarse grid approximation for multi-dimensional tensor product splines arises from the Kronecker product of the coarse grid approximations for one dimensional splines. For hat functions, monotone coarse grid approximation can

be found in [Kor94, Man84] and for higher order splines without coinciding knots in [HK07]. In this section we assume that d knot vectors $\Theta^{(j)}$ with $1 \leq j \leq d$ are given. Associated with the knot sequence $\Theta^{(j)}$ with uniform grid size h_j there is a mesh Θ, that is a partition of $\Omega \subset \mathbb{R}^d$ into d-dimensional open knot sequences $\Theta = \{\otimes_{j=1}^{d} \Theta^{(j)}\}$. We define by $\mathbb{S}_{k,\Theta^{(j)}}$ the space of splines regarding the knot sequence $\Theta^{(j)}$, then the d-dimensional tensor-product spline space of order k is given by

$$\mathbb{S}_{k,\Theta}^{d} := \otimes_{j=1}^{d} \mathbb{S}_{k,\Theta^{(j)}}. \tag{7.17}$$

To this end we consider tensor-product splines of order k with given expansion coefficients $c_{i_1,\ldots,i_d}$

$$\mathcal{S}(x) := \mathcal{S}(x_1,\ldots,x_d) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} c_{i_1,\ldots,i_d} N_{i_1,k,\Theta_1}(x_1) \cdots N_{i_d,k,\Theta_d}(x_d) \tag{7.18}$$

regarding the knot sequence $\Theta^{(j)}$ with grid size h_j for $j = 1,\ldots,d$. Let further the tensor-product spline of order k with given expansion coefficients $\tilde{c}_{i_1,\ldots,i_d}$ given by

$$\tilde{\mathcal{S}}(x) := \tilde{\mathcal{S}}(x_1,\ldots,x_d) = \sum_{i_1=1}^{\tilde{n}_1} \cdots \sum_{i_d=1}^{\tilde{n}_d} \tilde{c}_{i_1,\ldots,i_d} N_{i_1,k,\tilde{\Theta}_1}(x_1) \cdots N_{i_d,k,\tilde{\Theta}_d}(x_d) \tag{7.19}$$

regarding the knot sequence $\tilde{\Theta}_j$ with grid size $H_j := 2h_j$ for $j = 1,\ldots,d$. Then we define the tensor-product space of order k on a coarser grid $\tilde{\Theta}$ by

$$\mathbb{S}_{k,\widetilde{\Theta}}^{d} := \otimes_{j=1}^{d} \mathbb{S}_{k,\widetilde{\Theta}^{(j)}}. \tag{7.20}$$

First, we introduce the definition of upper monotone coarse grid approximations for d-dimensional tensor-product splines.

Definition 7.1. *(Monotone upper coarse grid approximation for tensor-product splines) Let $\mathbb{S}_{k,\Theta}^{d}$ be the space of d-dimensional splines of order k on the fine grid regarding the knot sequence Θ with grid size h_j. Let further $\mathbb{S}_{k,\widetilde{\Theta}}^{d}$ be the space of splines on the coarse grid regarding the knot sequence $\tilde{\Theta}$ with grid size H_j. We call $\tilde{\mathcal{S}}(x) := \tilde{\mathbf{c}}^T \tilde{\mathbf{N}}_k^d \in \mathbb{S}_{k,\widetilde{\Theta}}^{d}$ a monotone upper coarse grid approximation to the lower obstacle $\mathcal{S}(x) := \mathbf{c}^T \mathbf{N}_k^d \in \mathbb{S}_{k,\Theta}^{d}$, if $\tilde{\mathcal{S}}(x) \geq \mathcal{S}(x)$ holds for all $x \in \Omega \subset \mathbb{R}^d$.*

For a given spline obstacle function $\mathcal{S}(x)$ regarding the knot sequence Θ on a fine grid we provide a coarse grid approximation $\tilde{\mathcal{S}}(x)$ with respect to a coarser grid $\tilde{\Theta}$, which satisfies

i) $\tilde{\mathcal{S}}(x) \geq \mathcal{S}(x)$ for all $x \in \Omega \subset \mathbb{R}^d$;

ii) the number of arithmetic operations must be of order $\mathcal{O}(n^d)$;

iii) knots $\theta_i^{(j)} \in \Theta^{(j)}$ can be repeated at least $k-1$ times in the interior of $\Omega \subset \mathbb{R}^d$.

The first condition ensures the *robustness* and *monotonicity* of the MMG-method and the second condition the *optimal complexity*. The third condition is needed to *stabilize the computation of the Greeks* (see Section 4.2.2 or Section 5.2.2).
To derive monotone restriction operators according condition (iii) we will distinguish between the following two knot sequences. Therefore, let J^{rep} the index set for the dimensions, where the knots $\theta^{(j)}_{\wp} = \ldots = \theta^{(j)}_{\wp+k-2}$ in the extended sequence of knots $\Theta^{(j)}$ are repeated $k-1$ times

$$\theta^{(j)}_1 = \ldots = \theta^{(j)}_k < \theta^{(j)}_{k+1} < \ldots < \theta^{(j)}_{\wp} = \ldots = \theta^{(j)}_{\wp+k-2} < \ldots < \theta^{(j)}_{n_j} = \ldots = \theta^{(j)}_{n_j+k}$$
$$\text{with} \quad n_j := 2^{\ell_j} + 2k - 3 \text{ for } j \in J^{rep}. \tag{7.21}$$

Let further $\overline{J^{rep}} := \{1, \ldots, d\} \setminus J^{rep}$ denote the index set for the dimensions where the extended sequence of knots contains no internal repeated knots

$$\theta^{(j)}_1 = \ldots = \theta^{(j)}_k < \theta^{(j)}_{k+1} < \ldots < \ldots < \theta^{(j)}_n = \ldots = \theta^{(j)}_{n_j+k}$$
$$\text{with} \quad n_j := 2^{\ell_j} + k - 1 \text{ for } j \in \overline{J^{rep}}. \tag{7.22}$$

A graphical illustration of the knot series on a fine grid $\Theta^{(j)}$ and the corresponding knot series on a coarser grid $\tilde{\Theta}^{(j)}$ with $k-1$ repeating knots at $\theta^{(j)}_{2\tilde{\wp}}$, i.e. for $j \in J^{\text{rep}}$, can be found in Figure 7.1.

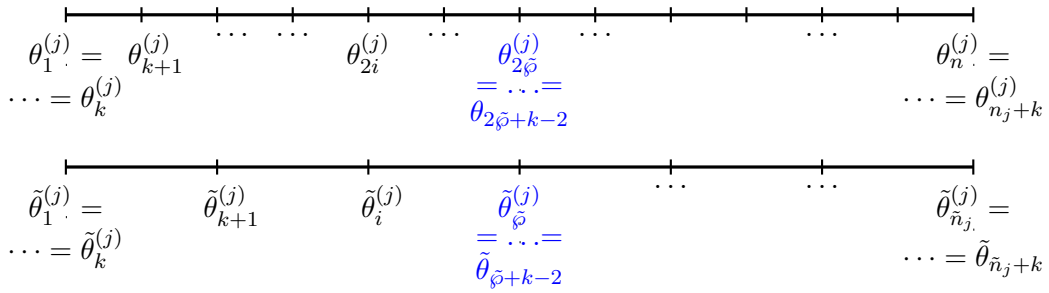

Figure 7.1.: Example for a knot series on a fine grid $\Theta^{(j)}$ (above) and the corresponding knot series on a coarser grid $\tilde{\Theta}^{(j)}$ (below) with $k-1$ repeating knots at $\theta^{(j)}_{2\tilde{\wp}}$, i.e. for $j \in J^{\text{rep}}$.

First, we consider coarse grid approximation for splines $N_{i,k,\Theta_j} \in \mathbb{S}_{k,\Theta^{(j)}}$ associated with the extended sequence of knots Θ_j for $j \in \overline{J^{rep}}$. More precisely we consider the directions where *no internal knots are repeated*. Therefore, one has to express the B-splines $N_{i_j,k,\tilde{\Theta}^{(j)}} \in \mathbb{S}_{k,\tilde{\Theta}^{(j)}}$ in the space $\mathbb{S}_{k,\Theta^{(j)}}$. It is well known (c.f. [Sch07]), that due to the convolution and scaling properties of B-splines, the following refinement equation holds

$$N_{i_j,k,\tilde{\Theta}^{(j)}}(x_j) = \sum_{q=0}^{k} a_q N_{2i_j-k+q,k,\Theta^{(j)}}(x_j) \quad \text{for } i_j = k, \ldots, \tilde{n}_j - k + 1,\ j \in \overline{J^{rep}} \tag{7.23}$$

with $\tilde{n}_j = 2^{\ell_j - 1} + k - 1$ B-splines on the level $\ell_j - 1$ and refinement coefficients

$$a_q := 2^{1-k} \binom{k}{q} \text{ for } q = 0, \dots, k. \tag{7.24}$$

The remaining refinement coefficients can be calculated via an exact interpolation as follows: We have

$$N_{i_j,k,\tilde{\Theta}^{(j)}}(x_j) = (\Phi^L_{(\bullet,i_j),k})^T \mathbf{N}^L_k \qquad \text{for } i_j = 1, \dots, k-1 \text{ and} \tag{7.25}$$

$$N_{i_j,k,\tilde{\Theta}^{(j)}}(x_j) = (\Phi^R_{(\bullet,i_j-(\tilde{n}_j-k+1)),k})^T \mathbf{N}^R_k \quad \text{for } i_j = \tilde{n}_j - k + 2, \dots, \tilde{n}_j,\ j \in \overline{J^{rep}} \tag{7.26}$$

where

$$\mathbf{N}^L_k := (N_{1,k,\Theta^{(j)}}(x_j), \dots, N_{2k-2,k,\Theta^{(j)}}(x_j))^T$$

and

$$\mathbf{N}^R_k := (N_{n_j-2k+3,k,\Theta^{(j)}}(x_j), \dots, N_{n_j,k,\Theta^{(j)}}(x_j))^T$$

and $\Phi^L_k \in \mathbb{R}^{(2k-2)\times(k-1)}$ includes the boundary adapted refinement coefficients on the left boundary and $\Phi^R_k \in \mathbb{R}^{(2k-2)\times(k-1)}$ the boundary adapted refinement coefficients on the right interval.
More precisely, the i_j'th row of the matrix Φ^L_k or Φ^R_k contains the refinement coefficients of the B-spline $N_{i_j,k,\tilde{\Theta}^{(j)}}$ for $i_j = 1, \dots, k-1$ or $i_j = \tilde{n}_j - k + 2, \dots, \tilde{n}_j$.

In matrix vector notation the refinement equation for B-splines can be expressed as follows

$$\widetilde{\mathbf{N}}^{(j)}_k = \boldsymbol{\mathcal{R}}^{(j)}_k \mathbf{N}^{(j)}_k \tag{7.27}$$

with $\mathbf{N}^{(j)}_k := (N_{1,k,\Theta_j}, \dots, N_{n_j,k,\Theta_j})^T$ and $\widetilde{\mathbf{N}}^{(j)}_k := (N_{1,k,\tilde{\Theta}_j}, \dots, N_{\tilde{n}_j,k,\tilde{\Theta}_j})^T$ and

$$\boldsymbol{\mathcal{R}}^{(j)}_k := \left(\begin{array}{c|cccccccc} (\Phi^L)^T & \mathbf{0} & & & & & & & \\ \hline \mathbf{0} & a_0 & a_1 & a_2 & \dots & a_k & & & \\ & & a_0 & \dots & a_{k-2} & a_{k-1} & a_k & & \\ & & & & & & \ddots & & \\ & & & & & & & a_0\ a_1 \dots a_k & \mathbf{0} \\ & & & & & & & \mathbf{0} & (\Phi^R)^T \end{array}\right),$$

$$\boldsymbol{\mathcal{R}}^{(j)}_k \in \mathbb{R}^{\tilde{n}_j \times n_j},\ k \geq 2, \quad \text{for } j \in \overline{J^{rep}}. \tag{7.28}$$

The above matrix is the restriction operator corresponding a knot series with *no internal repeated knots*, i.e. for directions where $j \in \overline{J^{rep}}$.

In the direction, where knots are *repeated* $k-1$ *times* at the knot $\theta_{\wp}^{(j)}$ B-splines $N_{i,k,\tilde{\Theta}^{(j)}} \in \mathbb{S}_{k,\tilde{\Theta}^{(j)}}$ can be expressed in the space $\mathbb{S}_{k,\Theta^{(j)}}$ for $j \in J^{rep}$ as follows:

$$N_{i_j,k,\tilde{\Theta}_j}(x_j) = \sum_{q=0}^{k} a_q N_{2i_j-k+q,k,\Theta_j}(x_j) \qquad \text{for } i_j = k,\dots,\tilde{\wp}-k \text{ and} \tag{7.29}$$

$$N_{i_j,k,\tilde{\Theta}_j}(x_j) = \sum_{q=0}^{k} a_q N_{2i_j-2k+2+q,k,\Theta_j}(x_j) \qquad \text{for } i_j = \tilde{\wp}+k-2,\dots,\tilde{n}_j-k+1 \tag{7.30}$$

with $\tilde{n}_j = 2^{\ell_j-1}+2k-3$ and refinement coefficients a_q as defined in (7.24).

The refinement equation with boundary adapted refinement coefficients for B-splines $N_{i_j,k,\tilde{\Theta}^{(j)}}(x)$ for $i_j = 1,\dots,k-1$ or $i_j = \tilde{n}_j-k+2,\dots,\tilde{n}_j$ can be derived analogously to (7.25) and (7.26).

Furthermore, the B-splines on the coarse grid regarding the *internal coinciding knots* can be represented on the fine grid by

$$N_{i_j,k,\tilde{\Theta}^{(j)}}(x_j) = (\Phi^I_{(\bullet,i_j-(\tilde{\wp}-k)),k})^T \mathbf{N}^I_k \qquad \text{for } i_j = \tilde{\wp}-k+1,\dots,\tilde{\wp}+k-3, \quad j \in J^{rep}, \tag{7.31}$$

where

$$\mathbf{N}^I_k := (N_{2\tilde{\wp}-3k+2,k,\Theta^{(j)}},\cdots,N_{2\tilde{\wp}+k-4,k,\Theta^{(j)}})^T$$

denotes the B-splines intersecting the B-splines on the coarser grid $N_{i_j,k,\tilde{\Theta}^{(j)}}(x_j)$ for $i_j = \tilde{\wp}-k+1,\dots,\tilde{\wp}+k-3$. Moreover, $\Phi^I \in \mathbb{R}^{(4k-5)\times(2k-3)}$ contains the adapted refinement coefficients for the internal repeating knots.

Now we can formulate the refinement equation for internal coinciding knots in matrix vector notation according to (7.27) with

$$\boldsymbol{\mathcal{R}}_k^{(j)} := \begin{pmatrix} (\Phi^L)^T & \mathbf{0} & & & & & \\ \mathbf{0}a_0\ a_1\ a_2 & \dots\ a_k & & & & & \\ & a_0 \dots a_{k-2} a_{k-1} a_k & & & & & \\ & & \ddots & & & & \\ & & & a_0\ a_1 & \dots\ a_k & & \\ & & & & (\Phi^I)^T & & \\ & & & & a_0\ a_1\ a_2 \dots\ a_k & & \\ & & & & & a_0 \dots a_{k-2} a_{k-1}\, a_k & \\ & & & & & & \ddots \\ & & & & & & a_0\ a_1 \dots\ a_k \mathbf{0} \\ & & & & & \mathbf{0} & (\Phi^R)^T \end{pmatrix},$$

$$\boldsymbol{\mathcal{R}}_k^{(j)} \in \mathbb{R}^{\tilde{n}_j \times n_j},\ k \geq 3,\ \text{for } j \in J^{rep}. \tag{7.32}$$

The above matrix is the restriction operator corresponding a knot series with $k-1$ *repeating internal knots*. Note, that in comparison to (7.28) the above matrix includes the coefficients matrix Φ^I due to the internal coinciding knots. It should be clear that for $k = 2$ no knots are repeated, thus one can use the same restriction operator as in (7.28). For piecewise continuous ($k = 2$), piecewise quadratic ($k = 3$) and piecewise cubic ($k = 4$) B-splines the corresponding coefficient matrices are given as follows

$$\Phi_2^L := \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}, \quad \Phi_2^R := \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}, \tag{7.33}$$

$$\Phi_3^L := \begin{pmatrix} 1 & \\ \frac{1}{2} & \frac{1}{2} \\ & \frac{3}{4} \\ & \frac{1}{4} \end{pmatrix}, \quad \Phi_3^R := \begin{pmatrix} \frac{1}{4} & \\ \frac{3}{4} & \\ & \frac{1}{2}\ \frac{1}{2} \\ & 1 \end{pmatrix}, \quad \Phi_3^I := \begin{pmatrix} \boxed{\Phi_3^R} & \\ & \boxed{1} \\ & \boxed{\Phi_3^L} \end{pmatrix}, \tag{7.34}$$

$$\Phi_4^L := \begin{pmatrix} 1 & & \\ \frac{1}{2} & \frac{1}{2} & \\ & \frac{3}{4} & \frac{1}{4} \\ & \frac{3}{16} & \frac{11}{16} \\ & & \frac{1}{2} \\ & & \frac{1}{8} \end{pmatrix}, \quad \Phi_4^R := \begin{pmatrix} \frac{1}{8} & & \\ \frac{1}{2} & & \\ \frac{11}{16} & \frac{3}{16} & \\ \frac{1}{4} & \frac{3}{4} & \\ & \frac{1}{2} & \frac{1}{2} \\ & & 1 \end{pmatrix}, \quad \Phi_4^I := \begin{pmatrix} \boxed{\Phi_4^R} & \\ & \boxed{1} \\ & \boxed{\Phi_4^L} \end{pmatrix}. \tag{7.35}$$

Theorem 7.2 provides monotone upper coarse grid approximation for tensor-product splines with $k-1$ coinciding knots in the interior by controlling the B-spline expansion coefficients. A proof similar to Theorem 7.2 and Corollary 7.4 where the boundary adapted refinement coefficients are ignored and where no repetition of the knots at $x = \theta_\wp$ are used can be found in [Hol04, HK07].

Theorem 7.2. *The spline $\tilde{\mathcal{S}}(x) \in \mathbb{S}^d_{k,\tilde{\Theta}}$ is an upper monotone coarse grid approximation to the lower obstacle $\mathcal{S}(x) \in \mathbb{S}^d_{k,\Theta}$ on the fine grid, if their B-spline expansion coefficients satisfy the linear inequality system*

$$\boldsymbol{\mathcal{P}}^d_k \tilde{\mathbf{c}} \geq \mathbf{c} \tag{7.36}$$

with tensor product matrix $\boldsymbol{\mathcal{P}}^d_k := \boldsymbol{\mathcal{P}}^{(1)}_k \otimes \ldots \otimes \boldsymbol{\mathcal{P}}^{(d)}_k$, where the matrices are defined as

$$\boldsymbol{\mathcal{P}}^{(j)}_k := \begin{pmatrix}
\boxed{\Psi^L_k} & & & & & & & & & & & & & & & & & & & & \\
\mathbf{0}\ a_{k-1} & a_{k-3} & \ldots & a_1 & & & & & & & & & & & & & & & & & \\
a_k & a_{k-2} & \ldots & a_2 & a_0 & & & & & & & & & & & & & & & & \\
 & a_{k-1} & \ldots & a_3 & a_1 & & & & & & & & & & & & & & & & \\
 & & \ddots & \vdots & \vdots & \ddots & & & & & & & & & & & & & & & \\
 & & & a_{k-1} & \ddots & \ddots & & & & & & & & & & & & & & & \\
 & & & a_k & & \ddots & & & & & & & & & & & & & & & \\
 & & & & \ddots & \ddots & \ddots & & & & & & & & & & & & & & \\
 & & & & & a_{k-1} & \ldots & a_1 & & & & & & & & & & & & & \\
 & & & & & a_k & \ldots & a_2 & a_0 & & & & & & & & & & & & \\
 & & & & & & \ddots & \vdots & \vdots & & & & & & & & & & & & \\
 & & & & & & & a_k & a_{k-2} & \ldots & a_0 & & & & & & & & & & \\
 & & & & & & & & a_{k-1} & \ldots & a_1 & & & & & & & & & & \\
 & & & & & & & & \boxed{\Psi^I_k} & & & & & & & & & & & & \\
 & & & & & & & & a_{k-1} & \ldots & a_1 & & & & & & & & & & \\
 & & & & & & & & a_k & \ldots & a_0 & & & & & & & & & & \\
 & & & & & & & & & \ddots & \vdots & \ddots & & & & & & & & & \\
 & & & & & & & & & & a_k & \ddots & \ddots & & & & & & & & \\
 & & & & & & & & & & & \ddots & \ddots & \ddots & & & & & & & \\
 & & & & & & & & & & & & a_{k-1} & \ldots & a_1 & & & & & & \\
 & & & & & & & & & & & & a_k & \ldots & a_2 & a_0 & & & & & \\
 & & & & & & & & & & & & & \ddots & \vdots & \vdots & & & & & \\
 & & & & & & & & & & & & & & a_k & a_{k-2} & \ldots & a_0 & & & \\
 & & & & & & & & & & & & & & & a_{k-1} & \ldots & a_1 & \mathbf{0} & & \\
 & & & & & & & & & & & & & & & & & & \boxed{\Psi^R_k} &
\end{pmatrix}, \tag{7.37}$$

$$\boldsymbol{\mathcal{P}}^{(j)}_k \in \mathbb{R}^{n_j \times \tilde{n}_j},\ k \geq 3,\ \textit{for } j \in J^{rep} \textit{ and}$$

$$\boldsymbol{\mathcal{P}}_k^{(j)} := \left(\begin{array}{c|cccccccccccccccccc} \multicolumn{1}{c|}{\multirow{2}{*}{Ψ_k^L}} & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & \\ \hline \mathbf{0} & a_{k-1} & a_{k-3} & \dots & a_1 & & & & & & & & & & & & & & \\ & a_k & a_{k-2} & \dots & a_2 & a_0 & & & & & & & & & & & & & \\ & & a_{k-1} & \dots & a_3 & a_1 & & & & & & & & & & & & & \\ & & a_k & \dots & a_4 & a_2 & a_0 & & & & & & & & & & & & \\ & & & \ddots & \vdots & \vdots & \vdots & & & & & & & & & & & & \\ & & & & a_{k-1} & a_{k-3} & \vdots & \ddots & & & & & & & & & & & \\ & & & & a_k & a_{k-2} & \vdots & \ddots & & & & & & & & & & & \\ & & & & & a_{k-1} & a_{k-3} & \ddots & & & & & & & & & & & \\ & & & & & a_k & a_{k-2} & \ddots & & & & & & & & & & & \\ & & & & & & a_{k-1} & \ddots & & & & & & & & & & & \\ & & & & & & a_k & & \ddots & & & & & & & & & & \\ & & & & & & & & \ddots & & & & & & & & & & \\ & & & & & & & & & \ddots & & & & & & & & & \\ & & & & & & & & & & a_{k-1} & \dots & a_1 & & & & & & \\ & & & & & & & & & & a_k & \dots & a_2 & a_0 & & & & & \\ & & & & & & & & & & & \ddots & \vdots & \vdots & & & & & \\ & & & & & & & & & & & & a_k & a_{k-2} & \dots & a_0 & & & \\ & & & & & & & & & & & & & a_{k-1} & \dots & a_1 & \mathbf{0} & & \\ \hline & & & & & & & & & & & & & & & & & \multicolumn{1}{c|}{\multirow{2}{*}{Ψ_k^R}} \\ & & & & & & & & & & & & & & & & & & \end{array}\right)$$

$$\boldsymbol{\mathcal{P}}_k^{(j)} \in \mathbb{R}^{n_j \times \tilde{n}_j},\ k \geq 2,\ \textit{for } j \in \overline{J^{rep}} \tag{7.38}$$

with refinement coefficients a_q as defined in (7.24), zero vector $\mathbf{0} \in \mathbb{R}^{k-1}$ and $\Psi_k^L, \Psi_k^R \in \mathbb{R}^{2k-2\times(3k-2)/2}$, $\Psi_k^I \in \mathbb{R}^{4k-5\times 3k-3}$ defined as

$$\Psi_k^L := \left(\begin{array}{c|ccc} \Phi_k^L & \mathbf{0} & & \\ & a_0 & & \\ & a_1 & & \\ & a_2 & a_0 & \\ & \vdots & & \ddots \\ & a_{k-2} & a_{k-4} & \dots\ a_0 \end{array}\right),\quad \Psi_k^R := \left(\begin{array}{ccccc|c} a_k & a_{k-2} & a_{k-4} & \dots & a_2 & \\ & a_{k-1} & a_{k-3} & \dots & a_3 & \\ & a_k & a_{k-2} & \dots & a_4 & \\ & & & \ddots & \vdots & \Phi_k^R \\ & & & & a_{k-1} & \\ & & & & a_k & \\ & & & & \mathbf{0} & \end{array}\right),$$

$$\textit{and } \Psi_k^I := \left(\begin{array}{ccc} \boxed{\Psi_k^R} & & \\ & \boxed{1} & \\ & & \boxed{\Psi_k^L} \end{array}\right). \tag{7.39}$$

The matrix in (7.37) corresponds to the prolongation operator when $k-1$ internal knots are repeated ($j \in \overline{J^{rep}}$) and the matrix in (7.38) is regarding a knot series with no internal coinciding knots.

Proof. The proof relies on the convolution and scaling properties and on the nonnegativity of B-splines. We consider the case k even. To shorten the notation we define $N_{i_j,k} := N_{i_j,k,\Theta_j}$ and $\tilde{N}_{i_j,k} := N_{i_j,k,\tilde{\Theta}_j}$. First, we express the spline $\tilde{\mathcal{S}}(x_j) \in \mathbb{S}_{k,\tilde{\Theta}^{(j)}}$ for $j \in J^{rep}$ by splines on the fine grid

$$\begin{aligned}\tilde{\mathcal{S}}(x_j) = \sum_{i_j=1}^{\tilde{n}_j} \tilde{c}_{i_j}\tilde{N}_{i_j,k} &= \sum_{i_j=1}^{k-1}\left(\tilde{c}_{i_j}(\Phi^L_{(\bullet,i_j),k})^T\mathbf{N}^L_k\right) + \sum_{i_j=k}^{\tilde{\wp}-k}\left(\tilde{c}_{i_j}\sum_{q=0}^{k} a_q N_{2i_j-k+q,k}(x_j)\right)\\
&+ \sum_{i_j=\tilde{\wp}-k+1}^{\tilde{\wp}+k-3}\left(\tilde{c}_{i_j}(\Phi^I_{(\bullet,i_j-(\tilde{\wp}-k)),k})^T\mathbf{N}^I_k\right) + \sum_{i_j=\tilde{\wp}+k-2}^{\tilde{n}-k+1}\left(\tilde{c}_{i_j}\sum_{q=0}^{k} a_q N_{2i_j-2k+2+q,k}(x_j)\right)\\
&+ \sum_{i_j=\tilde{n}_j-k+2}^{\tilde{n}_j}\left(\tilde{c}_{i_j}(\Phi^R_{(\bullet,i_j-(\tilde{n}_j-k+1)),k})^T\mathbf{N}^R_k\right).\end{aligned}$$

In the next step, we sort the above equation according to the basis functions $N_{i_j,k}$. Therefore, we have defined the matrices for the boundary adapted refinement coefficients in (7.39). The i_j'th row from Ψ^L_k contains the sorted refinement coefficients according the basis function $N_{i_j,k}$. Then we obtain:

$$\begin{aligned}\tilde{\mathcal{S}}(x_j) = &\sum_{i_j=1}^{2k-2}\Psi^L_k(i_j,\bullet)\tilde{\mathbf{c}}^{L^T}_k N_{i_j,k}(x_j)\\
&+ \sum_{i_j=2k-1}^{2\wp-3k+1}\left(a_{k-1}\tilde{c}_{(i_j+1)/2} + a_{k-3}\tilde{c}_{(i_j+3)/2} + \ldots + a_1\tilde{c}_{(i_j+k-1)/2}\right)N_{i_j,k}(x_j)\\
&+ \sum_{i_j=2k}^{2\tilde{\wp}-3k}\left(a_k\tilde{c}_{i_j/2} + a_{k-2}\tilde{c}_{(i_j+2)/2} + \ldots + a_0\tilde{c}_{(i_j+k)/2}\right)N_{i_j,k}(x_j)\\
&+ \sum_{i_j=2\tilde{\wp}-3k+2}^{2\tilde{\wp}+k-4}\Psi^I_k(i_j-(2\tilde{\wp}-3k+1),\bullet)\tilde{\mathbf{c}}^{I^T}_k N_{i_j,k}(x_j)\\
&+ \sum_{i_j=2\tilde{\wp}+k-3}^{n_j-2k+1}\left(a_{k-1}\tilde{c}_{(i_j+k-1)/2} + a_{k-3}\tilde{c}_{(i_j+k-3)/2} + \ldots + a_1\tilde{c}_{(i_j+2k-3)/2}\right)N_{i_j,k}(x_j)\\
&+ \sum_{i_j=2\tilde{\wp}+k-2}^{n_j-2k+2}\left(a_k\tilde{c}_{(i_j+k-2)/2} + a_{k-2}\tilde{c}_{(i_j+k-2)/2} + \ldots + a_0\tilde{c}_{(i_j+2k-2)/2}\right)N_{i_j,k}(x_j)\\
&+ \sum_{i_j=n-2k+3}^{n}\Psi^R(i_j-(n-2k+2),\bullet)\tilde{\mathbf{c}}^{R^T}_k N_{i_j,k}(x_j).\end{aligned}$$

$\tilde{\mathcal{S}}(x_j)$ is an upper monotone coarse grid approximation to the lower obstacle $\mathcal{S}(x_j)$, if

$$\tilde{\mathcal{S}}(x_j) - \mathcal{S}(x_j) = \sum_{i_j=1}^{n_j} d^{(j)}_{i_j}N_{i_j,k}(x_j) := \sum_{i_j=1}^{n_j}\left(\boldsymbol{\mathcal{P}}^{(j)}_k\tilde{\mathbf{c}}^{(j)} - \mathbf{c}^{(j)}\right)_{i_j}N_{i_j,k}(x_j) \geq 0 \qquad (7.40)$$

is satisfied.

Due to the non-negativity of B-splines, we only have to require $d_{i_j} \geq 0$, where d_{i_j} for $j \in J^{rep}$ and k even is defined as

$$d_{i_j}^{(j)} :=$$

$$\begin{cases} \Psi_k^L(i_j, \bullet)\tilde{\mathbf{c}}_k^{L^T} - c_{i_j} & 1 \leq i_j \leq 2k-2, \\ a_{k-1}\tilde{c}_{(i_j+1)/2} + \ldots + a_1\tilde{c}_{(i_j+k-1)/2} - c_{i_j} & 2k-1 \leq i_j \leq 2\tilde{\wp} - 3k + 1,\ i_j \text{ odd}, \\ a_k\tilde{c}_{i_j/2} + \ldots + a_0\tilde{c}_{(i_j+k)/2} - c_{i_j} & 2k \leq i_j \leq 2\tilde{\wp} - 3k,\ i_j \text{ even}, \\ \Psi_k^I(i_j - (2\tilde{\wp} - 3k + 1), \bullet)\tilde{\mathbf{c}}_k^{I^T} - c_{i_j} & 2\tilde{\wp} - 3k + 2 \leq i_j \leq 2\tilde{\wp} + k - 4, \\ a_{k-1}\tilde{c}_{(i_j+k-1)/2} + \ldots + a_1\tilde{c}_{(i_j+2k-3)/2} - c_{i_j} & 2\tilde{\wp} + k - 3 \leq i_j \leq n - 2k + 1, i_j \text{ odd} \\ a_k\tilde{c}_{(i_j+k-2)/2} + \ldots + a_0\tilde{c}_{(i_j+2k-2)/2} - c_{i_j} & 2\tilde{\wp} + k - 2 \leq i_j \leq n - 2k + 2, i_j \text{ even} \\ \Psi^R(i_j - (n-2k+2), \bullet)\tilde{\mathbf{c}}_k^{R^T} - c_{i_j} & n - 2k + 3 \leq i_j \leq n, \end{cases} \tag{7.41}$$

which leads to the inequality system $\boldsymbol{\mathcal{P}}_k^{(j)}\tilde{\mathbf{c}}^{(j)} \geq \mathbf{c}^{(j)}$. An analogous procedure provides the result for the splines regarding the extended sequence of knots without repeated knots in the interior with d_{i_j} for $j \in \overline{J^{rep}}$ and k even defined as

$$d_{i_j}^{(j)} := \begin{cases} \Psi_k^L(i_j, \bullet)\tilde{\mathbf{c}}_k^{L^T} - c_{i_j} & 1 \leq i_j \leq 2k-2, \\ a_{k-1}\tilde{c}_{(i_j+1)/2} + \ldots + a_1\tilde{c}_{(i_j+k-1)/2} - c_{i_j} & 2k-1 \leq i_j \leq n - 2k + 1,\ i_j \text{ odd}, \\ a_k\tilde{c}_{i_j/2} + \ldots + a_0\tilde{c}_{(i_j+k)/2} - c_{i_j} & 2k \leq i_j \leq n - 2k + 2,\ i_j \text{ even}, \\ \Psi^R(i_j - (n-2k+2), \bullet)\tilde{\mathbf{c}}_k^{R^T} - c_{i_j} & n - 2k + 3 \leq i_j \leq n. \end{cases} \tag{7.42}$$

By defining $\mathbf{c} := \mathbf{c}^{(1)} \otimes \ldots \otimes \mathbf{c}^{(d)}$ and $\tilde{\mathbf{c}} := \tilde{\mathbf{c}}^{(1)} \otimes \ldots \otimes \tilde{\mathbf{c}}^{(d)}$ we obtain the inequality system (7.36) by

$$\tilde{\mathcal{S}}(x) - \mathcal{S}(x) = \sum_{\mathbf{i}:=(i_1,\ldots,i_d)} \left(\boldsymbol{\mathcal{P}}_k^{(d)}\tilde{\mathbf{c}} - \mathbf{c}\right)_{\mathbf{i}} N_{\mathbf{i},k}(x) \geq 0. \tag{7.43}$$

An analogous procedure provides the result for k odd. □

For a better understanding we give an example for Theorem 7.2 in the case of cubic $k = 4$ B-splines with $k-1$ coinciding knots at $x^{(j)} = \theta_{\wp}^{(j)}$. An example without internal coinciding knots can be found in [Bos15].

Example 7.3. *(Monotone upper coarse grid approximation for $k = 4$)*
Exact interpolation as in (7.25),(7.26),(7.31) and the refinement equations in (7.29),(7.30) leads to

$$\tilde{\mathcal{S}}(x_j) = \sum_{i_j=1}^{\tilde{n}_j} \tilde{c}_{i_j}\tilde{N}_{i_j,4} = \tilde{c}_1\left(1N_{1,4} + \frac{1}{2}N_{2,4}\right) + \tilde{c}_2\left(\frac{1}{2}N_{2,4} + \frac{3}{4}N_{3,4} + \frac{3}{16}N_{4,4}\right)$$

$$
\begin{aligned}
&+ \tilde{c}_3 \left(\frac{1}{4} N_{3,4} + \frac{11}{16} N_{4,4} + \frac{1}{2} N_{5,4} + \frac{1}{8} N_{6,4} \right) \\
&+ \sum_{i_j=4}^{\tilde{\varphi}-4} \tilde{c}_{i_j} \left(a_0 N_{2i_j-4,4} + a_1 N_{2i_j-3,4} + a_2 N_{2i_j-2,4} + a_3 N_{2i_j-1,4} + a_4 N_{2i_j,4} \right) \\
&+ \tilde{c}_{\tilde{\varphi}-3} \left(\frac{1}{8} N_{2\tilde{\varphi}-10,4} + \frac{1}{2} N_{2\tilde{\varphi}-9,4} + \frac{11}{16} N_{2\tilde{\varphi}-8,4} + \frac{1}{4} N_{2\tilde{\varphi}-7,4} \right) \\
&+ \tilde{c}_{\tilde{\varphi}-2} \left(\frac{3}{16} N_{2\tilde{\varphi}-8,4} + \frac{3}{4} N_{2\tilde{\varphi}-7,4} + \frac{1}{2} N_{2\tilde{\varphi}-6,4} \right) \\
&+ \tilde{c}_{\tilde{\varphi}-1} \left(\frac{1}{2} N_{2\tilde{\varphi}-6,4} + 1 N_{2\tilde{\varphi}-5,4} + \frac{1}{2} N_{2\tilde{\varphi}-4,4} \right) \\
&+ \tilde{c}_{\tilde{\varphi}} \left(\frac{1}{2} N_{2\tilde{\varphi}-4,4} + \frac{3}{4} N_{2\tilde{\varphi}-3,4} + \frac{3}{16} N_{2\tilde{\varphi}-2,4} \right) \\
&+ \tilde{c}_{\tilde{\varphi}+1} \left(\frac{1}{4} N_{2\tilde{\varphi}-3,4} + \frac{11}{16} N_{2\tilde{\varphi}-2,4} + \frac{1}{2} N_{2\tilde{\varphi}-1,4} + \frac{1}{8} N_{2\tilde{\varphi},4} \right) \\
&+ \sum_{i_j=\tilde{\varphi}+2}^{\tilde{n}_j-3} \tilde{c}_{i_j} \left(a_0 N_{2i_j-6,4} + a_1 N_{2i_j-5,4} + a_2 N_{2i_j-4,4} + a_3 N_{2i_j-3,4} + a_4 N_{2i_j-2,4} \right) \\
&+ \tilde{c}_{\tilde{n}_j-2} \left(\frac{1}{8} N_{n_j-5,4} + \frac{1}{2} N_{n_j-4,4} + \frac{11}{16} N_{n_j-3,4} + \frac{1}{4} N_{n_j-2,4} \right) \\
&+ \tilde{c}_{\tilde{n}_j-1} \left(\frac{3}{16} N_{n_j-3,4} + \frac{3}{4} N_{n_j-2,4} + \frac{1}{2} N_{n_j-1,4} \right) + \tilde{c}_{\tilde{n}_j} \left(\frac{1}{2} N_{n_j-1,4} + 1 N_{n_j,4} \right).
\end{aligned}
$$

Sorting according $N_{i_j,4}$ yields

$$
\begin{aligned}
\tilde{\mathcal{S}}(x_j) &= \tilde{c}_1 N_{1,4} + \left(\frac{1}{2}\tilde{c}_1 + \frac{1}{2}\tilde{c}_2 \right) N_{2,4} + \left(\frac{3}{4}\tilde{c}_2 + \frac{1}{4}\tilde{c}_3 \right) N_{3,4} + \left(\frac{3}{16}\tilde{c}_2 + \frac{11}{16}\tilde{c}_3 + a_0\tilde{c}_4 \right) N_{4,4} \\
&+ \left(\frac{1}{2}\tilde{c}_3 + a_1\tilde{c}_4 \right) N_{5,4} + \left(\frac{1}{8}\tilde{c}_3 + a_2\tilde{c}_4 + a_0\tilde{c}_5 \right) N_{6,4} + \sum_{\substack{i_j=7 \\ i_j\, odd}}^{2\tilde{\varphi}-11} \left(a_3\tilde{c}_{(i_j+1)/2} + a_1\tilde{c}_{(i_j+3)/2} \right) N_{i_j,4} \\
&+ \sum_{\substack{i_j=8 \\ i_j\, even}}^{2\tilde{\varphi}-12} \left(a_4\tilde{c}_{i_j/2} + a_2\tilde{c}_{(i_j+2)/2} + a_0\tilde{c}_{(i_j+4)/2} \right) N_{i_j,4} + \left(a_4\tilde{c}_{\tilde{\varphi}-5} + a_2\tilde{c}_{\tilde{\varphi}-4} + \frac{1}{8}\tilde{c}_{\tilde{\varphi}-3} \right) N_{2\tilde{\varphi}-10,4} \\
&+ \left(a_3\tilde{c}_{\tilde{\varphi}-4} + \frac{1}{2}\tilde{c}_{\tilde{\varphi}-3} \right) N_{2\tilde{\varphi}-9,4} + \left(a_4\tilde{c}_{\tilde{\varphi}-4} + \frac{11}{16}\tilde{c}_{\tilde{\varphi}-3} + \frac{3}{16}\tilde{c}_{\tilde{\varphi}-2} \right) N_{2\tilde{\varphi}-8,4} \\
&+ \left(\frac{1}{4}\tilde{c}_{\tilde{\varphi}-3} + \frac{3}{4}\tilde{c}_{\tilde{\varphi}-2} \right) N_{2\tilde{\varphi}-7,4} + \left(\frac{1}{2}\tilde{c}_{\tilde{\varphi}-2} + \frac{1}{2}\tilde{c}_{\tilde{\varphi}-1} \right) N_{2\tilde{\varphi}-6,4} + 1\tilde{c}_{\tilde{\varphi}-1} N_{2\tilde{\varphi}-5,4} \\
&+ \left(\frac{1}{2}\tilde{c}_{\tilde{\varphi}-1} + \frac{1}{2}\tilde{c}_{\tilde{\varphi}} \right) N_{2\tilde{\varphi}-4,4} + \left(\frac{3}{4}\tilde{c}_{\tilde{\varphi}} + \frac{1}{4}\tilde{c}_{\tilde{\varphi}+1} \right) N_{2\tilde{\varphi}-3,4} \\
&+ \left(\frac{3}{16}\tilde{c}_{\tilde{\varphi}} + \frac{11}{16}\tilde{c}_{\tilde{\varphi}+1} + a_0\tilde{c}_{\tilde{\varphi}+2} \right) N_{2\tilde{\varphi}-2,4} + \left(\frac{1}{2}\tilde{c}_{\tilde{\varphi}+1} + a_1\tilde{c}_{\tilde{\varphi}+2} \right) N_{2\tilde{\varphi}-1,4}
\end{aligned}
$$

$$
\begin{aligned}
&+\left(\frac{1}{8}\tilde{c}_{\tilde{\wp}+1}+a_2\tilde{c}_{\tilde{\wp}+2}+a_0\tilde{c}_{\tilde{\wp}+3}\right)N_{2\tilde{\wp},4}+\sum_{\substack{i_j=2\tilde{\wp}+1\\ i_j\,odd}}^{n_j-7}\left(a_3\tilde{c}_{(i_j+3)/2}+a_1\tilde{c}_{(i_j+5)/2}\right)N_{i_j,4}\\
&+\sum_{\substack{i_j=2\tilde{\wp}+2\\ i_j\,even}}^{n_j-6}\left(a_4\tilde{c}_{(i_j+2)/2}+a_2\tilde{c}_{(i_j+4)/2}+a_0\tilde{c}_{(i_j+6)/2}\right)N_{i_j,4}\\
&+\left(\frac{1}{8}\tilde{c}_{\tilde{n}-2}+a_2\tilde{c}_{\tilde{n}-3}+a_0\tilde{c}_{\tilde{n}-4}\right)N_{n-5,4}+\left(\frac{1}{2}\tilde{c}_{\tilde{n}-2}+a_1\tilde{c}_{\tilde{n}-3}\right)N_{n-4,4}\\
&+\left(\frac{3}{16}\tilde{c}_{\tilde{n}-1}+\frac{11}{16}\tilde{c}_{\tilde{n}-2}+a_0\tilde{c}_{\tilde{n}-3}\right)N_{n-3,4}+\left(\frac{3}{4}\tilde{c}_{\tilde{n}-1}+\frac{1}{4}\tilde{c}_{\tilde{n}-2}\right)N_{n-2,4}\\
&+\left(\frac{1}{2}\tilde{c}_{\tilde{n}}+\frac{1}{2}\tilde{c}_{\tilde{n}-1}\right)N_{n-1,4}+\tilde{c}_{\tilde{n}}N_{n,4}.
\end{aligned}
$$

Then $\tilde{\mathcal{S}}(x_j)$ is an upper monotone coarse grid approximation to the lower obstacle $\mathcal{S}(x_j)$, if

$$\tilde{\mathcal{S}}(x_j)-\mathcal{S}(x_j)=\sum_{i_j=1}^{n_j}\left(\boldsymbol{\mathcal{P}}_4^{(j)}\tilde{\mathbf{c}}^{(j)}-\mathbf{c}^{(j)}\right)_{i_j}N_{i_j,4}(x_j)\geq 0 \tag{7.44}$$

is satisfied. Due to the non-negativity of B-splines, we obtain the inequality system $\boldsymbol{\mathcal{P}}_4^{(j)}\tilde{\mathbf{c}}^{(j)}\geq\mathbf{c}^{(j)}$.

Finally, we can immediately derive the quasi-optimal monotone upper coarse grid approximation for tensor product B-splines with coinciding knots.

Corollary 7.4. *(Quasi-optimal monotone upper coarse grid approximation) The Spline $L_k=\tilde{\mathbf{r}}^T\tilde{\boldsymbol{N}}_k\in\mathbb{S}^d_{k,\tilde{\Theta}}$ with expansion coefficients*

$$\tilde{r}_{\mathbf{i}}(\mathbf{c}):=\max\left\{c_{\mathbf{l}}:\mathbf{l}:=(\mathfrak{q}_{i_1},\ldots,\mathfrak{q}_{i_d}),\mathfrak{q}_{i_j}=\{2i_j-k,\ldots,2i_j\}\right\} \tag{7.45}$$

is an upper monotone coarse grid approximation to the lower obstacle $S:=\mathbf{c}^T\mathbf{N}_k\in\mathbb{S}^d_{k,\Theta}$.

Proof. The proof follows directly from Theorem 7.2. Let us define $\tilde{\mathbf{r}}:=\tilde{\mathbf{r}}^{(1)}\otimes\ldots\otimes\tilde{\mathbf{r}}^{(d)}$ and $\mathbf{c}:=\mathbf{c}^{(1)}\otimes\ldots\otimes\mathbf{c}^{(d)}$. As all row sums of $\boldsymbol{\mathcal{P}}_k^{(j)}$ are equal to one, the vector $\tilde{\mathbf{r}}^{(j)}$ containing the expansion coefficients from (7.45) satisfies $\boldsymbol{\mathcal{P}}_k^{(j)}\tilde{\mathbf{r}}^{(j)}\geq\mathbf{c}^{(j)}$. Due to the property of the Kronecker product one has

$$\left(\boldsymbol{\mathcal{P}}_k^{(1)}\otimes\ldots\otimes\boldsymbol{\mathcal{P}}_k^{(d)}\right)\left(\tilde{\mathbf{r}}^{(1)}\otimes\ldots\otimes\tilde{\mathbf{r}}^{(d)}\right)=\boldsymbol{\mathcal{P}}_k^{(1)}\tilde{\mathbf{r}}^{(1)}\otimes\ldots\otimes\boldsymbol{\mathcal{P}}_k^{(d)}\tilde{\mathbf{r}}^{(d)}\geq\mathbf{c}^{(1)}\otimes\ldots\otimes\mathbf{c}^{(d)}=:\mathbf{c}. \tag{7.46}$$

Thus the vector $\tilde{\mathbf{r}}$ containing the expansion coefficient from (7.45) satisfies the inequality system $\boldsymbol{\mathcal{P}}_k^d\tilde{\mathbf{r}}\geq\mathbf{c}$. □

For determining an upper monotone coarse grid approximation, a minimum has to be found n^d times. This can be solved by using a standard sorting algorithm with com-

plexity $\mathcal{O}(k \cdot (n^d)) = \mathcal{O}(n^d)$, such that the optimal complexity of the MMG method is preserved.

7.2.3. Convergence theory

In the next section we discuss the convergence theory for the MMG method. The theory for linear finite element has been established by [Kor94]. A detailed review article on monotone multigrid methods with linear finite elements can also be found in [GK09]. Moreover, the author in [Hol04] has transferred the theory from [Kor94] to B-splines. In particular, it was proven that the MMG method is asymptotically convergent and reduces to a linear subspace correction method provided quasi-optimal and monotone restriction operators are used. As mentioned earlier, in comparison to [Hol04] the quasi-optimal and monotone restriction operators, derived in this thesis, are based on tensor product B-splines with possible internal coinciding knots. Since the theory is based on the monotonicity and quasi-optimality of the restriction operators as provided in Corollary 7.4, the proofs are analogous to [Hol04, Kor94]. Moreover, the author in [Kor94] has provided nearly optimal a posteriori estimates for the asymptotic convergence rates in the case of linear finite element basis functions. A corresponding result for tensor product B-splines is still outstanding. The results presented in the next section are very close to [Kor94].
Our aim is to analyze the MMG method for the discrete variational inequality to find the solution $u_L \in \mathcal{K}_L^0$ such that

$$a(u_L, \varphi_L - u_L) \geq f(\varphi_L - u_L) \quad \text{for all } \varphi_L \in \mathcal{K}_L^0 \tag{7.47}$$

where the bilinear form is assumed to be symmetric and coercive in $H_0^1(\Omega)$ satisfying

$$M_1 \leq a(u,u) \leq M_2 \quad \text{for all } u \in H_0^1(\Omega). \tag{7.48}$$

Furthermore, the considerations are restricted to a polygonal domain $\Omega \subset \mathbb{R}^2$. Since $a(\cdot,\cdot)$ is symmetric the variational inequality is equivalent to the following optimization problem, that is to find $u_L \in \mathcal{K}_L^0$ such that

$$\mathcal{J}(u_L) \leq \mathcal{J}(\varphi_L) \quad \text{for all } \varphi_L \in \mathcal{K}_L^0 \tag{7.49}$$

and the linear functional

$$\mathcal{J}(\varphi_L) := \frac{1}{2} a(\varphi_L, \varphi_L) - f(\varphi_L). \tag{7.50}$$

Let us further define the active set

$$\mathcal{K}_L^{\bullet}(\varphi) := \{\mathbf{i} := (i_1, i_2), i_j \in \{2, \ldots, n_L - 1\} : \varphi_{\mathbf{i}}^L = \tilde{\psi}_{\mathbf{i}}^L\} \tag{7.51}$$

on Level L.

The remaining set

$$\mathcal{K}_L^\circ(\varphi) := \{\mathbf{i} := (i_1, i_2), i_j \in \{2, \dots, n_L - 1\} : \varphi_{\mathbf{i}}^L > \tilde{\psi}_{\mathbf{i}}^L\} \tag{7.52}$$

is called inactive set. Moreover, we say that the bilinear form satisfies the *strict complementarity condition* if

$$a(u_L, N_{\mathbf{i},k}) > f(N_{\mathbf{i},k}) \text{ for all } \mathbf{i} \in \mathcal{K}_L^\bullet(u_L). \tag{7.53}$$

Once the active set $\mathcal{K}_L^\bullet$ of the solution u_L is identified the discrete variational inequality reduces to the variational equation

$$a(u_L, \varphi_L) = f(\varphi_L) \quad \text{for all } \varphi_L \in \mathcal{V}_L^\circ \tag{7.54}$$

with the reduced subspace $\mathcal{V}_L^\circ \subset \mathcal{V}_L$ defined as

$$\mathcal{V}_L^\circ := \{\varphi_L \in \mathcal{V}_L | \varphi_{\mathbf{i}} = 0 \text{ for } \mathbf{i} \in \mathcal{K}_L^\bullet\}. \tag{7.55}$$

In order to analyze the convergence of the MMG method we first introduce the *extended relaxation method*. This general approach includes the classical multigrid method to solve linear variational equations as well as various extension to solve nonlinear problems. As pointed out by [Kor94] this approach can also be considered as nonlinear subspace correction method.
As mentioned earlier the projected Gauss-Seidel method suffers from unsatisfactory convergence rates when the the grid size is decreased. In order to accelerate basic iterative schemes the smoothing property of the projected Gauss-Seidel is exploited together with a correction step by successive minimization of the energy $\mathcal{J}$ on $\mathcal{K}_L$. To be more precise the basic iterative scheme with high frequent search direction Λ_L is extended by additional low frequent search directions M_c^ς with basis function on a coarser grid.

For simplicity we denote the set of tensor product B-spline basis function on the fine grid Θ_L by $\{N_{\mathbf{i},\Theta_L}\}_{i \in \mathcal{I}_L}$ with cardinality $\#\mathcal{I}_L = (n_L - 2)^2 := N_L^2$. Let

$$\Lambda_L := \{\mu_1, \dots, \mu_{N_L^2}\} := \{N_{\mathbf{i},\Theta_L}\}_{i \in \mathcal{I}}$$

denote the set of basis function of $\mathcal{V}_L$. Moreover, we denote the extended set of search direction by

$$M^\varsigma := \Lambda_L \cup M_c^\varsigma \tag{7.56}$$

where $M_c^\varsigma := \{\mu_{N_L^2+1}^\varsigma, \dots, \mu_{m_\varsigma}^\varsigma\}$ denotes the coarse grid function, which can vary in each iteration step ς.
The extended relaxation method is to find a numerical solution of (7.49) by the following iterates: For a given iterate w_0^ς the iterates $w^{\varsigma,j} = w^{\varsigma,j-1} + v^{*,\eta,j}$ for $j = 1, \dots, m_\varsigma$ are successively computed. The corrections $v^{*,\varsigma,j}$ are the unique solutions of the following local subproblems: Find $v^{*,\varsigma,j} \in D^{*,\varsigma,j}$ such that

$$\mathcal{J}(w^{\varsigma,j-1} + v^{*,\varsigma,j}) \leq \mathcal{J}(w^{\varsigma,j-1} + v) \quad \text{for all } v \in D^{*,\varsigma,j}, \tag{7.57}$$

where the closed convex subsets are defined as

$$D^{*,\varsigma,j} := \{v \in \operatorname{span}\{\mu_j^\varsigma\} : v_{\mathbf{i}} \geq \tilde{\psi}_{\mathbf{i}} - w_{\mathbf{i}}^{\varsigma,j-1} \text{ for all } \mathbf{i} \in \{1,\dots,N_L^2\}\}. \tag{7.58}$$

To ensure that the local corrections $v^{*,\varsigma,j} \in \operatorname{span}\{\mu_j^\varsigma\}$ for $j = N_L^2+1,\dots,m_\varsigma$ on the coarse grid are in the subsets $D^{*,\varsigma,j}$ the functions usually needs to interpolate on a finer grid. This destroys the optimal complexity of the multigrid method. Thus, the correction $v^{*,\varsigma,j} \in \operatorname{span}\{\mu_j^\varsigma\}$ is replaced by the following approximation

$$D^{\varsigma,j} := \{v \in \operatorname{span}\{\mu_j^\varsigma\} : v_{\mathbf{i}} \geq \tilde{\psi}_{\mathbf{i}}^{\varsigma,j} \text{ for all } \mathbf{i} \in \{1,\dots,N_L^2\}\}. \tag{7.59}$$

The corresponding approach is also called *approximated extended relaxation method*. The local approximated obstacle functions $\tilde{\psi}_{\mathbf{i}}^{\varsigma,j}$ are computed according to the quasi-optimal monotone restriction operator in (7.16) (see Corollary 7.4 for the construction). This ensures the monotonicity of the obstacle functions, i.e. the conclusion

$$D^{\varsigma,j} \subset D^{*,\varsigma,j} \tag{7.60}$$

is valid. The following Lemma ensures the global convergence of the MMG method. A proof in the context of linear finite element functions can be found in [GK09]. Since the same proof is also valid for tensor product B-splines we refer the reader to [GK09].

Lemma 7.5. *Let us assume that the iterates u_L^ς are produced by an algorithm of the form*

$$\overline{u}_L^\varsigma = \mathsf{Sm}_L(u_L, \tilde{\psi}_L), \quad u_L^{\varsigma+1} = \mathcal{C}_L(\overline{u}_L^\varsigma) \tag{7.61}$$

with $\mathsf{Sm}_L : \mathcal{K}_L^0 \to \mathcal{K}_L^0$ denoting the projected Gauss-Seidel iteration and some $\mathcal{C}_L : \mathcal{K}_L^0 \to \mathcal{K}_L^0$ satisfying the monotonicity condition

$$\mathcal{J}(\mathcal{C}_L(w)) \leq \mathcal{J}(w) \quad \textit{for all } w \in \mathcal{K}_L^0. \tag{7.62}$$

Then $u_L^\varsigma \to u_L$ holds for any initial iterate $u_L^0 \in \mathcal{K}_L^0$.

A direct consequence of Lemma 7.5 is the global convergence of the MMG method formulated in the following Theorem.

Theorem 7.6. *(Global convergence)*
The monotone multigrid method converges for any initial iterate $u_L^0 \in \mathcal{K}_L^0$.

Proof. The conclusion in (7.60) is valid, since the coarse grid approximations provided in Corollary 7.4 are monotone, thus the result follows directly from Lemma 7.5. □

Assuming a discretization with piecewise linear finite element basis functions, quasi-optimal and monotone restriction operators and the strict complementarity condition the author in [Kor94] has shown that the MMG method is asymptotically reducing to a *linear extended relaxation method*. This offers the analysis of the asymptotic convergence rates of the MMG method. Since we have provided quasi-optimal monotone restriction

operators for tensor product B-splines the proof in [Kor94] works also for tensor product B-splines. For the same proof in the context of B-splines we refer to [Hol04]. Thus, we state the following Theorem without proof.

Theorem 7.7. *Assume that the strict complementarity condition in (7.53) is satisfied. Then after a sufficiently large number of iteration steps ς_0, the MMG method is reducing to a linear extended relaxation method for the reduced linear problem in (7.54).*

In order to analyze the asymptotic convergence rates of the MMG method for finite elements the author in [Kor94] has applied the theory of subspace correction method, for a detailed explanation of subspace methods see [Yse93]. To apply this theory two assumptions have to be satisfied. One assumption is the stability of the decomposition and the other assumption is a Cauchy-Schwarz type inequality, the result can be found in [Yse93, Theorem 5.1].
One main difficulty in the context of the MMG method is to construct an appropriate subspace decomposition of the reduced subspace in (7.55) since the corresponding subdomain where the solution is given by the variational equation in 7.55 is a priori unknown. Thus, this subdomain has no exact representation on coarser grids. In order to prove the stability of the decomposition the idea of [KY94] is to construct appropriate subspace decompositions of the corresponding finite element spaces by way of an embedding of the domain under consideration into a simpler domain like a square or a cube. To do so corresponding interpolation like or L^2-like projection and corresponding estimates have to be developed, see [Yse86] or [Yse90]. To the best of my knowledge such theory is still unknown for (tensor product) splines or hierarchical splines. The available literature for instance in isogeometric analysis (see [HMS19, CH19]) deals with polygonal domains and L^2-projections to get h-independent stable subspace decompositions.
However, assuming the strict complementarity condition in (7.53) and quasi-optimal restriction operators the author in [Kor94] has provided the following a posteriori estimates for the MMG method with linear finite elements restricted to the 2D case

$$\|u_L - u_L^{\varsigma+1}\| \leq (1 - c(L+1)^{-3})\|u_L - u_L^{\varsigma}\| \tag{7.63}$$

for large enough $\varsigma \geq \varsigma_0$ and a constant $c < 1$. Thus, the asymptotic convergence rates are bounded by $(1 - c(L+1))^{-3}$ in the worst case.

8. Numerical results for variational inequalities

8.1. Test problem

In this section we construct a test problem for an elliptic variational inequality to verify the B-spline discretization and the corresponding solution algorithm. In order to check the convergence for B-splines of order $k = 2, 3, 4$ a suitable test problem for a one dimensional elliptic variational inequality shall satisfy the following requirements:

i) Let us assume that the constraints are non-binding for $S \leq S_f$. The solution is supposed to be smooth enough, where the constraints are non-binding, i.e. $w \in H^2(I) \cap H^4((S_f, S_{\max}])$. Thus, due to the Sobolev embedding $H^2(I) \hookrightarrow C^1(I)$ for $I \subset \mathbb{R}$ the solution satisfies the following conditions at the contact point S_f:

$$w(S_f) = \mathcal{H}(S_f), \quad w'(S_f) = \mathcal{H}'(S_f), \tag{8.1}$$

where the obstacle is defined as $\mathcal{H}(S) := \max\{K - S, 0\}$.

ii) The solution is supposed to fulfill the following free boundary problem:
Find $w := w(S)$ such that

$$\begin{aligned} w(S) &= \mathcal{H}(S), \quad \text{for } S \leq S_f \\ \mathcal{L}w &= f(S), \quad \text{for } S > S_f, \end{aligned} \tag{8.2}$$

where $H^2(I) \to L^2(I)$ is a bounded linear operator. Formulating (8.2) as linear complementarity problem leads to

$$\mathcal{L}w > f(S), \qquad w(S) = \mathcal{H}(S) \ \text{ for } S \leq S_f \tag{8.3}$$

and

$$\mathcal{L}w = f(S), \qquad w(S) > \mathcal{H}(S) \ \text{ for } S > S_f. \tag{8.4}$$

In order to verify the B-spline discretization for a variational inequality and the corresponding PSOR-algorithm, we consider the following test problem:

Test problem 8.1. *(Elliptic free boundary problem)*
Let us consider the following free boundary problem: Find $w(S)$ for $S \in [0, 20]$

$$w(S) = \mathcal{H}(S), \quad \text{for } S \leq S_f \tag{8.5}$$

$$\mathcal{L}^1 w = f(S), \quad \text{for } S > S_f, \quad \text{with} \quad \mathcal{L}^1 w := -\frac{\partial^2 w}{\partial S^2} \text{ and } w(20) = 0. \tag{8.6}$$

It is assumed that $S_f < K = 10$ as in the Black-Scholes Problem. First we have to construct a right hand side $f^1(S)$ such that (8.3) and also (8.4) are satisfied. Since $\mathcal{L}^1\mathcal{H}(S) = 0$ for $S \leq S_f$ we choose $f^1(S)$ such that $f^1(S) < 0$ for $S \leq S_f$ and $\mathcal{L}^1 w = f(S)$ for $S > S_f$. A suitable choice, where $w(S)$ is smooth for $S > S_f$, is

$$\begin{aligned} f^1(S) :&= \mathcal{L}^1(-15.5\cos\left(\frac{1}{40}\pi S\right) + a_2(S_f)S + a_3(S_f)) \\ &= -15.5\frac{\pi^2}{40^2}\cos\left(\frac{1}{40}\pi S\right) < 0 \quad \text{for} \quad S \in [0, 20). \end{aligned} \tag{8.7}$$

The coefficients $a_2(S_f)$, $a_3(S_f)$ and the free boundary S_f result from the conditions (8.1) and the boundary condition

$$w(S_f) = \mathcal{H}(S_f), \; w'(S_f) = \mathcal{H}'(S_f) \text{ and } w(20) = 0,$$

which leads to $S_f = 3.153096802086939$. Finally, the exact solution of the free boundary problem, where the right hand side is defined as in (8.7), is explicitly given by

$$w(S) = \begin{cases} \mathcal{H}(S) & \text{for } S \leq S_f \\ a_1\cos(\frac{1}{40}\pi S) + a_2(S_f)S + a_3(S_f) & \text{for } S > S_f. \end{cases} \tag{8.8}$$

and $a_1 := -15.5$. By similar arguments as in Section 2.2.1 we obtain the associated variational inequality: Find $w := w(S) \in \mathcal{K}_1 := \{\varphi \in H^1(I) : \varphi \geq \mathcal{H}(S), \varphi(20) = 0\}$ for $S \in I = [0, 20]$ such that

$$a^1(w, \varphi - w) \geq (f^1, \varphi - w) \quad \text{for all } \varphi \in \mathcal{K}_1, \tag{8.9}$$

with $a^1(w, \varphi) := \int_I \frac{\partial w}{\partial S}\frac{\partial \varphi}{\partial S}\mathrm{d}S$.

For the discretization of the elliptic variational inequality defined in (8.9), we proceed as in Section 4.2.2. Figure 8.1 shows the two different knot series used for the discretization of the elliptic variational inequality. Due to a homogenization with the function $\mathcal{H}(S)$ the solution $u(S) := w(S) - \mathcal{H}(S)$ is not differentiable at $S = K$. Therefore, a knot series with $k - 1$ repeating knots at this point is used to stabilize the numerical computation. The $k - 2$ coinciding knots at $S = S_f$ should ensure that the resulting spline solution has the same regularity as the exact solution. Note, that in practice the free boundary is a priori unknown and a uniform knot series without repeating knots

at the free boundary is used. To solve the discrete variational inequality, the B-spline based projected Gauss-Seidel method as described in Section 7.1 is used.

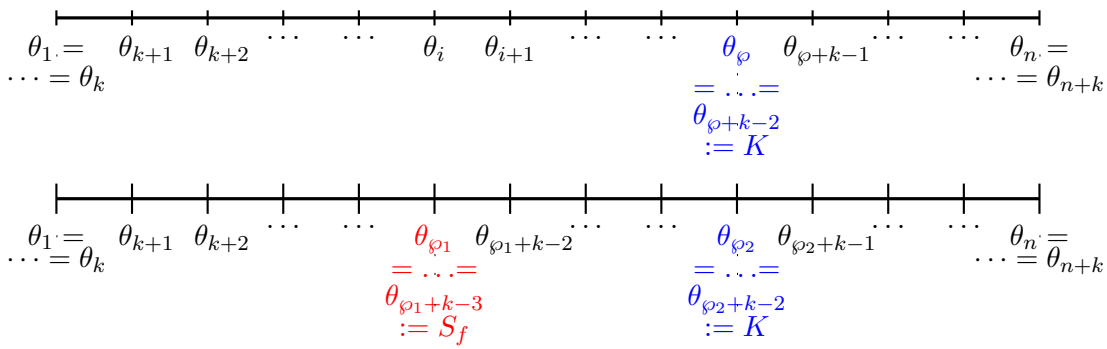

Figure 8.1.: Knot series used for the discretization of test problem 8.1: Knot series with uniform grid size and $k-1$ repeating knots at $S = K$ (above). Knot series with uniform grid size and $k-2$ repeating knots at $S = S_f$ and $k-1$ repeating knots at $S = K$ (below).

To study the discretization error, calculations are performed on a sequence of grids by doubling the number of knots. Table 8.1 contains the results of computations for different orders of B-spline basis functions $k = 2, 3, 4$. In comparison to linear basis ($k = 2$) functions the $H^1(I)$-error is significantly reduced for $k = 3, 4$ but is barely improved for cubic basis functions ($k = 4$). As expected it can be observed that the convergence rate is not optimal for $k = 3, 4$. This indicates the fact that the solution does not belong to $H^{k-1}(I)$ for $k = 3, 4$ due to the limited smoothness at the contact point S_f. We also consider the error between the exact and the approximated free boundary. For all considered orders of B-spline basis functions the error $|S_f - S_f^h|$ decreases as N is increased but is not improved for increasing order k.
In order to verify the B-spline-discretization for variational inequalities, calculations are performed corresponding to a knot series with uniform grid size and $k-2$ repeating knots at S_f, such that due to the properties of splines (see properties 3.16) the corresponding spline approximation has the same smoothness as the exact solution at the contact point S_f. The results of the computations for $k = 3$ and $k = 4$ are presented in Table 8.2. It can be observed that, indeed, the same optimal rate of convergence as in the unrestricted case is achieved. This is in agreement with the following theoretical consideration:
Since $a^1(\cdot,\cdot)$ in Test problem 8.1 is bounded and coercive, the following a priori estimation from Lemma 3.22 is satisfied

$$\|w - w_h\|^2_{H^1(I)} \lesssim \inf_{\varphi_h \in \mathcal{K}_h^1} \Big(\|\varphi_h - w\|^2_{H^1(I)} + \langle Aw - f, \varphi_h - w \rangle \Big) \quad \text{for all } \varphi_h \in \mathcal{K}_h^1 \subset \mathcal{K}^1. \tag{8.10}$$

Since the free boundary is detected exactly and the approximated solution is given by the exact obstacle function $\mathcal{H}(S)$ one has $\varphi_h = w$ for all $S \leq S_f$ and

$$\begin{aligned}\langle Aw - f, \varphi_h - w\rangle &= \int_0^{S_f} (Aw - f)(\varphi_h - w)\mathrm{d}S + \int_{S_f}^{S_{\max}} (Aw - f)(\varphi_h - w)\mathrm{d}S \\ &= \int_{S_f}^{S_{\max}} (Aw - f)(\varphi_h - w)\mathrm{d}S. \end{aligned} \tag{8.11}$$

For $S \geq S_f$ the solution satisfies $Aw - f = 0$ such that

$$\langle Aw - f, \varphi_h - w\rangle = \int_{S_f}^{S_{\max}} (Aw - f)(\varphi_h - w)\mathrm{d}S = 0. \tag{8.12}$$

Applying (8.12) to (8.10) the approximation properties for splines with $\varphi_h = I_h w$ and $w \in H^4((S_f, S_{\max}))$ indicates

$$\|w - w_h\|^2_{H^1(I)} \lesssim \inf_{I_h w \in \mathcal{K}_h^1} \left(\|I_h w - w\|^2_{H^1(I)}\right) \lesssim h^{k-1}. \tag{8.13}$$

Hence, in this special case the B-spline approximation leads to the same optimal convergence rate as in the unrestricted case.
In Figure 8.2 we present some results for a B-spline discretization with a refined knot series at $(S_f, S_f + h)$ for $k = 3, 4$ and a uniform knot series for $k = 2, 3, 4$. They confirm that the discretization error $\|w - w_h\|_{H^1(I)}$ for quadratic and cubic B-splines is only perturbed by the free boundary.
Finally, it is also remarkable that for all considered knot series the PSOR-algorithm identifies the knot next to the contact point as the approximated free boundary. In particular, the PSOR-algorithm converges to the exact free boundary S_f, if the contact point S_f is aligned with $k - 2$ knots.

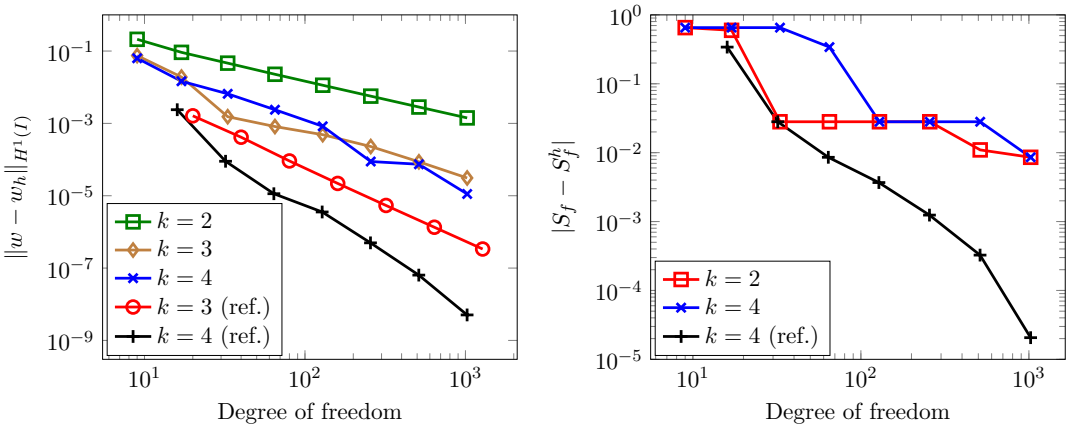

Figure 8.2.: Convergence of the error $\|w - w_h\|_{H^1(I)}$ (above) and $|S_f - S_f^h|$ (below) for test problem 8.1 computed with a B-spline discretization of order k for a uniform knot series and a refined knot series at $[S_f, S_f + h]$ (ref.).

	$k=2$			$k=3$		
N	$\|w-w_h\|_{H^1(I)}$	Rate	$\|S_f-S_f^h\|$	$\|w-w_h\|_{H^1(I)}$	Rate	$\|S_f-S_f^h\|$
9	$2.11e{-1}$	–	$6.53e{-1}$	$7.48e{-2}$	–	$6.53e{-1}$
17	$9.20e{-2}$	1.20	$5.97e{-1}$	$1.91e{-2}$	1.97	$6.53e{-1}$
33	$4.63e{-2}$	0.99	$2.81e{-2}$	$1.53e{-3}$	3.64	$2.81e{-2}$
65	$2.28e{-2}$	1.02	$2.81e{-2}$	$8.19e{-4}$	0.90	$2.81e{-2}$
129	$1.14e{-2}$	1.01	$2.81e{-2}$	$4.87e{-4}$	0.75	$2.81e{-2}$
257	$5.67e{-3}$	1.00	$2.81e{-2}$	$2.29e{-4}$	1.09	$2.81e{-2}$
513	$2.83e{-3}$	1.00	$1.10e{-2}$	$8.67e{-5}$	1.40	$2.81e{-2}$
1,025	$1.42e{-3}$	1.00	$8.57e{-3}$	$3.10e{-5}$	1.48	$8.57e{-3}$

	$k=4$		
N	$\|w-w_h\|_{H^1(I)}$	Rate	$\|S_f-S_f^h\|$
9	$6.32e{-2}$	–	$6.53e{-1}$
17	$1.46e{-2}$	2.11	$6.53e{-1}$
33	$6.57e{-3}$	1.15	$6.53e{-1}$
65	$2.39e{-3}$	1.46	$3.41e{-1}$
129	$8.32e{-4}$	1.52	$2.81e{-2}$
257	$8.83e{-5}$	3.24	$2.81e{-2}$
513	$7.41e{-5}$	0.25	$2.81e{-2}$
1,025	$1.12e{-5}$	2.73	$8.57e{-3}$

Table 8.1.: Errors for Test problem 8.1 with a B-spline discretization of order k and a uniform grid size. S_f is not aligned with a knot, the corresponding knot series is depicted in Figure 8.1 (above).

	$k=3$			$k=4$		
N	$\|w-w_h\|_{H^1(I)}$	Rate	$\|S_f-S_f^h\|$	$\|w-w_h\|_{H^1(I)}$	Rate	$\|S_f-S_f^h\|$
10	$5.93e{-3}$	–	0	$1.35e{-4}$	–	0
18	$1.40e{-3}$	2.08	0	$1.71e{-5}$	2.98	0
34	$3.48e{-4}$	2.01	0	$2.14e{-6}$	3.00	0
66	$8.63e{-5}$	2.01	0	$2.68e{-7}$	3.00	0
130	$2.15e{-5}$	2.00	0	$3.37e{-8}$	2.99	0
258	$5.38e{-6}$	2.00	0	$4.24e{-9}$	2.99	0
514	$1.35e{-6}$	2.00	0	$5.31e{-10}$	3.00	0
1,026	$3.37e{-7}$	2.00	0	$6.86e{-11}$	2.95	0

Table 8.2.: Errors for test Problem 8.1 with a B-spline discretization of order $k = 3, 4$ and a uniform grid size. S_f is aligned with $k-2$ repeating knots, the corresponding knot series is depicted in Figure 8.1 (below).

8.2. Pricing American options with Black-Scholes and Heston's model

The code developed for this thesis is implemented in Matlab R2018b. Since the PSOR method cannot be vectorized due to the projection step, a code implemented in Matlab is very slow. Thus, the PSOR method is written in a MEX function based on a C code. In this section, we present several numerical results for American options with Black-Scholes and Heston's model. Examples 8.2, 8.3, 8.4 and 8.5 with a small time interval $[0, T)$ are devoted to study the spatial discretization for different initial condition (payoff function). In Example 8.6 we choose a larger time interval to clarify the temporal discretization error.

Example 8.2. *(American put option in the Black-Scholes model)*
Consider an American put option for the discrete Black-Scholes Problem 4.13 with strike price $K = 5$ and expiration date $T = 0.5$. The volatility is $\sigma = 0.6$, the dividend rate is $D_0 = 0$ and the risk free interest rate is taken as $r = 0.01$. With these parameters the option price satisfies the homogeneous boundary condition for $S_{\max} = 20$. Thus, the spatial domain is choosen as $I = (0, 20)$.

Example 8.3. *(American call option in the Black-Scholes model)*
Consider an American call option for the discrete Black-Scholes Problem 4.13 with strike price $K = 7.5$ and expiration date $T = 0.5$. The volatility is $\sigma = 0.4$, the dividend rate is $D_0 = 0.05$ and the risk free interest rate is taken as $r = 0.01$. With these parameters the optimal exercise prices $S_f(t)$ are smaller than $S_{\max} = 20$ for all times t. Thus, the spatial domain is chosen as $I = (0, 20)$.

Example 8.4. *(American butterfly-spread option in the Black-Scholes model)*
Consider an American butterfly-spread option for the discrete Black-Scholes Problem 4.13 with strike prices $K_1 = 5$, $K_2 = 15$, $K = (K_1 + K_2)/2$ and expiration date $T = 0.5$. The volatility is $\sigma = 0.4$, the dividend rate is $D_0 = 0.10$ and the risk free interest rate is taken as $r = 0.01$. In particular, the parameters are chosen such that the optimal exercise prices satisfy $K_1 \leq S_{f_1}(t) < K \leq S_{f_2}(t) \leq K_2$ for all times $t > 0$. Thus, the spatial domain is chosen as $I = (0, 40)$.

Example 8.5. *(American put option in Heston's model)*
Consider an American put option for the discrete Heston Problem 5.14 with strike price $K = 10$ and expiration date $T = 0.25$. The parameters are chosen as $\rho = 0.1$, $\xi = 0.9$, $\kappa = 5.0$, $\gamma = 0.16$, $\lambda = 0$ and $r = 0.04$.
The domain is chosen as $\Omega := (x_{\min}, x_{\max}) \times (v_{\min}, v_{\max}) = (-2, 2) \times (0.075, 1)$. Thus, the transformed option price satisfies the boundary condition on $\partial\Omega$.

Example 8.6. *(American call option in the Black-Scholes model)*
Consider an American call option for the discrete Black-Scholes Problem 4.13 with strike price $K = 7.5$ and expiration date $T = 2$. The volatility is $\sigma = 0.2$, the dividend rate is $D_0 = 0.05$ and the risk free interest rate is taken as $r = 0.01$. With these parameter the optimal exercise prices $S_f(t)$ are smaller than $S_{\max} = 20$ for all times t. Thus, the spatial domain is chosen as $I = (0, 20)$.

For the examples corresponding the *Black-Scholes model* we denote by N the dimension of the discrete solution space $\mathcal{V}_h^0$ as defined in (4.33) with a uniform grid size h. Since no exact solution formula is known, we compute a reference solution $V_{\text{ref}}(0,S)$ with cubic B-splines ($k=4$) at time $t=0$ with $\#\mathbb{T}=2^{19}$ time steps and $N=2^{12}+1+2|\text{rep}|$ degrees of freedom. We choose a uniform sequence of knots with 3 repeating knots at $S=K$ for a put and call option, thus the sequence of knots has $2|\text{rep}|=2$ additional knots. In the case of the butterfly-spread option we use 3 repeating knots at $S=K_1,K,K_2$, this results in $2|\text{rep}|=6$ additional knots. For each grid size h the approximation of the discretization error in the $H^1(I)$-norm for $t=0$ is computed by the following formula

$$\begin{aligned}\left(E_h^{V,\Delta}(0)\right)^2 &:= \|V_{\text{ref}}(0,\cdot)-V_h(0,\cdot)\|_{H^1(I)}^2 \\ &= \|V_{\text{ref}}(0,\cdot)-V_h(0,\cdot)\|_{L^2(I)}^2+\|\Delta_{\text{ref}}(0,\cdot)-\Delta_h(0,\cdot)\|_{L^2(I)}^2 \end{aligned} \tag{8.14}$$

with y_h,Δ_h as defined in (6.1), (6.3). Since $V_{\text{ref}}(0,\cdot)$ and $V_h(0,\cdot)$ are polynomials of degree less than or equal $k-1=3$ on $[\theta_i^{\text{ref}},\theta_{i+1}^{\text{ref}}]$, the integrals can be computed exactly with the Gauss-Legendre quadrature rule. Thus, the following formula is valid for the approximation of the convergence rate:

$$\text{rate} \approx \frac{\log\left(\frac{E_h^{V,\Delta}}{E_{\frac{h}{2}}^{V,\Delta}}\right)}{\log(2)}. \tag{8.15}$$

Since we are also interested in the approximation error of the second derivatives for American options we approximate the error of $\Gamma_h(0,S)$ in the L^2-norm. It follows from the discussion in Chapter 6 that the second derivative for an American option is discontinuous at the free boundary. Thus, the error of Gamma in the neighbourhood of the free boundary is much higher as there where the solution is twice differentiable. As a result it cannot be excepted that the error of Γ_h in the L^2-norm on the whole spatial domain is reduced when the B-spline order is increased. Thus, we ignore the error of Γ_h in a neighbourhood, where the support of B-splines intersects the free boundary. More precisely, the error for the approximation of Gamma is approximated by the following formula

$$\widetilde{E}_h^{\Gamma}(0) := \|\Gamma_{\text{ref}}(0,\cdot)-\Gamma_h(0,\cdot)\|_{L^2(I\setminus\mathfrak{S}_{S_f^H})}, \tag{8.16}$$

where $\mathfrak{S}_{S_f^H}$ denotes the union of the supports of B-spline basis functions whose support intersects $S_f^H(0)$ on the coarsest grid with grid size H.
For *Heston's model*, we denote by N and M the dimension of the solution space $\mathcal{V}_h^0$ as defined in (5.36) in the x- and v- direction. Then we compute a reference solution $V_h(0,S,v)$ for the set of boundary conditions in (2.111) with cubic B-splines and a uniform knot sequence with 3 repeating knots at $x=0$. The reference solution is computed with $NM=(2^{10}+3)(2^9+2)$ degrees of freedom and $\#\mathbb{T}=2^{14}$ time steps.

For each grid size $h := (h_x, h_v)$ the discretization error at $t = 0$ in the $H^1(\Omega_{\mathcal{L}})$-norm is given by

$$\begin{aligned}(E_h^{V,\Delta,\nu}(0))^2 : = &\|K\exp(x)(y_{\text{ref}}(0,x,v) - y_h(0,x,v))\|^2_{L^2(\Omega)} \\ &+ \|K\exp(x)(\Delta_{\text{ref}}(0,x,v) - \Delta_h(0,x,v))\|^2_{L^2(\Omega)} \\ &+ \|K\exp(x)(\nu_{\text{ref}}(0,x,v) - \nu_h(0,x,v))\|^2_{L^2(\Omega)} \end{aligned} \tag{8.17}$$

with y_h, Δ_h and ν_h as defined in (6.5), (6.8), (6.10). The additional term $K\exp(x)$ results from the transformation of the integral from to S to x coordinates. The integrals are approximated with a Gauss-Legendre quadrature for two dimensional integrals. Thus, the approximation of the convergence rate is computed analogously to (8.15). To study the convergence of Gamma in Heston's model, we proceed as described for the Black-Scholes model. For Heston's model the free boundary at a fixed time is a line depending on the variance v. Thus, the error for the approximation of Gamma is approximated by the following formula

$$\widetilde{E}_h^{\Gamma}(0) := \|K\exp(x)(\Gamma_{\text{ref}}(0,\cdot,\cdot) - \Gamma_h(0,\cdot,\cdot))\|_{L^2(\Omega\setminus\mathfrak{S}_{x_f^H})}, \tag{8.18}$$

where $\mathfrak{S}_{x_f^H}$ denotes the union of the supports of B-spline basis functions whose support intersects $x_f^H(0,v)$ on the coarsest grid with grid size H. The formula for the approximation of Γ_h in Heston's model can be found in (6.9)

8.2.1. Choice of the boundary conditions for Heston's model

In this section, we briefly discuss the choice of the boundary condition for an American put option in Heston's model on

$$\Upsilon_3 := \{x \in (x_{\min}, x_{\max}) : v = v_{\min}\}.$$

We distinguish between the boundary conditions in (2.110) with a Dirichlet boundary condition on Υ_3 and (2.111) with the Neumann boundary data on Υ_3. A graphical illustration of the approximation of Vega for the different boundary data on Υ_3 is presented in Figure 8.3 (above). As you can see imposing a Dirichlet boundary condition on Υ_3 results in an unstable approximation of Vega while the approximation of Vega for a Neumann boundary condition is stable. This is due to the fact that for $v \to v_{\min}$ the Heston equation becomes increasingly hyperbolic and imposing a Dirichlet boundary condition is not appropriate, as pointed out by [CP99]. The option price becomes clearly flat in this region that is why the authors in [CP99] recommend a Neumann boundary condition on Υ_3.
The approximation of Gamma for different boundary conditions on Υ_3 is depicted in Figure 8.3 (below). Since the Dirichlet boundary condition is given by the non-smooth payoff function, the approximation of Gamma is also unstable on Υ_3. Thus, for further

numerical computations a Neumann boundary condition on Υ_3 is imposed to stabilize the approximation of the partial derivatives Vega and Gamma.

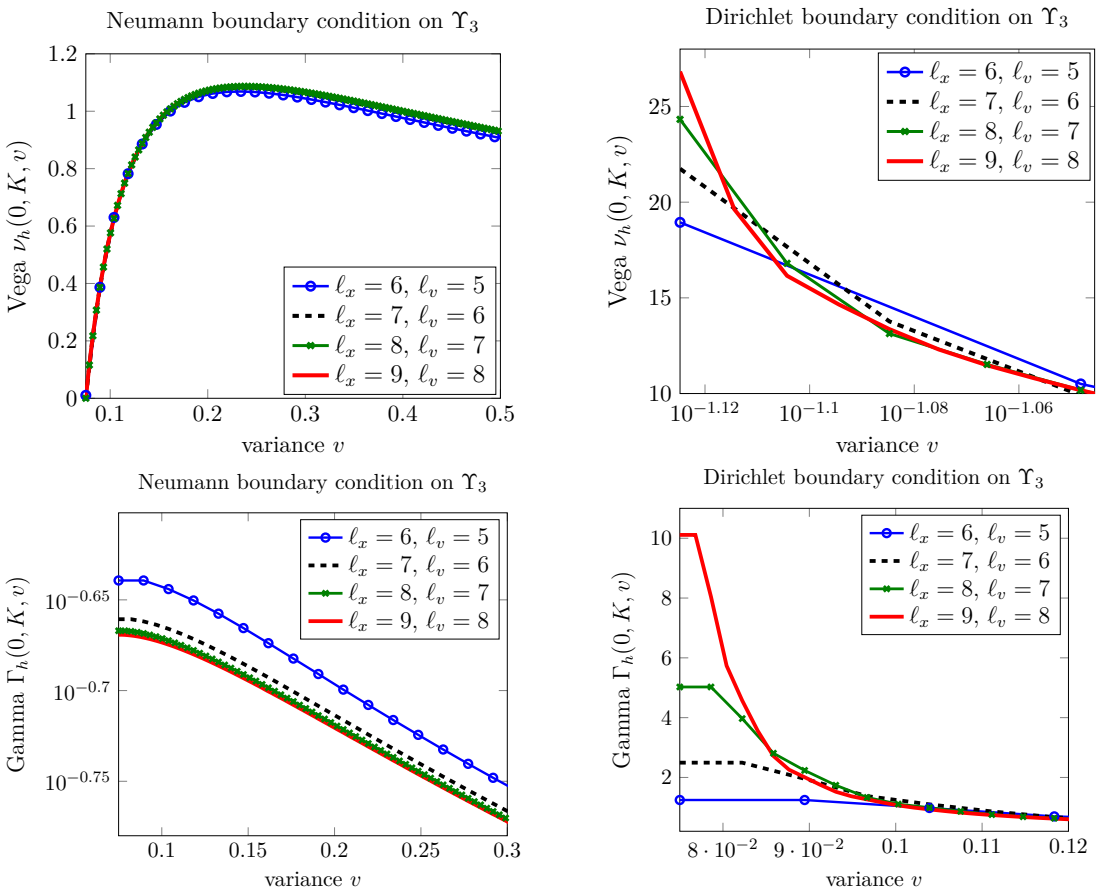

Figure 8.3.: Cubic spline approximation of *Vega* (above) and *Gamma* (below) for an American option in Heston's model at $S = K$ with the set of boundary conditions from (2.111) (left) and (2.110) (right) for different levels ℓ_x, ℓ_v. The parameters for Heston's model are chosen as in Example (8.5)

8.2.2. Influence of coinciding knots on the approximated option value and Greeks

In this section, we analyze the influence of coinciding knots on the Greeks for American options in the Black-Scholes and Heston's model. Therefore, we consider the approximation of today's American option price $V_h(0, \cdot)$ and its sensitivities at $t = 0$.

First, we study the influence of repeating knots on the approximated American option value and Greeks in the Black-Scholes model for different payoffs as described in Example 8.2, Example 8.3 and 8.4. The right plots in Figure 8.4, Figure 8.5 and Figure 8.6 show that using a spatial cubic B-spline discretization with no repeating knots (no rep.) at the critical points $S = K$ for a put and call option or $S = K_1, K, K_2$ for a butterfly option being the locations of the discontinuities in the first derivatives of the terminal conditions, causes significant oscillations near these points.
The right plots in the figures named above confirm that using a cubic B-spline discretization with $k - 1 = 3$ repeating knots (rep.) at $S = K$ for a put and call option or at $S = K_1, K, K_2$ for a butterfly-spread option facilitates a pointwise accurate approximation of the partial derivatives up to order two. As expected, the second derivatives Γ_h are discontinuous at the free boundary $S = S_f^h$ for a put or call option and at $S = S_{f_1}^h, S_{f_2}^h$ for a butterfly-spread option.
Furthermore, Delta for a butterfly-spread option value, the first partial derivative, is discontinuous at $S = S_{f_2}$. This is due to the fact that the butterfly-spread option value is given by the payoff function (obstacle) at $S = S_{f_2}^h(0) = 10$, where the first derivative of the payoff is discontinuous.
Similar considerations apply to the approximation of today's American put option price and its partial derivatives with respect to the underlying price S in Heston's model. A graphical illustration of the approximation of today's put option price in Heston's model, its first derivative Delta and its second derivative Gamma using a spatial cubic tensor product B-spline discretization are presented in Figure 8.7. The right plots show the numerical approximation with no coinciding knots at $S = K$, which is the location of the discontinuity in the first derivative of the terminal condition $\mathcal{H}_{\mathrm{P}}(S)$. It can be observed here, that a cubic spline approximation with no repeating knots at $S = K$ causes significant oscillations along the points (K, v), $v \in (v_{\min}, v_{\max})$. The left plots confirm that a cubic B-spline discretization with coinciding knots at $S = K$ facilitates a pointwise accurate approximation of the partial derivatives up to order two. Moreover, it can also be seen that the approximation of Gamma, the second derivative of today's American put option price, is discontinuous along the approximated free boundary $S_f^h(0, v)$.
Since the terminal condition is constant in v, this concept of coinciding knots has no visible influences on Vega, the first partial derivative with respect to the variance v. A graphical illustration of the approximation of Vega is presented in Figure 8.8. From the approximation of Vega it can be seen, that the Neumann boundary conditions on Υ_v are clearly satisfied. This verifies the derived bilinear form in (2.128) in some sense.

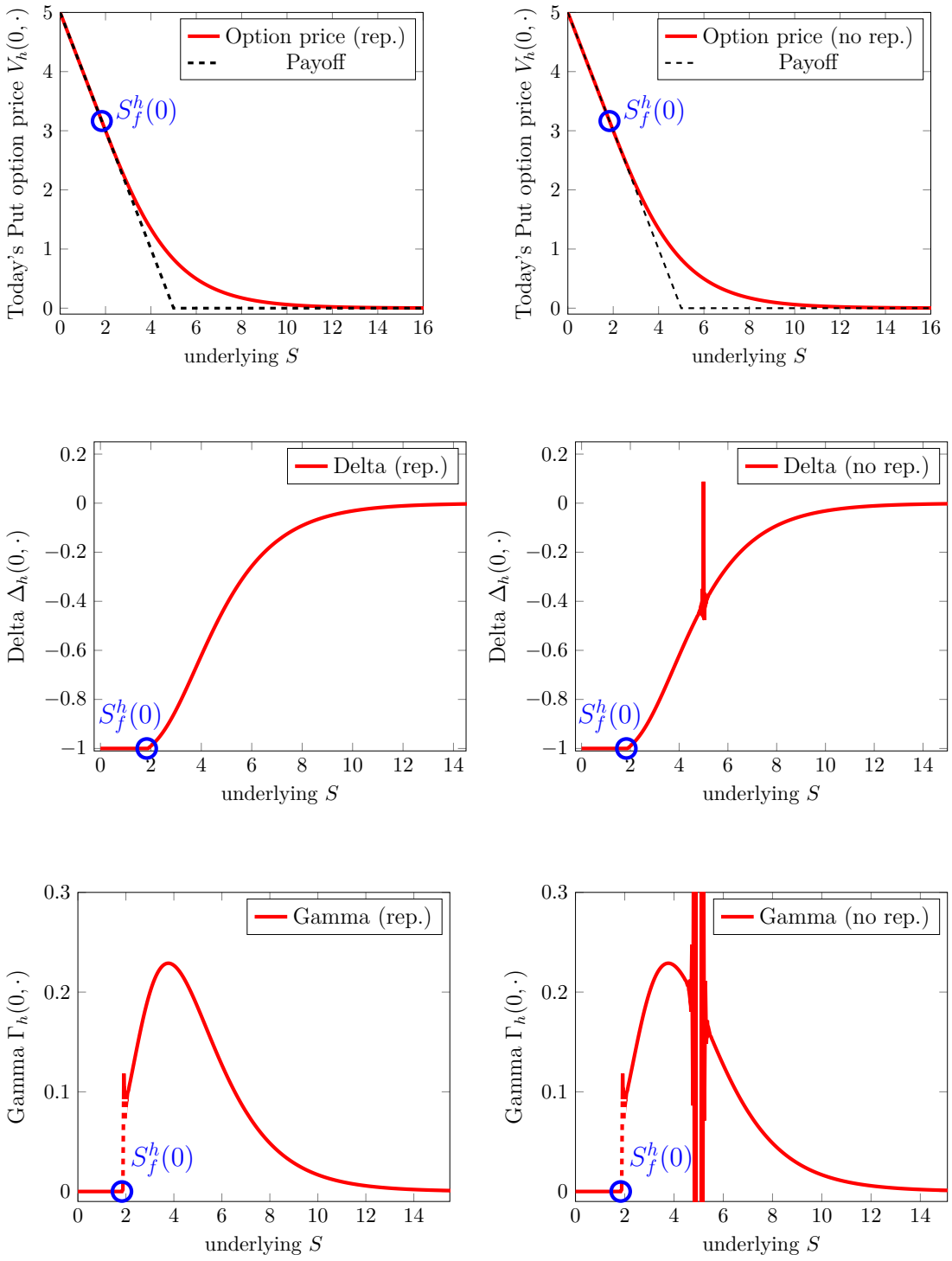

Figure 8.4.: Numerical approximation of today's American put option price in *Black-Scholes model*, its first derivative *Delta* and its second derivative *Gamma* for cubic B-splines ($k = 4$) with repeating knots (rep.) at $S = K = 5$ and no repeating knots (no rep.) on level $\ell = 9$. The parameters for the Black-Scholes model are chosen as in Example 8.2.

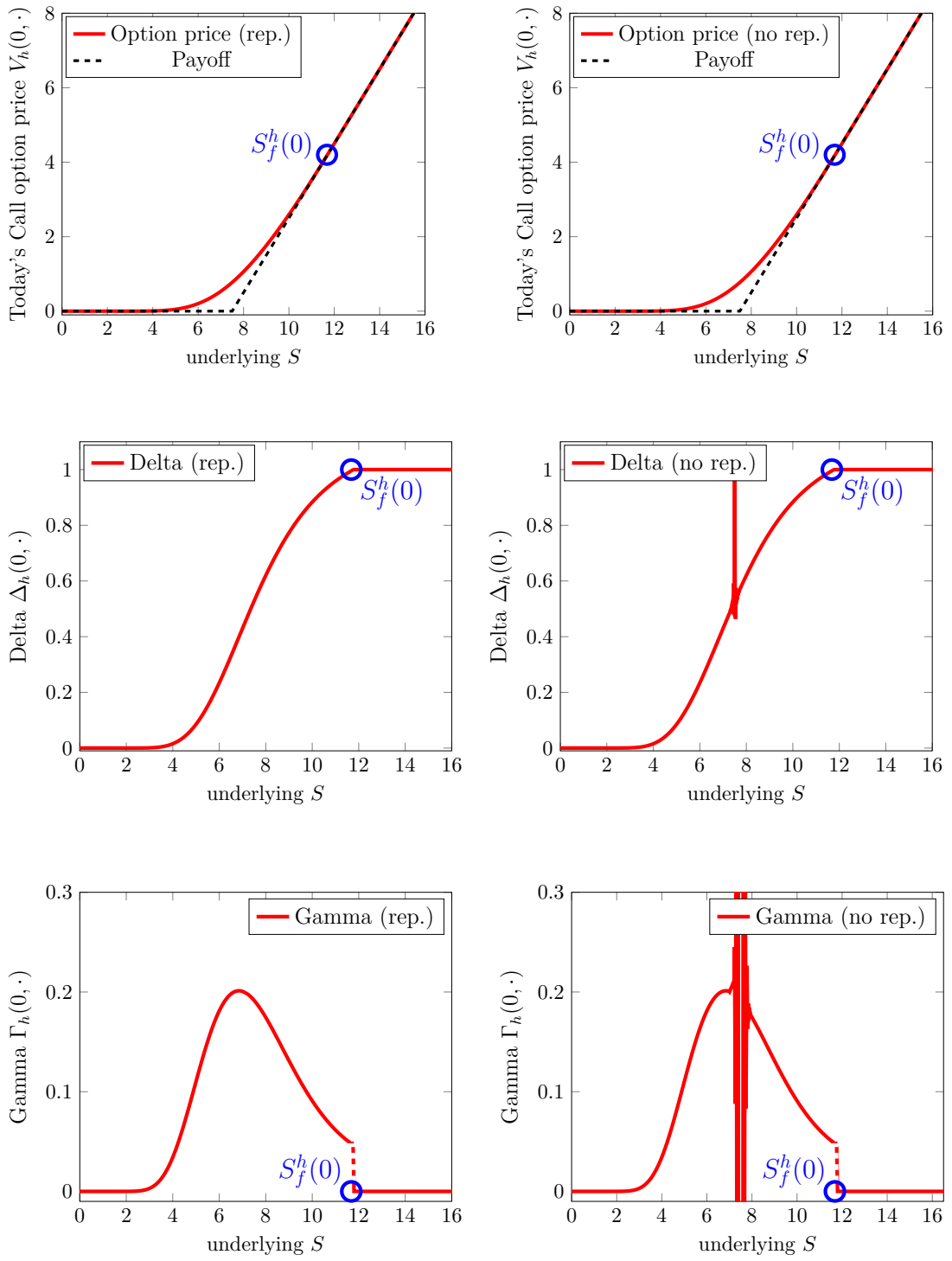

Figure 8.5.: Numerical approximation of today's American call option price in the *Black-Scholes model*, its first derivative *Delta* and its second derivative *Gamma* for cubic B-splines ($k = 4$) with repeating knots (rep.) at $S = K = 7.5$ and no repeating knots (no rep.) on level $\ell = 9$. The parameters for the Black-Scholes model are chosen as in Example 8.3.

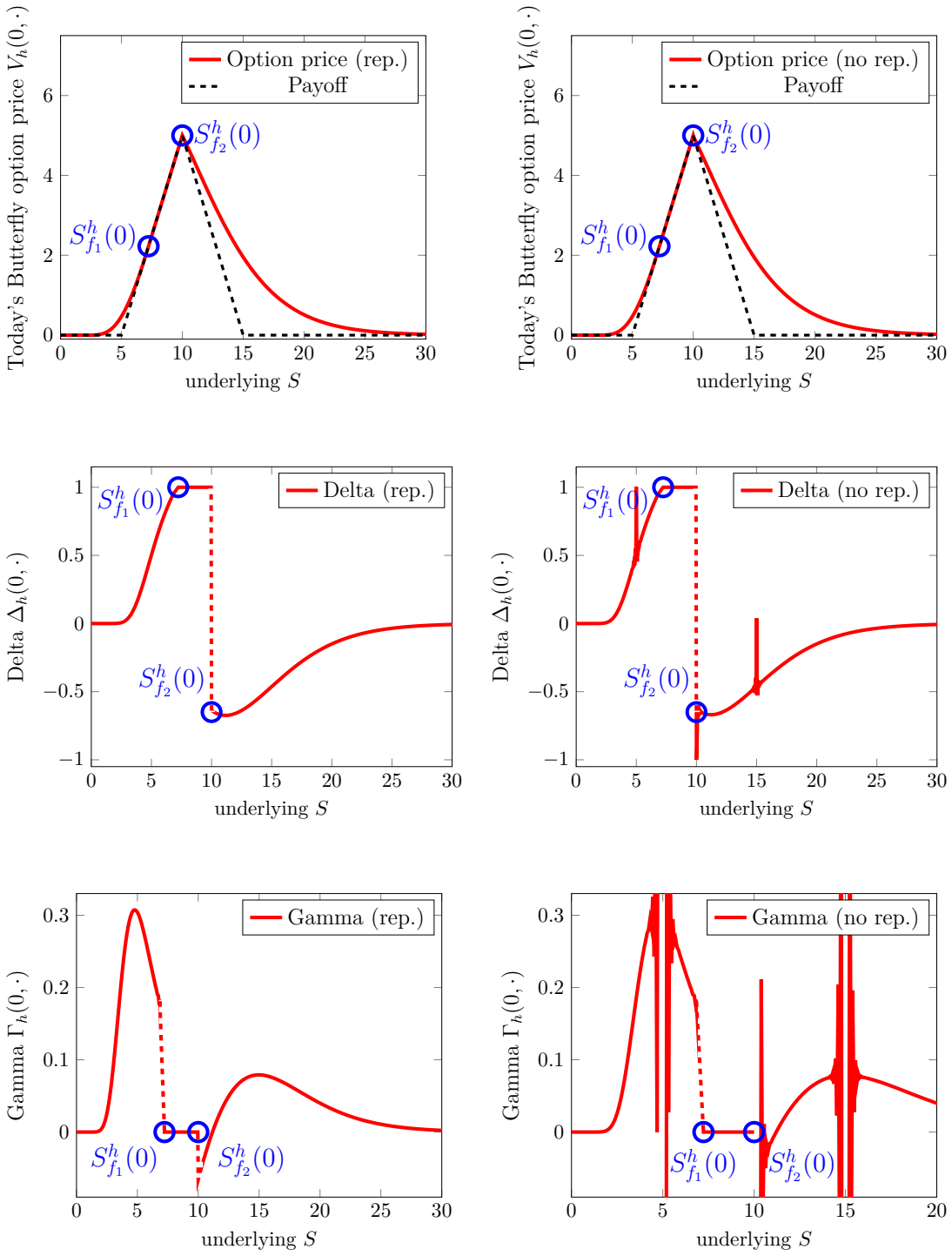

Figure 8.6.: Numerical approximation of today's American butterfly-spread option price in the *Black-Scholes model*, its first derivative *Delta* and its second derivative *Gamma* for cubic B-splines ($k = 4$) with repeating knots (rep.) at $S = K_1, K, K_2$ with $K_1 = 5$, $K = 10$, $K_2 = 15$ and no repeating knots (no rep.) on level $\ell = 9$. The parameters for the Black-Scholes model are chosen as in Example 8.4.

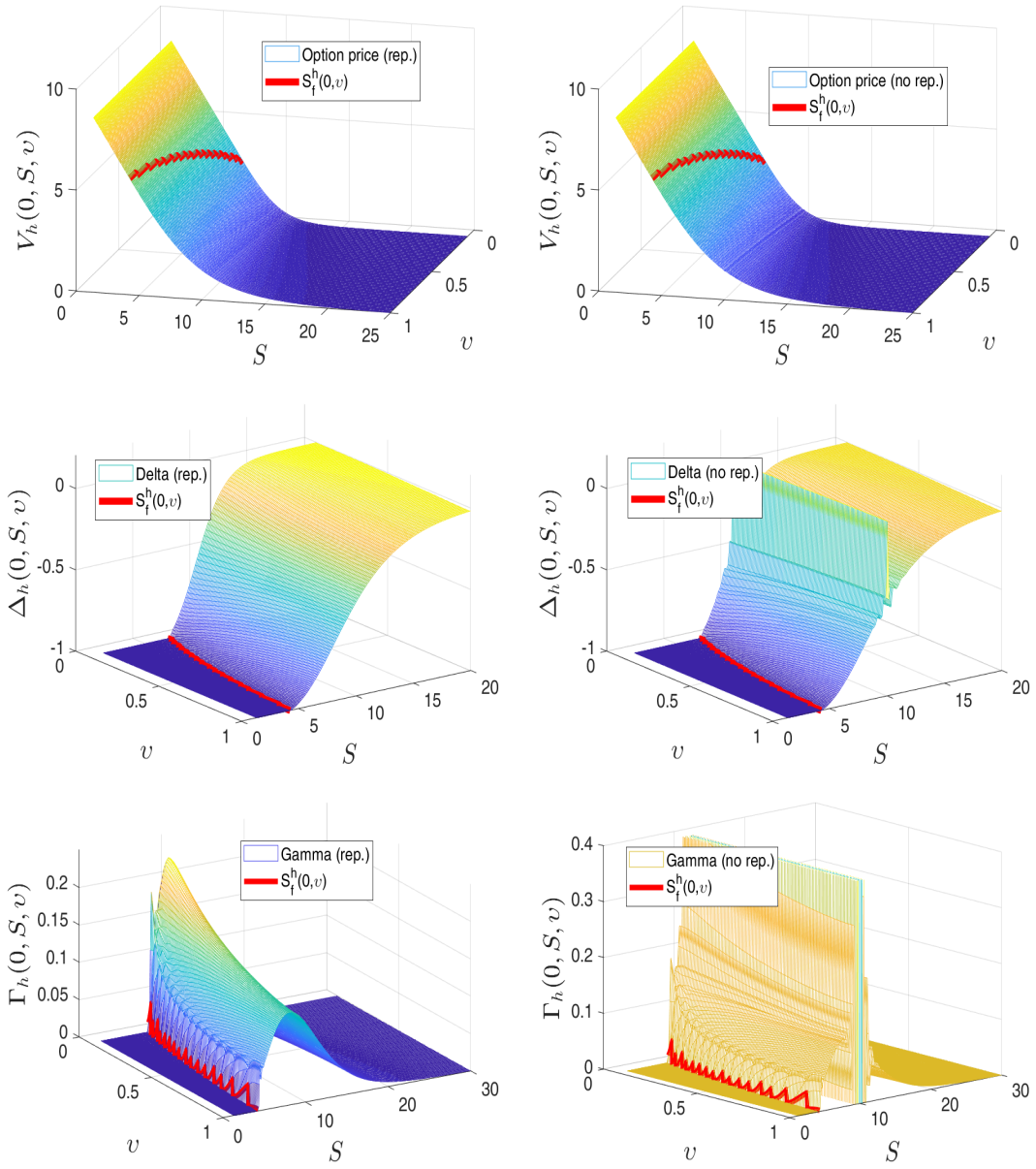

Figure 8.7.: Numerical approximation of today's American put option price in *Heston's model*, its first partial derivative *Delta* and its second partial derivative *Gamma* for cubic tensor product B-splines ($k = 4$) with repeating knots (rep.) at $S = K = 10$ and no repeating knots (no rep.) on level $\ell_x = 7$ and $\ell_v = 6$. The parameters for the Heston model are chosen as in Example 8.5.

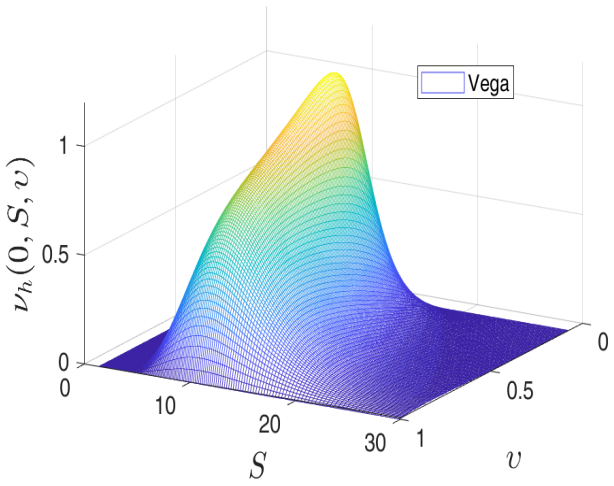

Figure 8.8.: Numerical approximation of Vega in *Heston's model*, the first partial derivative with respect to the variance v with cubic B-splines on level $\ell_x = 7$ and $\ell_v = 6$. The parameters for the Heston model are chosen as in Example 8.5.

Now, we study the spatial discretization error for American options in the H^1-norm at $t = 0$ as stated in (8.14) for the Black-Scholes or (8.17) for Heston's model. We distinguish between a cubic spline approximation with and without repeating knots at the locations where the partial derivative of the terminal condition are discontinuous. When determining the spatial discretization error, the number of time steps $\#\mathbb{T}$ are chosen sufficiently large to minimize the approximation error in the time variable, and calculations are performed using a sequence of grids with doubling the number of knots (degree of freedom) for the spatial B-spline discretization.
The corresponding results are depicted in Figure 8.9. As one can see, the error is significantly reduced for a cubic spline discretization with coinciding knots (rep.) in comparison to the discretization with no coinciding knots (no rep.). The convergence rates (≈ 1.5) for a spatial discretization with coinciding knots are in agreement with the theoretical results for elliptic variational inequalities. Thus, it can be reasonably assumed that the semi-discrete solution of the Black-Scholes variational inequality for a put or call option and of the Heston variational inequality for a put option satisfies the maximal smoothness for variational inequalities $W^{\mathfrak{s},\mathfrak{p}}(I)$ or $W^{\mathfrak{s},\mathfrak{p}}(\Omega)$ for all $\mathfrak{s} < 2 + \frac{1}{\mathfrak{p}}$ and $1 < \mathfrak{p} < \infty$ for which a theoretical result is still outstanding.
A remarkable result is that for the American butterfly-spread option with coinciding knots at $S = K_1, K, K_2$ in Figure 8.9 (below, left). Although the solution in $t = 0$ is not smooth enough at $S = K$, the error is of order ≈ 1.5. This is due to the fact that in $S = K$ the knots are repeated $k - 1$ times and this point is identical to a nodal point for finite elements. Thus, the cubic spline approximation is well defined and it is sufficient that the solution is smooth enough in $(0, K)$ and $(K, S_{\max})$. It should be also mentioned that a convergence rate of $\mathcal{O}(h^{1/2})$ is also obtained for a linear or quadratic

B-spline approximation, when $S = K$ are not aligned with $k - 1$ knot/s.

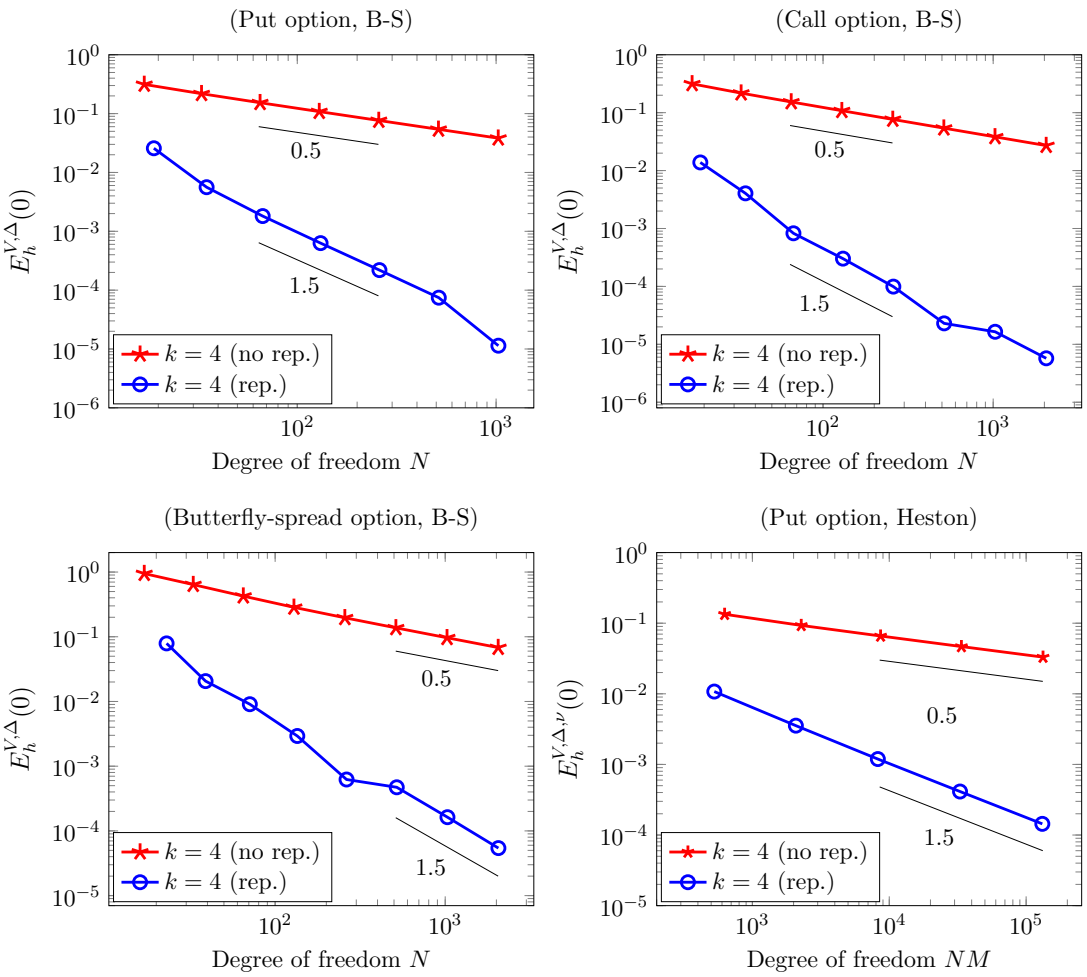

Figure 8.9.: Convergence of the error of today's American option price in the H^1-norm as stated in (8.14) or (8.17) for Example 8.2 (put option B-S, above left), Example 8.3 (call option B-S, above right), Example 8.4 (butterfly-spread option B-S, below left) and Example 8.5 (put option Heston, below right) with a cubic ($k = 4$) B-spline discretization with $k - 1$ repeating knots at $S = K$ or $S = K_1, K, K_2$ (rep.) and no repeating knots (no rep.) applying an implicit Euler method with fixed sufficiently large number of time steps $\#\mathbb{T}$.

8.2.3. Influence of the time discretization

The experiments in the next subsection are dedicated to clarify the temporal convergence of today's American option price and its first partial derivatives. In particular, we

concentrate on the convergence of the Euler, Crank-Nicholson and Rannachner time-stepping method for an American call option in the Black-Scholes model. To study the temporal discretization error, we consider Example 8.6 with a greater time interval $[0, T]$ with $T = 2$.
When determining the temporal discretization error, the number of knots in the spatial variable is chosen sufficiently large to minimize the spatial discretization error, and calculations are performed using a sequence of meshes with doubling the number of time steps $\#\mathbb{T}$. The corresponding results can be found in Figure 8.10. It can be seen that the Crank-Nicolson method for a smaller number of time steps converges with a rate of ≈ 0.5 and for a higher number of time steps it starts to converge with a rate of ≈ 1.2. For the Rannacher timestepping scheme the rate of convergence for a lower number of time steps is of order ≈ 1 and for a higher number of time steps of order ≈ 1.2.

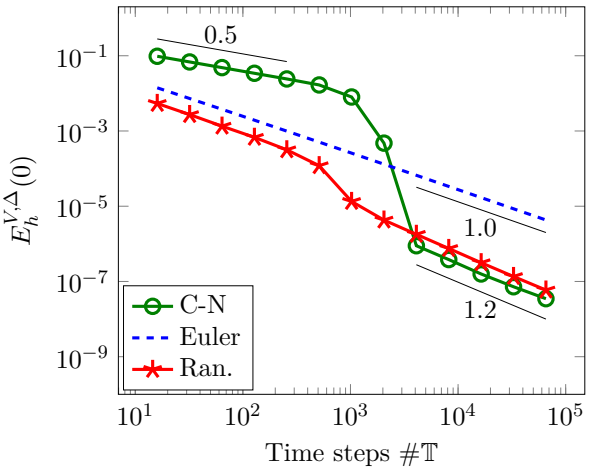

Figure 8.10.: $H^1(I)$-convergence of today's call option price ($t = 0$) for Example 8.6 (B-S) with a cubic ($k = 4$) B-spline discretization with $k - 1$ repeating knots at $S = K$ and fixed sufficiently large N applying different time stepping schemes with time steps $\#\mathbb{T}$.

The results for the Crank-Nicolson scheme and the Rannacher timestepping method seem to be confusing, see also Remark 4.7 for a detailed explanation. The reason for the poor convergence for a smaller number of time steps can be explained by the plots of the approximation of Delta in Figure 8.11 and Figure 8.12. The first top plots in Figure 8.11 present the approximations of the first partial derivative with respect to the underlying price, $\Delta_h(S, 0)$, for a Crank-Nicolson and Euler method. The plots show that in the approximation of Delta oscillations appear when using a Crank-Nicolson method with a low number of time steps $\#\mathbb{T}$ and the oscillations vanish for a very high number of time steps. Note, that the oscillations also occur at $S = K$, where the derivative of the terminal condition is discontinuous. Oscillations at $S = K$ for the Crank-Nicolson method can also be observed for simple partial differential equations with non-smooth data (see [GC06, FV02]).

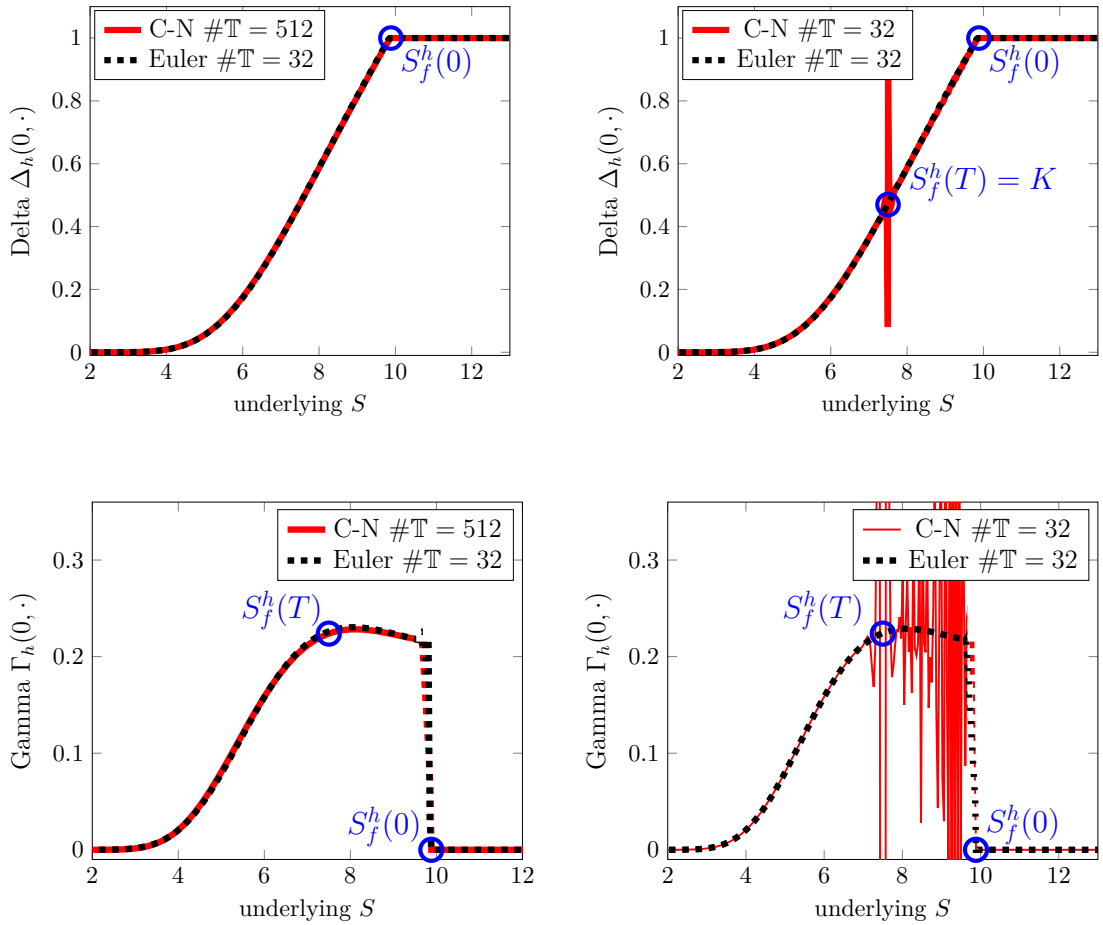

Figure 8.11.: Numerical approximation of the first derivative *Delta* of today's call option price and its second derivative *Gamma* for Example 8.6 (B-S) with cubic B-splines ($k = 4$), $k - 1$ repeating knots at $S = K$ and $N = 515$ applying the Crank-Nicolson (C-N) and Euler method (Euler).

An extensive analysis of the Crank-Nicolson and Rannacher method with a spatial finite difference discretization for the valuation of European options with the Black-Scholes partial differential equation can be found in [GC06]. To damp the errors in the initial data caused by the irregularity of the terminal condition, the authors recommend the Rannacher timestepping method for European options. Corresponding results for American options with the Rannacher method and a spatial cubic B-spline discretization can be found in Figure 8.12. As one can see the error for the approximation of Gamma for a small number of time steps at $S = K$ is significantly reduced but Gamma is still oscillating along $S_f^h(t)$ for $t < T$. Choosing a higher number of time steps for the Rannacher timestepping method (but lower as necessary for the Crank-Nicolson

method) eliminates the oscillations in the partial derivatives. To this end, this different behavior of the Crank-Nicolson or Rannacher method for small or higher number of time steps can explain the steep descent of the error in Figure 8.10.
The low order of the convergence rate for the Rannacher or Crank-Nicolson method in Figure 8.10 can be explained as follows: These methods are second order schemes and are appropriate for problems whose solution are Lipschitz continuous in time. Since solutions of parabolic variational inequalities do not satisfy this regularity assumption (see Remark 3.15), due to the free boundary, a second order convergence for parabolic variational inequalities cannot be expected.
Since the partial derivatives of the solution regarding S are of particular importance from the financial point of view and the implicit Euler method provides stable (non oscillating) results for the partial derivatives up to order two regardless of the number of time steps, all subsequent runs will use the implicit Euler method.

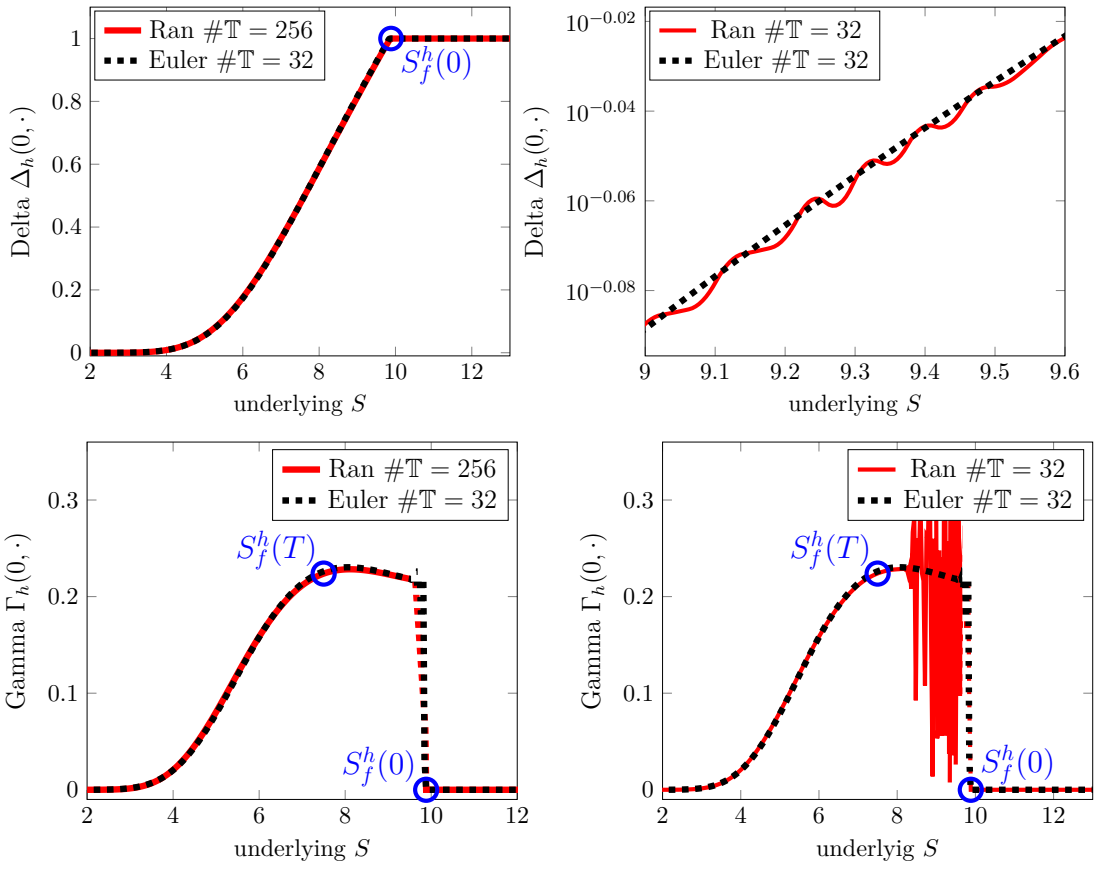

Figure 8.12.: Numerical approximation of the first derivative *Delta* of today's call option price and its second derivative *Gamma* for Example 8.6 (B-S) with cubic B-splines ($k = 4$), repeating knots at $S = K$ and $N = 515$ applying the Rannacher timestepping (Ran) and Euler method (Euler).

8.2.4. Influence of the B-spline order

The next experiments are dedicated to clarifying the convergence bahavior for the semi-discrete Black-Scholes or Heston variational inequality, when the order of B-spline basis functions k is increased. To study the spatial discretization error for today's American option price in the Black-Scholes or Heston's model, we proceed as described in Section 8.2.2. We choose the number of time steps sufficiently large to minimize the error in the time variable, and calculations are performed by doubling the number of knots for the spatial B-spline discretization on uniform grids with $k-1$ repeated knots at $S = K$ for a put or call option, or at $S = K_1, K, K_2$ for a butterfly-spread option.
The results for the discretization error in the $H^1(I)$-norm (or $H^1(\Omega)$-norm) are presented in Figure 8.13. From the financial point of view this error corresponds to the error of V_h and Δ_h in the Black-Scholes model and to the error of V_h, Δ_h and ν_h in Heston's model. As one can see, for all problems considered, the error for quadratic and cubic (tensor product) B-splines is significantly reduced in comparison to linear basis functions, but is not improved for $k > 3$.
It can be also observed that for the one-dimensional Black-Scholes variational inequality the error for quadratic and cubic basis functions varies slightly. This behavior can be explained as follows: In the one dimensional case the free boundary for fixed time is only a point and the discretization error strongly depends on the position of the knots next to the free boundary (see Section 8.1). Thus, the discretization error is reduced rapidly, if $|S_f^h - S_f^{\text{ref}}|$ is decreased. However, the convergence rates for linear basis functions are of order $\mathcal{O}(h)$ and $\mathcal{O}(h^{3/2})$ for quadratic and cubic basis functions. The approximated convergence rates are in agreement with the theoretical results for variational inequalities.
Now, we are interested in the error of Gamma for American options. As pointed out before, Gamma, the second derivative of today's American option price with respect to S, is discontinuous at the free boundary. Moreover, for a cubic B-spline discretization of the Black-Scholes variational inequality the approximation Γ_h is a piecewise linear function in S and the point of discontinuity in Gamma is approximated by a linear polynomial. Thus, in this neighbourhood the rate of convergence is of order $\mathcal{O}(h^{1/2})$. In fact, the approximation could be improved by using a knot series with $k-2$ repeating knots at the free boundary, but in practice this is difficult to realize, since the free boundary is a priori unknown.
Similar considerations can be applied for Gamma in Heston's model. Note, that due to the transformation $x = \log(S/K)$ for Heston's model the approximation of Gamma is a product of $(1/K\exp(x))^2$ and the first and second derivative of u_h (see (6.9)). It's clear that in this case the error of Γ_h in the L_2-norm is dominated by the error of the second derivative of u_h. However, Gamma is a smooth function except at the free boundary. Therefore, we consider the error of Γ_h in the L_2-norm ignoring the error in a neighbourhood of the approximated free boundary as stated in (8.16) for Black-Scholes model or (8.18) for Heston's model. Figure 8.14 contains the results of computation. As one can see, the error of Γ_h for cubic B-splines is significantly reduced in comparison to quadratic B-splines.

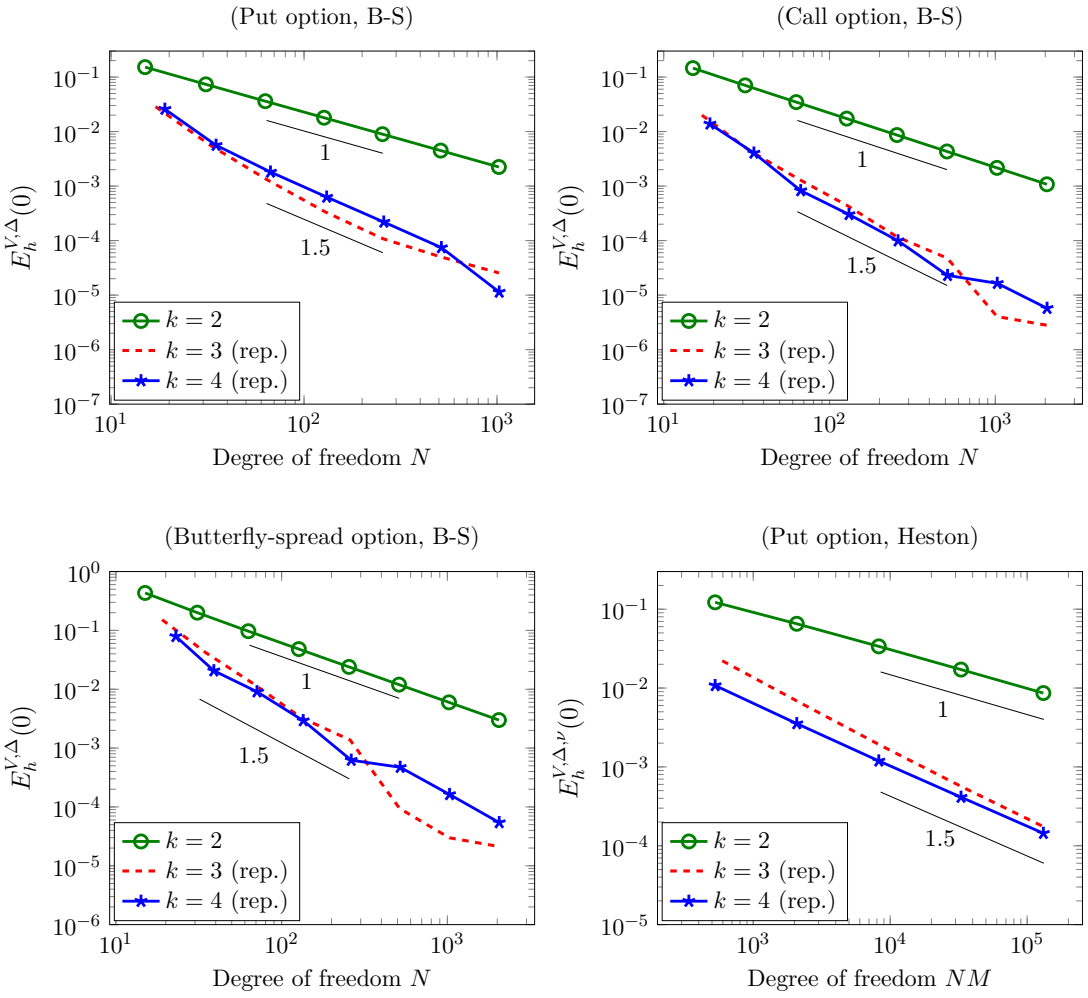

Figure 8.13.: H^1-convergence at $t = 0$ for Example 8.2 (put option (B-S), above left), Example 8.3 (call option (B-S), above right), Example 8.4 (butterfly-spread option (B-S), below left) and for Example 8.5 (put option (Heston), below right) with different orders of B-spline basis functions k applying an implicit Euler method for fixed time steps $\#\mathbb{T}$.

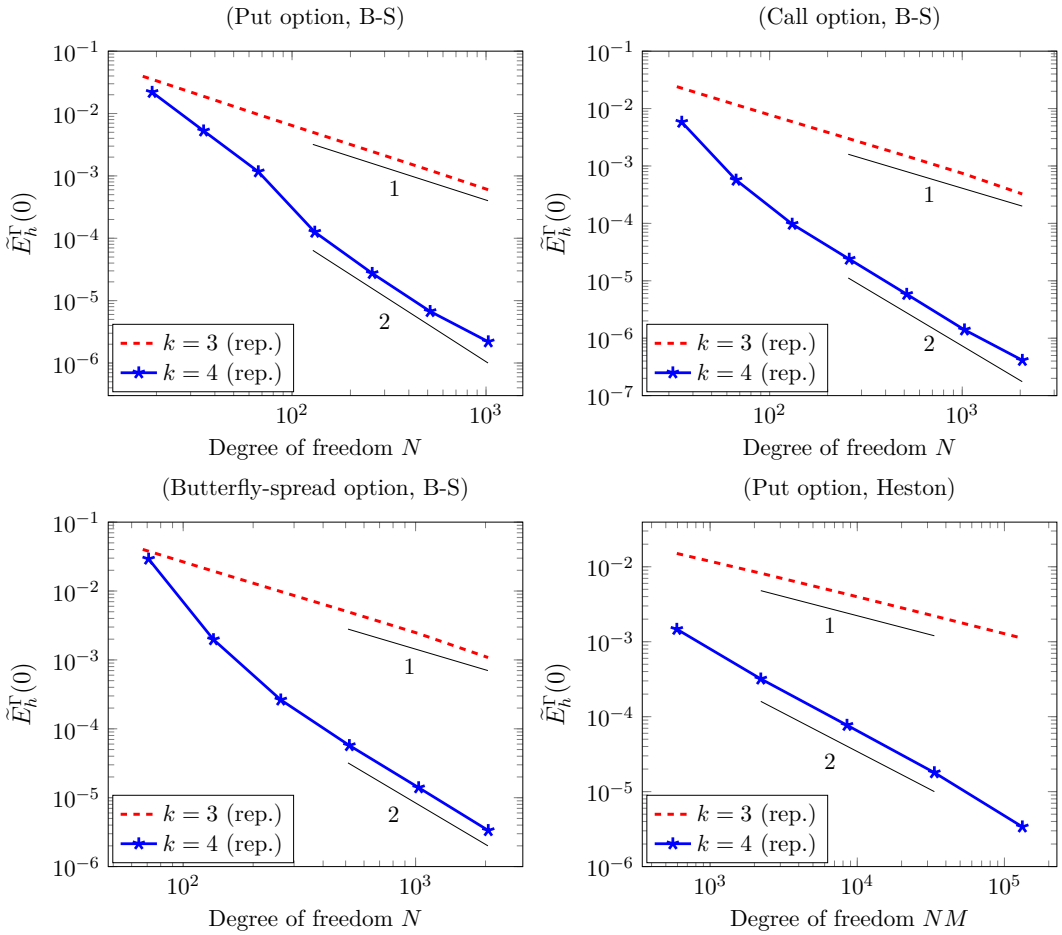

Figure 8.14.: Convergence of Gamma, the second derivative of today's American option price in the L^2-norm, ignoring the error in a neighbourhood of the free boundary as stated in (8.16) or (8.18) for Example 8.2 (put option B-S, above left), Example 8.3 (call option B-S, above right), Example 8.4 (butterfly-spread option B-S, below left) and Example 8.5 (put option Heston, below right) with different orders of B-splines k and $k-1$ repeating knots at $S = K$ or $S = K_1, K, K_2$ applying an implicit Euler method with fixed sufficiently large time steps $\#\mathbb{T}$.

8.2.5. Convergence of the monotone Multigrid method

The experiments in this subsection are dedicated to study the convergence behavior of the monotone Multigrid-method with respect to the refinement levels $\#\ell$, mesh size h, B-spline order k and number of a priori smoothing steps η_1. Therefore the stopping criteria in time step z is coupled on the spatial discretization error as follows

$$\|\mathbf{u}_*^{(z)} - \mathbf{u}_\varsigma^{(z)}\|_{\ell_2} \leq 10^{-6} h^{3/2} \quad \text{(B-S)} \quad \text{and} \quad \|\mathbf{u}_*^{(z)} - \mathbf{u}_\varsigma^{(z)}\|_{\ell_2} \leq 10^{-3} h^{3/2} \quad \text{(Heston)}, \tag{8.19}$$

where $\mathbf{u}_\varsigma^{(z)}$ denotes the iterate after ς iterations in time step z. Due to the log-transformation for Heston's problem the constant in the stopping criteria is chosen smaller as for the Black-Scholes Problem. In order to study the mesh dependence of the iterative methods we introduce the algebraic convergence rate for the linear complementarity problem (c.f. Problem for the B-S model and Problem for Heston's model) in one time step z as follows

$$\rho^{(z)} := \left(\frac{\|\mathbf{res}_{\varsigma_*}^{(z)}\|_{\ell_2}}{\|\mathbf{res}_{\varsigma_0}^{(z)}\|_{\ell_2}} \right)^{1/\varsigma_*}, \tag{8.20}$$

where $\mathbf{res}_{\varsigma_*}^{(z)}$ denotes the residual in one time step z after ς iterations of the MMG-method. Due to the projection step in the MMG- or PGS-method after each iteration, the constraint $\mathbf{u}^{(z)} \geq 0$ is maintained for all iteration steps, thus it is useful to consider the following residual

$$\mathbf{res}_\varsigma^{(z)} := \begin{cases} 0 & \text{if } u_i^{(z)} = 0, \\ C\mathbf{u}_\varsigma^{(z)} - \mathbf{f}^{(z-1)} & \text{if } u_i^{(z)} > 0, \end{cases} \tag{8.21}$$

see e.g. [Rei04]. In order to measure the algebraic convergence rate over all time steps, we calculate the averaged convergence rate

$$\overline{\rho} := \frac{\sum_{z=1}^{\#\mathbb{T}} \rho^{(z)}}{\#\mathbb{T}}. \tag{8.22}$$

Furthermore, the corresponding averaged iteration step per time step is specified as follows

$$\overline{\#\text{it}} := \frac{\sum_{z=1}^{\#\mathbb{T}} \#\text{it}^{(z)}}{\#\mathbb{T}}, \tag{8.23}$$

where $\#\text{it}^{(z)}$ denotes the number of iterations in one time step z.
For demonstrating the robustness in the number of levels, we fix the order of the B-spline basis function, the number of time steps $\#\mathbb{T}$, mesh size h and vary the number of refinement levels $\#\ell$. Table 8.4 contains the results for B-spline orders $k = 2, 3, 4$. We observe that the algebraic convergence rates and number of iterations are robust with respect to refinement levels $\#\ell$. The next experiments are dedicated to study the robustness of the MMG-method with respect to the mesh size. Therefore, we vary the mesh size h and fix the B-spline order k and time steps $\#\mathbb{T}$. Corresponding results for Heston's problem can be found in Table 8.3. The results show, that the algebraic

convergence rates and number of iterations steps for the MMG-method are robust with respect to the mesh size while the convergence rates for the PGS-method tend to 1. Moreover, we present results comparing the MMG-method to the PGS-method, where the mesh and time step size are chosen proportional to the discretization error. Corresponding results for the Black-Scholes and Heston's model with initial iterate obtained by Nested-iteration can be found in Table 8.5 and Table 8.6. It can be observed that a time step size $s = \mathcal{O}(h^{3/2})$ yields optimal convergence in the H^1-norm for the stopping criteria in (8.19). In comparison to the PGS-method the number of iterations and algebraic convergence rates are substantially reduced in the MMG-scheme. Table 8.6 also shows CPU-times in seconds required for the PGS-method in comparison to the MMG-method. For coarser grids the CPU-times for the PGS- and MMG-method are quite similar, while for finer grids the MMG-method is much faster than the PGS-method. Figure 8.15 and Figure 8.16 present some results for the MMG-method on a fixed level for polynomial order $k = 2, 3, 4$ when the number of a priori-smoothing steps is increased. The results show that the number of iterations inversely depends on the number of smoothing steps. Moreover, we see that for roughly up to 4 a priori-smoothing steps the CPU-times are significantly reduced.
Lastly, it should be noted from Table 8.3, Table 8.6 and Figure 8.16, that the number of iteration steps, the algebraic convergence rates and CPU-times increases with the B-spline order k. This behavior is clearly evident for the two dimensional Heston Problem.

Table 8.3.: $V(2,1)$-cycle convergence with initial iterate obtained by Nested-iteration applied to Example 8.5 with B-spline discretization of order $k = 2, 3, 4$, mesh sizes $h_x = 4/2^{\ell_x}$, $h_v = 0.925/2^{\ell_v}$ and an implicit Euler method with $\#\mathbb{T} = 128$ time steps.

$k = 2$

(ℓ_x, ℓ_v)	$(4,3)$		$(5,4)$		$(6,5)$		$(7,6)$		$(8,7)$		$(9,8)$	
method	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$
MMG	4	0.03	4	0.01	4	0.01	4	0.03	5	0.06	6	0.08
PGS	9	0.22	8	0.13	10	0.22	38	0.68	150	0.91	614	0.98

$k = 3$

(ℓ_x, ℓ_v)	$(4,3)$		$(5,4)$		$(6,5)$		$(7,6)$		$(8,7)$		$(9,8)$	
method	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$
MMG	7	0.18	7	0.17	6	0.11	6	0.07	7	0.08	8	0.09
PGS	15	0.43	17	0.43	18	0.41	22	0.50	74	0.81	282	0.95

$k = 4$

(ℓ_x, ℓ_v)	$(4,3)$		$(5,4)$		$(6,5)$		$(7,6)$		$(8,7)$		$(9,8)$	
method	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$	#it	$\overline{\rho}$
MMG	11	0.37	14	0.40	13	0.36	13	0.32	14	0.33	15	0.34
PGS	31	0.64	46	0.72	46	0.71	51	0.74	63	0.78	194	0.92

Table 8.4.: $V(3,1)$-cycle convergence with initial iterate on the obstacle with a linear ($k = 2$), quadratic ($k = 3$) and cubic ($k = 4$) B-spline discretization and an implicit Euler method.

Put option (B-S) with $h = 20/2^9$ and $\#\mathbb{T} = 1024$

refinement levels	$k = 2$		$k = 3$		$k = 4$	
$\#\ell$	$\overline{\#\text{it}}$	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\rho}$
2	4.0	$3e{-4}$	5.0	$8e{-4}$	9.4	0.038
3	4.0	$3e{-4}$	5.0	$8e{-4}$	9.4	0.038
4	4.0	$3e{-4}$	5.0	$8e{-4}$	9.4	0.038
5	4.0	$3e{-4}$	5.0	$8e{-4}$	9.4	0.038
6	4.0	$3e{-4}$	5.0	$8e{-4}$	9.4	0.038

Butterfly-spread option (B-S) with $h = 40/2^9$ and $\#\mathbb{T} = 1024$

refinement levels	$k = 2$		$k = 3$		$k = 4$	
$\#\ell$	$\overline{\#\text{it}}$	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\rho}$
2	4.0	$2e{-4}$	4.2	$2e{-4}$	5.3	0.001
3	4.0	$2e{-4}$	4.3	$2e{-4}$	5.3	0.001
4	4.0	$2e{-4}$	4.3	$2e{-4}$	5.3	0.001
5	4.0	$2e{-4}$	4.3	$2e{-4}$	5.3	0.001
6	4.0	$2e{-4}$	4.3	$2e{-4}$	5.3	0.001

Put option (Heston) with $h_x = 4/2^8$, $h_v = 0.925/2^7$, $\#\mathbb{T} = 64$

refinement levels	$k = 2$		$k = 3$		$k = 4$	
$\#\ell$	$\overline{\#\text{it}}$	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\rho}$
2	6	0.044	7	0.061	15	0.315
3	6	0.044	7	0.060	15	0.308
4	6	0.044	7	0.060	15	0.309
5	6	0.044	7	0.060	15	0.309
6	6	0.044	7	0.060	15	0.309

Table 8.5.: $V(1,1)$-cycle convergence with initial iterate obtained by Nested-iteration in comparison to the PGS-scheme applied to Example 8.2 (put option (B-S), above), Example 8.3 (call option (B-S), middle), Example 8.4 (butterfly-spread option (B-S), below) with cubic ($k = 4$) B-spline discretization and an implicit Euler method with $s = \mathcal{O}(h^{3/2})$.

Put option (B-S)

$h = 20/2^\ell$		PGS		MMG			
ℓ	$\#\mathbb{T}$	$\overline{\#\text{it}}$	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\rho}$	$E_h^{V,\Delta}(0)$	rate
6	16	33	0.56	5	0.024	$1.6e{-2}$	–
7	46	49	0.65	4	0.015	$5.6e{-3}$	1.48
8	128	75	0.73	3	0.008	$2.1e{-3}$	1.45
9	363	115	0.79	3	0.005	$7.4e{-4}$	1.48
10	1,024	179	0.85	3	0.004	$2.6e{-4}$	1.49
11	2,897	277	0.89	3	0.004	$9.3e{-5}$	1.50
12	8,192	428	0.92	3	0.004	$3.2e{-5}$	1.53

Call option (B-S)

$h = 20/2^\ell$		PGS		MMG			
ℓ	$\#\mathbb{T}$	$\overline{\#\text{it}}$	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\rho}$	$E_h^{V,\Delta}(0)$	rate
6	16	26	0.54	6	0.03	$1.6e{-2}$	–
7	46	35	0.61	5	0.04	$5.9e{-3}$	1.43
8	128	47	0.68	6	0.07	$2.2e{-3}$	1.41
9	363	62	0.73	6	0.12	$8.2e{-4}$	1.45
10	1,024	82	0.78	7	0.17	$3.0e{-4}$	1.46
11	2,897	109	0.82	8	0.22	$1.1e{-4}$	1.48
12	8,192	146	0.86	9	0.27	$3.8e{-5}$	1.49

Butterfly-spread option (B-S)

$h = 40/2^\ell$		PGS		MMG			
ℓ	$\#\mathbb{T}$	$\overline{\#\text{it}}$	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\rho}$	$E_h^{V,\Delta}(0)$	rate
6	16	26	0.55	8	0.10	$5.5e{-2}$	–
7	46	35	0.62	8	0.12	$1.9e{-2}$	1.51
8	128	48	0.69	8	0.11	$7.0e{-3}$	1.47
9	363	69	0.75	6	0.07	$2.5e{-3}$	1.48
10	1,024	101	0.80	5	0.04	$8.9e{-4}$	1.50
11	2,897	149	0.85	4	0.02	$3.1e{-4}$	1.51
12	8,192	222	0.88	4	0.01	$1.1e{-4}$	1.55

Table 8.6.: $V(4,1)$-cycle convergence with initial iterate obtained by Nested-iteration in comparison to the PGS-scheme applied to Example 8.5 (put option, Heston), with B-spline discretization of order $k = 2, 3, 4$ and an implicit Euler method with $s = \mathcal{O}(h^{3/2})$.

$k = 2$

$h_x := \frac{4}{2^{\ell_x}}$, $h_v := \frac{0.925}{2^{\ell_v}}$			PGS			MMG				
ℓ_x	ℓ_v	$\#\mathbb{T}$	$\overline{\#\text{it}}$	$\overline{\text{cpu}}$ (s)	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\text{cpu}}$ (s)	$\overline{\rho}$	$E_h^{V,\Delta,\nu}(0)$	rate
5	4	3	32	0.02	0.70	3.3	0.003	0.01	1.9e−1	–
6	5	8	77	0.01	0.85	4.0	0.004	0.01	8.9e−2	1.06
7	6	23	150	0.03	0.92	4.1	0.022	0.02	4.1e−2	1.13
8	7	64	271	0.27	0.95	4.1	0.115	0.02	1.9e−2	1.09
9	8	182	450	1.72	0.97	4.1	0.226	0.02	9.2e−3	1.06
10	9	512	727	11.25	0.98	5.0	1.119	0.03	4.5e−3	1.04

$k = 3$

$h_x := \frac{4}{2^{\ell_x}}$, $h_v := \frac{0.925}{2^{\ell_v}}$			PGS			MMG				
ℓ_x	ℓ_v	$\#\mathbb{T}$	$\overline{\#\text{it}}$	$\overline{\text{cpu}}$ (s)	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\text{cpu}}$ (s)	$\overline{\rho}$	$E_h^{V,\Delta,\nu}(0)$	rate
5	4	3	17	0.00	0.46	4.0	0.01	0.02	1.3e−1	–
6	5	8	38	0.01	0.71	4.0	0.01	0.02	5.5e−2	1.28
7	6	23	73	0.03	0.83	4.0	0.03	0.02	2.0e−2	1.45
8	7	64	129	0.18	0.89	5.0	0.15	0.02	7.4e−3	1.44
9	8	182	209	1.17	0.93	5.0	0.44	0.02	2.7e−3	1.48
10	9	512	331	7.43	0.95	5.9	2.06	0.02	9.6e−4	1.47

$k = 4$

$h_x := \frac{4}{2^{\ell_x}}$, $h_v := \frac{0.925}{2^{\ell_v}}$			PGS			MMG				
ℓ_x	ℓ_v	$\#\mathbb{T}$	$\overline{\#\text{it}}$	$\overline{\text{cpu}}$ (s)	$\overline{\rho}$	$\overline{\#\text{it}}$	$\overline{\text{cpu}}$ (s)	$\overline{\rho}$	$E_h^{V,\Delta,\nu}(0)$	rate
5	4	3	35	0.00	0.70	6.3	0.03	0.15	1.2e−1	–
6	5	8	46	0.01	0.75	6.8	0.04	0.14	4.8e−2	1.27
7	6	23	58	0.04	0.79	8.1	0.10	0.16	1.8e−2	1.45
8	7	64	95	0.26	0.86	9.3	0.41	0.17	6.5e−3	1.43
9	8	182	148	1.54	0.90	10.5	1.32	0.18	2.3e−3	1.48
10	9	512	225	8.67	0.93	11.6	5.70	0.19	8.4e−4	1.48
11	10	1,449	336	50.68	0.95	12.7	24.77	0.20	–	–

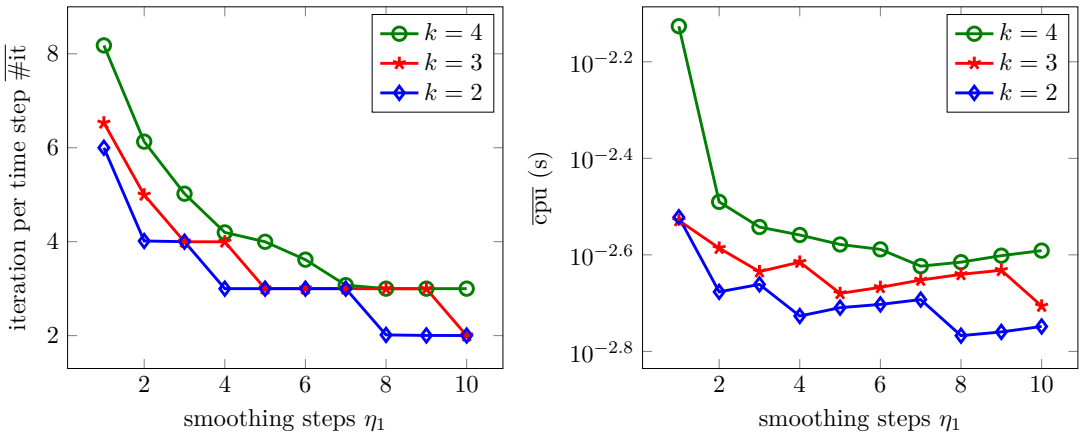

Figure 8.15.: $V(\eta_1, 1)$-cycle convergence with initial iterate obtained by Nested-iteration applied to Example 8.4 (butterfly-spread option, B-S), with B-spline discretization of order k on level $\ell = 9$ and an implicit Euler method with $\mathbb{T} = 1024$.

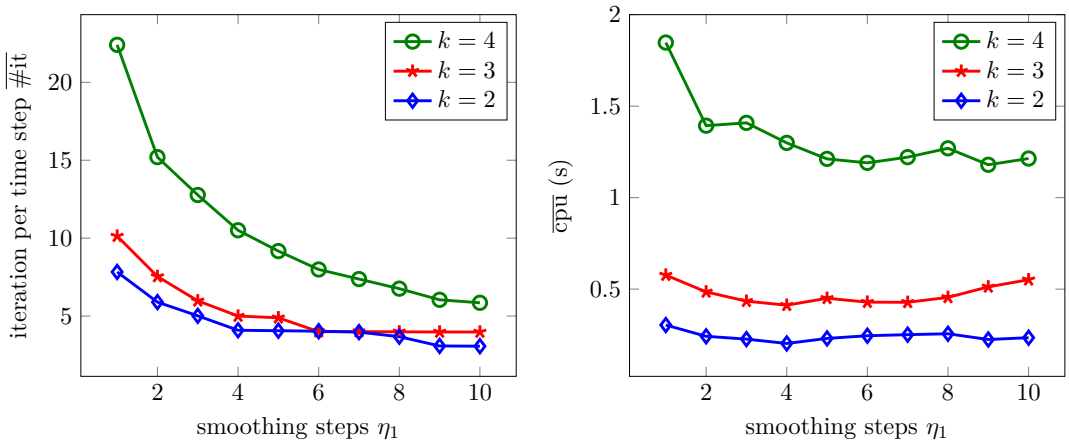

Figure 8.16.: $V(\eta_1, 1)$-cycle convergence with initial iterate obtained by Nested-iteration applied to Example 8.5 (put option, Heston) with B-spline discretization of order k on level $\ell_x = 9, \ell_v = 8$ and an implicit Euler method with $\#\mathbb{T} = 182$.

9. Pricing European option with the Black-Scholes-Barenblatt equation

9.1. Uncertain volatility model

In this chapter we focus on the non-linear Black-Scholes model, which was derived independently by Lyons and Avellaneda et al. in 1995 [ALP95, Lyo95]. In comparison with the Heston model where the volatility can explain the volatility smile or skewness effects observed from option implied volatilities in the Black-Scholes model (see Section 2.1.3), the volatility in the uncertain volatility model is assumed to lie within a range of extreme values. With these extreme values the option price can be computed in a worst or best case scenario. As pointed out by [ALP95] the bounds can be inferred from high-low peaks of implied volatilities and be considered as defining a confidence interval for future volatilities.
Since also in the simple case of an European option with a non-convex payoff function no closed form solution for the Black-Scholes-Barenblatt equation is available, we concentrate on a European option. In particular we consider a European butterfly-spread option, since for this type of option the volatilities varies significantly. Note that all analytical or numerical considerations of the Black-Scholes-Barenblatt equation in this chapter can be also applied to other once weakly differentiable payoff functions in a similar manner. Restricting our considerations to a butterfly-spread option only offers a more simple notation. In the following we denote by $\underline{\sigma} \in \Sigma$ the volatility in the worst case scenario where an infimum is taken and by $\overline{\sigma} \in \Sigma$ the volatility in the best case scenario where a supremum is taken. Recall the non-linear Black-Scholes equation from (2.34) and (2.35), then the valuation of a European butterfly-spread option in the worst- and best case scenario can be formulated as follows:

Problem 9.1. *(Black-Scholes-Barenblatt equation – worst case)*
Find $V(t,S)$ *such that*

$$V_t + \inf_{\underline{\sigma}\in\Sigma} (\mathcal{L}^{\underline{\sigma}} V) := V_t + \inf_{\underline{\sigma}\in\Sigma} \left(\frac{1}{2} S^2 \underline{\sigma}^2 V_{SS} \right) + (r - D_0) S V_S - rV = 0 \quad \text{in } [0,T) \times \mathbb{R}^+ \tag{9.1}$$

with $\Sigma := [\sigma_{\min}, \sigma_{\max}]$ *and boundary conditions*

$$\lim_{S\to 0} V(t,S) = \lim_{S\to 0} \mathcal{H}_{BS}(S), \quad \lim_{S\to\infty} V(t,S) = \lim_{S\to\infty} \mathcal{H}_{BS}(S)$$

and end condition $V(T,S) = \mathcal{H}_{BS}(S)$.

Problem 9.2. *(Black-Scholes-Barenblatt equation – best case)*
Find $V(t,S)$ such that

$$V_t + \sup_{\overline{\sigma}\in\Sigma}\left(\mathcal{L}^{\overline{\sigma}}V\right) := V_t + \sup_{\overline{\sigma}\in\Sigma}\left(\frac{1}{2}\overline{\sigma}^2 S^2 V_{SS}\right) + (r-D_0)SV_S - rV = 0 \quad \text{in } [0,T)\times\mathbb{R}^+ \tag{9.2}$$

with $\Sigma := [\sigma_{\min}, \sigma_{\max}]$ and boundary conditions

$$\lim_{S\to 0} V(t,S) = \lim_{S\to 0} \mathcal{H}_{BS}(S) = 0, \quad \lim_{S\to\infty} V(t,S) = \lim_{S\to\infty} \mathcal{H}_{BS}(S) = 0$$

and end condition $V(T,S) = \mathcal{H}_{BS}(S)$.

This model studies uncertain volatility models, in which the volatility σ is not known precisely but is assumed to lie between extreme values $\sigma_{\min}$ and $\sigma_{\max}$.
The Black-Scholes-Barenblatt equation is a special type of a Hamilton-Jacobi Bellman equation. One advantage of this special setting is that the infimum or supremum can be expressed analytically as pointed out in (2.37) and (2.38) (otherwise minimizing procedures should be considered). Considering the non-linear spatial term in (9.1) or (9.2) one of the main difficulties in comparison to linear elliptic partial differential equations arises from the fact that the spatial term of the non-linear Black-Scholes-Barenblatt equation does not admit a weak $H^1(I)$-formulation in the continuous setting. This is due to the fact that the non-linearity or analytic expression of the volatility function depends on the pointwise evaluation of the a-priori unknown second derivative of the option price. Hence no integration by parts is possible to pass the partial derivative onto a test function. This fact is nothing new and has already been established in [JS13] or [Sme15, Chapter 1.2] for a Hamilton-Jacobi-Bellman equation in a more general setting. To apply some known analysis from [Sme15] in the next section, it is convenient to introduce a time transformation $\tau := T - t$. Moreover, we truncate the spatial domain $\mathbb{R}^+$ to a bounded domain $I := (0, S_{\max})$, where $S_{\max}$ is chosen large enough such that the boundary conditions are satisfied. Hence the *time transformed option price*

$$y(\tau, S) := V(T-t, S)$$

for a butterfly-spread option satisfies the following *time transformed Black-Scholes-Barenblatt equation* (BSB equation) in the worst case scenario

$$\frac{\partial y}{\partial \tau} - \inf_{\underline{\sigma}\in\Sigma}\left(\mathcal{L}^{\underline{\sigma}} y\right) = 0 \qquad \text{in } (0,T]\times I \tag{9.3}$$

and in the best case scenario with

$$\frac{\partial y}{\partial \tau} - \sup_{\overline{\sigma}\in\Sigma}\left(\mathcal{L}^{\overline{\sigma}} y\right) = 0 \qquad \text{in } (0,T]\times I \tag{9.4}$$

with initial condition $y(0,S) = \mathcal{H}_{\mathrm{BS}}(S)$ and boundary conditions $y(\tau,0) = 0$ and $y(\tau, S_{\max}) = 0$.

9.2. Well-posedness

To prove well-posedness for classical elliptic or parabolic equations in Sobolev spaces usually a weak formulation is introduced and the well-posedness of the solution in $H_0^1(I)$ is then derived by showing that the bilinear form is bounded and coercive. As pointed out before the non-linear spatial term of the BSB equation does not admit a weak $H^1(I)$-formulation in general, thus the classical theory for Sobolev spaces is not applicable.
A well-known theory for non-linear partial differential equations is that of a *viscosity solution*. The uniqueness, existence and smoothness for viscosity solutions in the special case of the Black-Scholes-Barenblatt equation is analyzed in [Var01]. The author has shown, that if the payoff is locally Lipschitz, then the BSB equation has a unique solution with the first derivative defined almost everywhere. Moreover, he has proven that the solution has also Hölder continuous second derivatives for the log-transformed BSB equation.
In this thesis, we prove the well-posedness of the BSB equation by a novel theory, developed for elliptic and parabolic Hamilton-Jacobi-Bellman equations in the PhD thesis [Sme15] or in the later published article [SS16]. For this theory the notion of viscosity solution is not required. The authors prove well-posedness for a HJB equation in a second order Sobolev space or Bochner space within a second order Sobolev space assuming that the coefficients in the differential operator satisfy the Cordes condition. The Cordes condition arises from the literature on nondivergence form PDE, i.e. $a : D^2u$ with $a \in L^\infty(\Omega)^{d\times d}$, where d denotes the space dimension. Although the theory for HJB equations with Cordes coefficients has been introduced for higher space dimensions $d \geq 2$, it is straightforward to employ the theory for the one dimensional BSB equation. Thus, all results in this subsection are simple applications of the theory developed in [Sme15, Chapter 4.1].
But, since this theory has been developed for non-degenerate HJB equations on bounded domains, we introduce a log-transformation $x = \log(S/K)$ for the BSB equation, and the spatial domain $I := (0, S_{\max})$ is truncated to $\tilde{I} := (S_{\min}, S_{\max})$ in S-coordinates or $I_x := (x_{\min}, x_{\max})$ in x-coordinates. In order to avoid some significant truncation effects on the solution, the parameter $x_{\min}$ should be chosen small enough such that $-\infty < x_{\min} < 0 < x_{\max} < \infty$ is satisfied. To avoid repetitions we concentrate on the BSB equation in the best case scenario. Thus, by setting $y(\tau, S) := \tilde{y}(\tau, x)$ in (9.4) the corresponding problem in log-coordinates x reads as follows: Find $\tilde{y}(\tau, x)$ with $(\tau, x) \in (0, T] \times I_x$ such that

$$F[\tilde{y}] := \frac{\partial \tilde{y}}{\partial \tau} - \sup_{\overline{\sigma}} \left(\tilde{\mathcal{L}}^{\overline{\sigma}} \tilde{y} \right) = 0 \qquad \text{in } (0, T] \times I_x, \tag{9.5}$$

with the log-transformed operator $\tilde{\mathcal{L}}^{\overline{\sigma}} : H^2(I_x) \to L^2(I_x)$ defined as

$$\tilde{\mathcal{L}}^{\overline{\sigma}} \tilde{y} := \frac{1}{2}\overline{\sigma}^2 \tilde{y}_{xx} + \left(r - D_0 - \frac{1}{2}\overline{\sigma}^2 \right) \tilde{y}_x - r\tilde{y}. \tag{9.6}$$

The boundary condition on the degenerate boundary is set to $\tilde{y}(\tau, x_{\min}) = 0$ and the boundary condition for $S_{\max}$ transforms into $\tilde{y}(\tau, x_{\max}) = 0$. Moreover, the initial condition is given by $\tilde{y}(0, x) := \mathcal{H}_{\mathrm{BS}}(K \exp(x))$.
Let us define the Sobolev-space

$$\mathcal{W} := H^2(I_x) \cap H_0^1(I_x). \tag{9.7}$$

Note that the fully non-linear operator $F[\tilde{y}]$ is well defined as a mapping from the Bochner space $H(0,T;I_x)$ into the space $L^2(0,T;I_x)$, where $H(0,T;I_x)$ is defined as

$$H(0,T;I_x) := L^2(0,T;\mathcal{W}) \cap H^1(I_x; L^2(I_x)). \tag{9.8}$$

For $S_{\min} > 0$ we can see that $\tilde{y}(\tau, x)$ is a solution of (9.5) if and only if $y(\tau, S)$ is a solution of (9.4). Although we assume that the initial data are given by the payoff function of a butterfly-spread option, it should be clear that the results in this subsection can also be applied for any initial data given by a non-convex function in $H_0^1(I_x)$. As mentioned earlier for convex initial data the BSB equation reduces to the linear BS-equation with one of the extreme values $\sigma_{\min}$ or $\sigma_{\max}$. Thus, for this special case, standard theory for linear parabolic variational equations can be applied. For the analysis, the volatility is assumed to be strictly positive and bounded and the interest rate is assumed to be strictly positive, i.e.

$$0 < \sigma_{\min} \leq \overline{\sigma} \leq \sigma_{\max} < \infty \qquad \text{and} \qquad r > 0. \tag{9.9}$$

Moreover, we assume that the coefficients $\overline{\sigma}$, r and D_0 satisfy the *Cordes condition*, i.e. there exists $\lambda > 0$, $\omega > 0$ and $\varepsilon \in (0, 1]$ such that

$$\frac{\frac{1}{4}\overline{\sigma}^4 + |r - D_0 - \frac{1}{2}\overline{\sigma}^2|^2/(2\lambda) + (r/\lambda)^2 + (1/\omega)^2}{(\frac{1}{2}\overline{\sigma}^2 + r/\lambda + 1/\omega)^2} \leq \frac{1}{2+\varepsilon} \quad \text{for each } \overline{\sigma} \in \Sigma. \tag{9.10}$$

Let us further define a strictly positive function $\gamma : \Sigma \to \mathbb{R}_{>0}$ by

$$\gamma(\overline{\sigma}) := \frac{\frac{1}{2}\overline{\sigma}^2 + r/\lambda + 1/\omega}{\frac{1}{4}\overline{\sigma}^4 + |r - D_0 - \frac{1}{2}\overline{\sigma}^2|^2/(2\lambda) + (r/\lambda)^2 + (1/\omega)^2}. \tag{9.11}$$

For each $\overline{\sigma} \in \Sigma$ define $\gamma^{\overline{\sigma}} : (x, \tau) \mapsto \gamma(\overline{\sigma})$. Moreover, for the analysis we introduce the operator $F_\gamma : H(0,T;I_x) \to L^2(0,T;L^2(I_x))$ defined by

$$F_\gamma[\varphi] := \inf_{\overline{\sigma} \in \Sigma} \left(\gamma^{\overline{\sigma}} (\partial_\tau \varphi - \tilde{\mathcal{L}}^{\overline{\sigma}} \varphi) \right). \tag{9.12}$$

Now, the remainder of this subsection is devoted to prove that there exists a *unique strong solution* $\tilde{y}(\tau, x) \in H(0, T; I_x)$ to the following problem:
Find $\tilde{y}(\tau, x)$ with $(\tau, x) \in (0, T] \times I_x$ such that

$$F_\gamma[\tilde{y}] = 0 \tag{9.13}$$

with boundary conditions $\tilde{y}(\tau, x_{\min}) = 0$ and $\tilde{y}(\tau, x_{\max}) = 0$; and initial condition $y(0, x) = \mathcal{H}_{\mathrm{BS}}(K \exp(x))$.
Then, by proving that the solution of (9.13) is also the unique solution of (9.5) it can be concluded that there exists a unique strong solution $\tilde{y}(\tau, x) \in H(0, T; I_x)$ to the log-transformed BSB equation.
For $\omega > 0$ and $\lambda > 0$ we introduce the operators

$$L_\lambda \varphi := \varphi_{xx} - \lambda \varphi, \qquad L_\omega \varphi := \omega \partial_\tau - L_\lambda \varphi. \tag{9.14}$$

From [Sme15, Lemma 4.2] it is known that $\mathcal{W} \subset H_0^1(I_x) \subset L^2(I_x)$ is a Gelfand triple under the inner product $\langle \cdot, \cdot \rangle_{H_0^1}$ and the duality pairing

$$\langle f, \varphi \rangle_{L^2 \times \mathcal{W}} := \int_{I_x} f(-L_\lambda \varphi) \ \text{ for } f \in L^2(I_x) \mathrm{d}x, \ \varphi \in \mathcal{W}. \tag{9.15}$$

Moreover, the space $H(0, T; I_x)$ is continuously embedded in $C([0, T]; H_0^1(I_x))$. Let us further define the norm on $H(0, T, I_x)$ by

$$\|\varphi\|^2_{H(0,T,I_x)} := \int_0^T \omega^2 \|\partial_\tau \varphi\|^2_{L^2(I_x)} + |\varphi|^2_{H^2(I_x),\lambda} \mathrm{d}\tau \tag{9.16}$$

with

$$|\varphi|_{H^2(I_x),\lambda} := |\varphi|^2_{H^2(I_x)} + 2\lambda |\varphi|^2_{H^1(I_x)} + \lambda^2 \|\varphi\|^2_{L^2(I_x)}, \quad \omega, \lambda > 0. \tag{9.17}$$

The following theorem is strongly based on [Sme15, Lemma 4.1]. In particular, we provide an analogous proof for the log-transformed BSB equation.

Lemma 9.3. *Suppose that the data $\overline{\sigma} \in \Sigma$, r and D_0 satisfy (9.9) and the Cordes condition in (9.10) with $\lambda > 0$ and $\omega > 0$. Let $U \subset I_x$ and $J \subset (0, T)$ be open intervals, and let the function $\tilde{y}, \varphi \in L^2(J; H^2(U)) \cap H^1(J; L^2(U))$; and set $w := \tilde{y} - \varphi$. Then the following inequality holds a.e. in U, for a.e. $\tau \in J$*

$$|F_\gamma[\tilde{y}] - F_\gamma[\varphi] - L_\omega w| \leq \sqrt{1 - \varepsilon} \sqrt{\omega^2 |\partial_t w|^2 + |w_{xx}|^2 + 2\lambda |w_x|^2 + \lambda^2 |w|^2} \tag{9.18}$$

with the non-linear operator for the log-transformed BSB equation defined in (9.12).

Proof. The definition of $F_\gamma[\tilde{y}]$ in (9.12) and L_ω in (9.14) implies the identity

$$\begin{aligned} F_\gamma[\tilde{y}] - L_\omega \tilde{y} &= \inf_{\overline{\sigma} \in \Sigma} \left(\gamma^{\overline{\sigma}} (\partial_\tau \tilde{y} - \tilde{\mathcal{L}}^{\overline{\sigma}} \tilde{y}) - L_\omega \tilde{y} \right) \\ &= - \sup_{\overline{\sigma} \in \Sigma} \left(\gamma^{\overline{\sigma}} (\tilde{\mathcal{L}}^{\overline{\sigma}} \tilde{y} - \partial_\tau \tilde{y}) + L_\omega \tilde{y} \right). \end{aligned}$$

Now, $|\sup_{\overline{\sigma}}(x^{\overline{\sigma}}) - \sup_{\overline{\sigma}}(y^{\overline{\sigma}})| \leq \sup_{\overline{\sigma}} |y^{\overline{\sigma}} - x^{\overline{\sigma}}|$ leads to

$$
\begin{aligned}
&|F_\gamma[\tilde{y}] - F_\gamma[\varphi] - L_\omega w| \leq \sup_{\overline{\sigma}\in\Sigma} |\gamma^{\overline{\sigma}}(\tilde{\mathcal{L}}w - \partial_\tau w + L_\omega w)| \\
&\leq \sup_{\overline{\sigma}\in\Sigma} \left(\left| \frac{1}{2}\gamma^{\overline{\sigma}}\overline{\sigma}^2 - 1 \right| |w_{xx}| + |\gamma^{\overline{\sigma}}| \left| r - D_0 - \frac{1}{2}\overline{\sigma}^2 \right| |w_x| + |\lambda - \gamma^{\overline{\sigma}} r||w| + |\omega - \gamma^{\overline{\sigma}}||\partial_\tau w| \right).
\end{aligned}
$$

Applying Cauchy-Schwarz inequality results in

$$
|F_\gamma[\tilde{y}] - F_\gamma[\varphi] - L_\omega w| \leq \sup_{\overline{\sigma}\in\Sigma} \left(\sqrt{C^{\overline{\sigma}}} \right) \sqrt{\omega^2 |\partial_t w|^2 + |w_{xx}|^2 + 2\lambda |w_x|^2 + \lambda^2 |w|^2} \tag{9.19}
$$

with

$$
C^{\overline{\sigma}} := \left| \frac{1}{2}\gamma^{\overline{\sigma}}\overline{\sigma}^2 - 1 \right|^2 + \frac{\left| r - D_0 - \frac{1}{2}\overline{\sigma}^2 \right|^2}{2\lambda} + \frac{|\lambda - \gamma^{\overline{\sigma}} r|^2}{\lambda^2} + \frac{|\omega - \gamma^{\overline{\sigma}}|^2}{\omega^2}. \tag{9.20}
$$

Finally, expanding the terms in $C^{\overline{\sigma}}$ and the definition of $\gamma^{\overline{\sigma}}$ in (9.11) implies

$$
\begin{aligned}
C^{\overline{\sigma}} &= 3 - 2\gamma^{\overline{\sigma}} \left(\frac{1}{2}\overline{\sigma}^2 + \frac{r}{\lambda} + \frac{1}{\omega} \right) + (\gamma^{\overline{\sigma}})^2 \left(\frac{1}{4}\overline{\sigma}^4 + \frac{|r - D_0 - \frac{1}{2}\overline{\sigma}^2|^2}{2\lambda} + \frac{r^2}{\lambda^2} + \frac{1}{\omega^2} \right) \\
&= 3 - 2 \left(\frac{(\frac{1}{2}\overline{\sigma}^2 + r/\lambda + 1/\omega)^2}{\frac{1}{4}\overline{\sigma}^4 + |r - D_0 - \frac{1}{2}\overline{\sigma}^2|^2/(2\lambda) + (r/\lambda)^2 + (1/\omega)^2} \right) \\
&\quad + \left(\frac{(\frac{1}{2}\overline{\sigma}^2 + r/\lambda + 1/\omega)^2}{\frac{1}{4}\overline{\sigma}^4 + |r - D_0 - \frac{1}{2}\overline{\sigma}^2|^2/(2\lambda) + (r/\lambda)^2 + (1/\omega)^2} \right),
\end{aligned}
$$

which proves $C^{\overline{\sigma}} \leq 1 - \varepsilon$ by applying the cordes condition in (9.10). □

Finally, by the same proof as in [Sme15, Theorem 4.4] one can prove the existence and uniqueness of a strong solution for the log-transformed BSB equation. For a better understanding we provide a short sketch of this proof. Note, that the author considers the more general case where the coefficients are assumed to be continuous functions depending on (x, τ), while for the log-transformed BSB equation the coefficients $\overline{\sigma}, r, D_0$ are constant.

Theorem 9.4. *(Strong solution for the log-transformed BSB equation)*
Let the data $\overline{\sigma} \in \Sigma$, r and D_0 satisfy (9.9) and the cordes condition in (9.10) with $\lambda > 0$ and $\omega > 0$. Then there exists a unique strong solution $\tilde{y}(\tau, x) \in H(0, T; I_x)$ of the log-transformed BSB equation in (9.5) with boundary conditions $\tilde{y}(\tau, x) = 0$ for $x \in \partial I_x$ and initial data $\tilde{y}(0, x) = \mathcal{H}_{BS}(K \exp(x))$. Moreover, $\tilde{y}(\tau, x)$ is also the unique solution of (9.13).

Proof. [Sme15, Theorem 4.3] implies that (9.13) is equivalent to the following variational formulation: Find $\tilde{y} \in H(0, T; I_x)$ such that

$$
\langle A(\tilde{y}), \varphi \rangle = 0 \quad \text{for all } \varphi \in H(0, T; I_x) \tag{9.21}
$$

with operator $A : H(0,T;I_x) \to H(0,T;I_x)^*$ defined as

$$\langle A(\tilde{y}), \varphi\rangle := \int_0^T \int_I F_\gamma[\tilde{y}] L_\omega \varphi \mathrm{d}x\mathrm{d}\tau + \omega\langle \tilde{y}(0) - \mathcal{H}_{\mathrm{BS}}, \varphi(0)\rangle_{H_0^1(I_x)} \tag{9.22}$$

with inner product $\langle \tilde{y}, \varphi\rangle_{H_0^1(I_x)} := \int_{I_x} \tilde{y}_x\varphi_x + \lambda\tilde{y}\varphi \mathrm{d}x$ associated with the norm $\|\cdot\|_{H_0^1(I_x)}$. Now, the Browder-Minty Theorem, which can be seen as a non-linear generalization of the Lax-Milgram Theorem, provides the existence and uniqueness of the solution with a strongly monotonicity, boundedness, continuity and coercivity assumption on the operator. To show that A is Lipschitz continuous, let $\tilde{y}, \varphi, z \in H(0,T;I_x)$, then the compactness of Σ, assumption (9.9) and the Cauchy-Schwarz inequality implies

$$\begin{aligned} |\langle A(\tilde{y}) - A(\varphi), z\rangle| &\le \|F_\gamma[\tilde{y}] - F_\gamma[\varphi]\|_{L^2(0,T;L^2(I_x))}\|z\|_{L^2(0,T;L^2(I_x))} \\ &\quad + \omega\|\tilde{y}(0) - \varphi(0)\|_{H_0^1(I_x)}\|z(0)\|_{H_0^1(I_x)} \le C\|\tilde{y} - \varphi\|_{H(0,T;I_x)}\|z\|_{H(0,T;I_x)} \end{aligned} \tag{9.23}$$

with constant C depending on T, ω, the strictly positive coefficients $\overline{\sigma}, r > 0$ and $D_0 \ge 0$. Moreover, applying Lemma 9.3, the bound in [Sme15, Theorem 4.3] and the Cauchy-Schwarz inequality results in

$$\begin{aligned} \langle A(\tilde{y}) - A(\varphi), \tilde{y} - \varphi\rangle \ge \frac{\varepsilon}{2}\|\tilde{y} - \varphi\|^2_{H(0,T;I_x)} + \frac{\omega}{2}\|\tilde{y}(T) - \varphi(T)\|^2_{H_0^1(I_x)} \\ + \frac{\omega}{2}\|\tilde{y}(0) - \varphi(0)\|^2_{H_0^1(I_x)}. \end{aligned} \tag{9.24}$$

Thus, due to (9.23) and (9.24) the operator A is bounded, continuous, coercive and strongly monotone and the Browder-Minty Theorem implies that there exists a unique solution $\tilde{y} \in H(0,T;I_x)$ to (9.21). Moreover, $\tilde{y}$ is also the unique solution to $F_\gamma(\tilde{y}) = 0$ with initial data $\tilde{y}(0) = \mathcal{H}_{\mathrm{BS}}(K\exp(x))$.
Finally, the author in [Sme15] shows that $\tilde{y} \in H(0,T;I_x)$ solves $F_\gamma(\tilde{y}) = 0$ with initial data $\tilde{y}(0) = \mathcal{H}_{\mathrm{BS}}(K\exp(x))$ if and only if $\tilde{y}$ solves (9.5). □

The next Corollary provides the well-posedness of the BSB equation in the original variable S.

Corollary 9.5. *Let the data $\overline{\sigma} \in \Sigma$, r and D_0 satisfy (9.9) and the cordes condition in (9.10) with $\lambda > 0$ and $\omega > 0$. Let us further assume that $0 < S_{\min} < S_{\max} < \infty$.*
Then there exists a unique strong solution $y(\tau, S) \in H(0,T;\tilde{I})$ of the BSB equation in (9.4) with initial data $y(0,S) = \mathcal{H}_{BS}(S)$.

Proof. Working out the solution in the original variables $\tilde{y}(\tau, x) = y(\tau, S)$ with the transformation $S = K\exp(x)$ for $x \in (x_{\min}, x_{\max})$ and $S \in (S_{\min}, S_{\max})$, the result is a direct consequence of Theorem 9.4. □

9.3. Time semi-discretization and spatial B-spline discretization

In this section, we specify the discretization scheme for Problem 9.1 of Hamilton-Jacobi-Bellman type with B-splines of higher order in space and an implicit scheme for the time discretization. A former approach for the multidimensional Hamilton-Jacobi-Bellman equation is based on a finite-element discretization with piecewise linear hat functions on unstructured meshes with possibly degenerate, isotropic diffusion in [JS13].
First, we derive a time semi-discretization of time transformed problem in (9.3) with backward differentiation formulas (BDF), which approximate the time derivative with Lagrange interpolation polynomials. To avoid repetitions, we concentrate on the *BSB equation in the worst case scenario* as defined in (9.3). Let $0 = \tau^{(0)} < \tau^{(1)} < \cdots < \tau^{(\#\mathbb{T})} = T$ be a partition of the time interval $[0, T]$ with time increment $s := \tau^{(z+1)} - \tau^{(z)}$. We denote the semi-discrete solution in time step z by $y(\tau^{(z)}, S) := y^{(z)}$. Let us consider (9.3) with the bounded linear operator defined as

$$\mathcal{L}^{\underline{\sigma}} y := \frac{1}{2} S^2 \underline{\sigma}^2 y_{SS} + (r - D_0) S y_S - r y, \tag{9.25}$$

which leads to the following time semi-discretization.

- BDF1/implicit Euler scheme (*BDF of order one*):

$$\begin{aligned} &\left(\frac{y^{(z+1)} - y^{(z)}}{s} \right) - \inf_{\underline{\sigma} \in \Sigma} \left(\mathcal{L}^{\underline{\sigma}} y^{(z+1)} \right) = 0 \quad \text{for } 0 \le z \le \#\mathbb{T} - 1, \\ &y^{(0)} = \mathcal{H}(S); \end{aligned} \tag{9.26}$$

- BDF2 scheme (*BDF of order two*):

$$\begin{aligned} &\left(\frac{3y^{(z+2)} - 4y^{(z+1)} + y^{(z)}}{2s} \right) - \inf_{\underline{\sigma} \in \Sigma} \left(\mathcal{L}^{\underline{\sigma}} y^{(z+2)} \right) = 0 \quad \text{for } 0 \le z \le \#\mathbb{T} - 2, \\ &y^{(z)} = \widetilde{y}^{(z)} \quad \text{for } 0 \le z \le 1. \end{aligned} \tag{9.27}$$

BDF2 is a three level scheme, therefore in order to start the algorithm we need initial data $y^{(z)} = \widetilde{y}^{(z)}$ for $0 \le z \le z_0$, where $\widetilde{y}^{(0)}$ is given by the initial condition and other values are obtained by the implicit Euler method (BDF1). Since the non-smooth initial condition is given by the piecewise linear continuous payoff function an implicit Euler method for the initial time steps should be used to avoid oscillations in the numerical solution and its partial derivatives, see Remark 4.7. Moreover, the BDF2 scheme provides second order convergence rates for the time discretization if the solution is smooth enough. In the context of HJB equations the advantage of the BDF2 method in comparison to the Rannacher-timestepping method is that we only need to evaluate the volatility function once per time step.

The next step is the space discretization of the semi-discrete Black-Scholes-Barenblatt equation. Traditionally, with Galerkin schemes the space discretization is achieved by establishing a weak formulation of the partial differential equation with test functions and then by approximating the solution space with B-splines. As already mentioned, the non-linear spatial term of the HJB equation does not admit a weak formulation. But in the discrete setting multiplication with B-splines can be seen as the regularization of the residual. This idea has been established by [JS13] in the context of finite element methods with hat functions.
Let $\mathcal{V}_\Theta \subset \mathbb{S}_{k,\Theta}$, $\ell \in \mathbb{N}$, be a subspace of finite dimensional B-spline spaces of order $k \in \mathbb{N}$, $k \geq 2$ with knot series $\Theta := \{\theta_i\}_{i=1,\dots,n_\Theta}$, which satisfy homogeneous Dirichlet boundary conditions on ∂I. The knot series for the B-spline space is defined as follows

$$\theta_1 = \ldots = \theta_k < \theta_{k+1} < \ldots < \theta_\wp = \ldots = \theta_{\wp+k-2} < \ldots \leq \ldots < \theta_{n_\Theta+1} = \ldots = \theta_{n_\Theta+k} \tag{9.28}$$

with $\theta_1 := S_{\min}$, $\theta_{n_\Theta+1} := S_{\max}$ and $\theta_\wp = K$, where the maximum multiplicity of the knots we allow is $k-1$. Thus, the finite dimensional B-spline space is given by

$$\mathcal{V}_\Theta := \operatorname{span}\{N_{i,k}(S) : \mathcal{I} := \{i = 2, \dots, n_\Theta - 1\}\} \subset \mathbb{S}_{k,\Theta}, \tag{9.29}$$

where $N_{1,k}(S), N_{n_\Theta,k}(S)$ are omitted to depict the homogeneous Dirichlet boundary conditions. The grid size of the B-spline basis is referred to $h_\ell := \max\{\theta_{i+1} - \theta_i\}$.
Set $\tilde{N}_{i,k}(S) := N_{i,k}(S)/\|N_{i,k}(S)\|_{L^1(I)}$. Thus, the $N_{i,k}(S)$ are normalized in the $L^\infty(I)$-norm and $\tilde{N}_{i,k}(S)$ in the $L^1(I)$-norm. Moreover, the support of a normalized B-spline $\tilde{N}_{i,k}$ is $\tilde{\mathcal{Q}} := \operatorname{supp}(\tilde{N}_{i,k}(S)) \subseteq [\theta_i, \theta_{i+k}]$ and due to the non-negativity of B-splines $N_{i,k}(S) \geq 0$ one has $\tilde{N}_{i,k}(S) \geq 0$ for all $i \in \mathcal{I}$.
Let us consider the non-linear spatial term in (9.26) and (9.27) for a fixed time step z. By multiplying with test functions $\tilde{N}_{i,k}(S) \geq 0$ and integrating over $\tilde{\mathcal{Q}}$, if $x \in \tilde{\mathcal{Q}}$ on a fine grid one has

$$\begin{aligned}
&\inf_{\underline{\sigma}\in\Sigma} \left(\int_{\tilde{\tilde{\mathcal{Q}}}} \frac{1}{2}\underline{\sigma}^2 x^2 \; y_{SS}^{(z)}(x)\tilde{N}_{i,k}(S)\mathrm{d}S \right) \\
&= \inf_{\underline{\sigma}\in\Sigma} \left(\frac{1}{2}\underline{\sigma}^2 \int_{\tilde{\tilde{\mathcal{Q}}}} x^2 \; y_{SS}^{(z)}(x)\tilde{N}_{i,k}(S)\mathrm{d}S \right) \\
&\approx \inf_{\underline{\sigma}\in\Sigma} \left(\frac{1}{2}\underline{\sigma}^2 \int_{\tilde{\tilde{\mathcal{Q}}}} S^2 \; y_{SS}^{(z)}(S)\tilde{N}_{i,k}(S)\mathrm{d}S \right) \\
&= \inf_{\underline{\sigma}\in\Sigma} \left(-\frac{1}{2}\underline{\sigma}^2 \int_{\tilde{\tilde{\mathcal{Q}}}} y_S^{(z)}(S) \left(S^2\tilde{N}'_{i,k}(S) + 2S\tilde{N}_{i,k}(S) \right) \mathrm{d}S \right),
\end{aligned} \tag{9.30}$$

since $\tilde{N}_{i,k}$ approximates a Dirac-Delta as the element size is decreased.

Let us further denote by $\underline{\sigma}^-(\tau^{(z)})$ the exact volatility function and by $\underline{\sigma}_\theta(\tau^{(z)})$ the approximated volatility function for a fixed time level. The above approximation in (9.30) leads to an error in the volatility function, one has

$$\|\underline{\sigma}^-(\tau^{(z)}) - \underline{\sigma}_\theta(\tau^{(z)})\|_{L^\infty} \leq C(\theta_{i+1} - \theta_i) \tag{9.31}$$

with C depending also on k. For a linear finite element discretization this estimate can also be found in ([Sme15, Chapter 1.3]).
For further considerations we denote by $H := L^2(I)$ the space of square integrable functions and define the weighted Sobolev space with homogeneous Dirichlet boundary conditions as follows

$$\mathcal{V}^0 := \left\{ \varphi \in H : S\frac{\partial \varphi}{\partial S} \in H, \varphi = 0 \text{ on } \partial I \right\}. \tag{9.32}$$

Setting $u^{(z)} := y^{(z)} - \mathcal{H}_{\mathrm{BS}} \in \mathcal{V}^0$ leads to the following discretization of the BSB equation for the schemes (9.26) and (9.27) in a weak form with homogeneous Dirichlet boundary condition:

Problem 9.6. *(BDF1 scheme)*
Find $u^{(z+1)} \in \mathcal{V}^0$ for each $z \in \{0, \dots, \#\mathbb{T} - 1\}$ such that

$$\begin{aligned} &\left(\frac{u^{(z+1)} - u^{(z)}}{s}, \tilde{N}_{i,k} \right) - \inf_{\underline{\sigma} \in \Sigma} \left\langle \mathcal{A}^{\underline{\sigma}}(u^{(z+1)} + \mathcal{H}), \tilde{N}_{i,k} \right\rangle = 0 \text{ for all } i \in \mathcal{I}, \\ &u^{(0)} = 0, \end{aligned} \tag{9.33}$$

with $\mathcal{A}^{\underline{\sigma}} : \mathcal{V}^0 \to \mathbb{R}^{n_\Theta}$ defined as

$$\begin{aligned} \left\langle \mathcal{A}^{\underline{\sigma}} u^{(z)}, \tilde{N}_{i,k} \right\rangle := &- \frac{1}{2}\underline{\sigma}^2 \int u_S^{(z)} (S^2 \tilde{N}'_{i,k} + 2S\tilde{N}_{i,k}) \mathrm{d}S \\ &+ \int \left((r - D_0) S u_S^{(z)} - r u^{(z)} \right) \tilde{N}_{i,k} \mathrm{d}S. \end{aligned} \tag{9.34}$$

Problem 9.7. *(BDF2 scheme)*
Find $u^{(z+2)} \in \mathcal{V}^0$ for each $z \in \{0, \dots, \#\mathbb{T} - 2\}$ such that

$$\begin{aligned} &\left(\frac{3u^{(z+2)} - 4u^{(z+1)} + u^{(z)}}{2s}, \tilde{N}_{i,k} \right) - \inf_{\underline{\sigma} \in \Sigma} \left\langle \mathcal{A}^{\underline{\sigma}}(u^{(z+2)} + \mathcal{H}), \tilde{N}_{i,k} \right\rangle = 0 \quad \text{for all } i \in \mathcal{I}, \\ &u^{(z)} = \tilde{u}^{(z)} \quad \text{for } z \in \{0, 1\}, \end{aligned} \tag{9.35}$$

with $\left\langle \mathcal{A}^{\underline{\sigma}} u^{(z)}, \tilde{N}_{i,k} \right\rangle$ as defined in (9.34).

Now, we define the fully discrete form for the BDF1 and BDF2 scheme. Let $u_\Theta^{(z)} \in \mathcal{V}_\Theta$ the spline solution with

$$u_\Theta^{(z)} := u_\Theta^{(z)}(S) := \sum_{j \in \mathcal{I}} \mathbf{u}_j^{(z)} N_{j,k}(S). \tag{9.36}$$

Inserting the numerical solution in Problem 9.6 and Problem 9.7 leads to the following numerical schemes:

- BDF1 scheme: Find $u_\Theta^{(z+1)} \in \mathcal{V}_\Theta \subset \mathbb{S}_{k,\Theta}$ for each $z \in \{0, \dots, \#\mathbb{T} - 1\}$ such that

$$\left\langle \frac{u_\Theta^{(z+1)} - u_\Theta^{(z)}}{s}, \tilde{N}_{i,k} \right\rangle - \inf_{\underline{\sigma} \in \Sigma} \left\langle \mathcal{A}^{\underline{\sigma}}(u_\Theta^{(z+1)} + \mathcal{H}), \tilde{N}_{i,k} \right\rangle = 0 \quad \text{for all } i \in \mathcal{I},$$
$$u_\Theta^{(0)} = 0. \tag{9.37}$$

- BDF2 scheme: Find $u_\Theta^{(z+2)} \in \mathcal{V}_\Theta \subset \mathbb{S}_{k,\Theta}$ for each $z \in \{0, \dots, \#\mathbb{T} - 2\}$ such that

$$\left\langle \frac{3u_\Theta^{(z+2)} - 4u_\Theta^{(z+1)} + u_\Theta^{(z)}}{2s}, \tilde{N}_{i,k} \right\rangle - \inf_{\underline{\sigma} \in \Sigma} \left\langle \mathcal{A}^{\underline{\sigma}}(u_\Theta^{(z+2)} + \mathcal{H}), \tilde{N}_{i,k} \right\rangle = 0 \quad \text{for all } i \in \mathcal{I},$$
$$u_\Theta^{(z)} = \tilde{u}_\Theta^{(z)} \quad \text{for } z \in \{0, 1\}. \tag{9.38}$$

As already mentioned, one advantage of the BSB equation is that the infimum can be expressed as an easy formula otherwise numerical optimization algorithms are needed. Let us consider (9.37) or (9.38) then the infimum in time step $\tilde{z}$ is attained for

$$\left(\hat{\underline{\boldsymbol{\sigma}}}_\Theta(u_\Theta^{(\tilde{z})})\right)_i := \begin{cases} \sigma_{\min}^2, & \text{if } -\int_{\tilde{\mathcal{Q}}} \left(\frac{\partial u_\Theta^{(\tilde{z})} + \mathcal{H}(S)}{\partial S} \left(S^2 \tilde{N}'_{i,k} + 2S\tilde{N}_{i,k} \right) \right) \mathrm{d}S \geq 0 \\ \sigma_{\max}^2, & \text{if } -\int_{\tilde{\mathcal{Q}}} \left(\frac{\partial u_\Theta^{(\tilde{z})} + \mathcal{H}(S)}{\partial S} \left(S^2 \tilde{N}'_{i,k} + 2S\tilde{N}_{i,k} \right) \right) \mathrm{d}S < 0, \end{cases} \tag{9.39}$$

which corresponds to the volatility in the worst case scenario. Moreover, in the case of a cubic B-spline discretization the volatility can be found by an pointwise approximation of the second derivative. Therefore, let $u_\Theta^{(\tilde{z})}$ the cubic spline solution in time step $\tilde{z}$. Integration by parts and a similar approximation as in (9.30) for $x \in \tilde{\mathcal{Q}}$ leads to

$$-\int_{\tilde{\mathcal{Q}}} \left(\frac{\partial u_\Theta^{(z)} + \mathcal{H}(S)}{\partial S} \left(S^2 \tilde{N}'_{i,k} + 2S\tilde{N}_{i,k} \right) \right) \mathrm{d}S \approx S^2 \frac{\partial^2 \left(u_\Theta^{(\tilde{z})} \right)}{\partial S^2} \Big|_{S=x}. \tag{9.40}$$

Hence the volatility in the worst case scenario for a cubic B-spline discretization can be approximated as follows for $x \in \tilde{\mathcal{Q}}$

$$\left(\hat{\underline{\boldsymbol{\sigma}}}_\Theta(u_\Theta^{(\tilde{z})})\right)_i \approx \hat{\underline{\boldsymbol{\sigma}}}_\Theta^{\text{opt}} \left(u_\Theta^{(\tilde{z})}(x) \right) := \begin{cases} \sigma_{\min}^2 & \text{if } S^2 \frac{\partial^2 u_\Theta^{(\tilde{z})}}{\partial S^2} \Big|_{S=x} \geq 0, \\ \sigma_{\max}^2 & \text{if } S^2 \frac{\partial^2 u_\Theta^{(\tilde{z})}}{\partial S^2} \Big|_{S=x} < 0. \end{cases} \tag{9.41}$$

Comparing the approximation of the volatility function in (9.41) and (9.39) one can see that the approximation of the volatility function differs in its regularity requirements. The approximation in (9.41) needs a stable approximation of the second derivative

while the approximation in (9.39) only needs a stable approximation in $H^1(I)$.
We now specify the numerical scheme in matrix vector notation. The formula for the volatility in (9.39) or (9.41) allows us to express the infimum in (9.37) or (9.38) as follows

$$\inf_{\underline{\sigma}\in\Sigma}\left\langle \mathcal{A}^{\underline{\sigma}}(u_\Theta^{(z+1)}+\mathcal{H}),\tilde{N}_{i,k}\right\rangle = -\left((\mathfrak{A}_\Theta^{\hat{\boldsymbol{\sigma}}_\Theta^{(\tilde{z})}}+\mathfrak{B}_\Theta^{\hat{\boldsymbol{\sigma}}_\Theta^{(\tilde{z})}}+r\mathfrak{G}_\Theta)\mathbf{u}_\Theta\right)_i \quad \text{for all } i\in\mathcal{I}, \tag{9.42}$$

with B-spline coefficient vector $\mathbf{u}_\Theta\in\mathbb{R}^{n_\Theta}$ and discretization matrices $\mathfrak{A}_\Theta^{\hat{\boldsymbol{\sigma}}_\Theta^{(\tilde{z})}},\mathfrak{B}_\Theta^{\hat{\boldsymbol{\sigma}}_\Theta^{(\tilde{z})}},\mathfrak{G}\in\mathbb{R}^{n_\Theta\times n_\Theta}$ defined as

$$\left(\mathfrak{A}_\Theta^{\hat{\boldsymbol{\sigma}}_\Theta^{(\tilde{z})}}\right)_{i,j} := \frac{\left(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(\tilde{z})}\right)_i}{2}\int_{\mathcal{Q}}(S^2N'_{j,k}(S)\tilde{N}'_{i,k}(S))\mathrm{d}S, \tag{9.43}$$

$$\left(\mathfrak{B}_\Theta^{\hat{\boldsymbol{\sigma}}_\Theta^{(\tilde{z})}}\right)_{i,j} := \left(\left(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(\tilde{z})}\right)_i+D_0-r\right)\int_{\mathcal{Q}}(SN'_{j,k}(S)\tilde{N}_{i,k}(S))\mathrm{d}S, \tag{9.44}$$

$$\mathfrak{G}_{i,j} := \int_{\mathcal{Q}} N_{j,k}(S)\tilde{N}_{i,k}(S)\mathrm{d}S. \tag{9.45}$$

For the right hand side we also define the vector $\mathfrak{d}_\Theta^{\hat{\boldsymbol{\sigma}}_\Theta^{(\tilde{z})}}\in\mathbb{R}^{n_\Theta}$ as follows

$$\begin{aligned}\left(\mathfrak{d}_\Theta^{\hat{\boldsymbol{\sigma}}_\Theta^{(\tilde{z})}}\right)_i := &\int\left(\frac{\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(\tilde{z})}}{2}S^2\mathcal{H}'(S)\tilde{N}'_{i,k}(S)\right)+r\mathcal{H}(S)\tilde{N}_{i,k}(S)\mathrm{d}S\\ &+\int\left(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(\tilde{z})}+D_0-r\right)S\mathcal{H}'(S)\tilde{N}_{i,k}(S)\mathrm{d}S.\end{aligned}$$

Note that the discretization matrices are very similar to the matrices for the standard Black-Scholes model in (4.37). Finally, we can rewrite (9.37) and (9.38) to the following problem in matrix vector notation:

- BDF1: Find $\mathbf{u}_\Theta^{(z+1)}\in\mathbb{R}^{n_\Theta}$ for each $z\in\{0,\dots,\#\mathbb{T}-1\}$ such that $\mathbf{u}_\Theta^{(0)}=0$,

$$F^1(\mathbf{u}_\Theta^{(z+1)}) := C_\Theta^1\left(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(z+1)}\right)\mathbf{u}_\Theta^{(z+1)}-\mathbf{f}_\Theta^1\left(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(z+1)}\right)=0, \tag{9.46}$$

  where $C_\Theta^1\left(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(z+1)}\right)\in\mathbb{R}^{n_\Theta\times n_\Theta}$, and $\mathbf{f}_\Theta^1\left(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(z+1)}\right)\in\mathbb{R}^{n_\Theta}$ are defined as

$$\begin{aligned}C_\Theta^1\left(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(z+1)}\right) &:= \left(\mathfrak{A}_\Theta^{\hat{\boldsymbol{\sigma}}_\Theta^{(z+1)}}+\mathfrak{B}_\Theta^{\hat{\boldsymbol{\sigma}}_\Theta^{(z+1)}}+(r+s^{-1})\mathfrak{G}_\Theta\right),\\ \mathbf{f}_\Theta^1\left(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(z+1)}\right) &:= s^{-1}\mathfrak{G}_\Theta\mathbf{u}_\Theta^{(z)}-\mathfrak{d}_\Theta^{\hat{\boldsymbol{\sigma}}_\Theta^{(z+1)}}.\end{aligned} \tag{9.47}$$

- BDF2: Find $\mathbf{u}_\Theta^{(z+2)}\in\mathbb{R}^{n_\Theta}$ for each $z\in\{0,\dots,\#\mathbb{T}-2\}$ such that $\mathbf{u}_\Theta^{(z_0)}=\tilde{\mathbf{u}}_\Theta^{(z_0)}$,

$$F^2(\mathbf{u}_\Theta^{(z+1)}) := C_\Theta^2\left(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(z+2)}\right)\mathbf{u}_\Theta^{(z+2)}-\mathbf{f}_\Theta^2\left(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{(z+2)}\right)=0, \tag{9.48}$$

where $C^2_\Theta\left(\underline{\hat{\boldsymbol{\sigma}}}^{(z+2)}_\Theta\right) \in \mathbb{R}^{n_\Theta \times n_\Theta}$, $\mathbf{f}^2_\Theta\left(\underline{\hat{\boldsymbol{\sigma}}}^{(z+2)}_\Theta\right) \in \mathbb{R}^{n_\Theta}$ are defined as

$$\begin{aligned} C^2_\Theta\left(\underline{\hat{\boldsymbol{\sigma}}}^{(z+2)}_\Theta\right) &:= \left(\mathfrak{A}^{\underline{\hat{\boldsymbol{\sigma}}}^{(z+2)}_\Theta}_\Theta + \mathfrak{B}^{\underline{\hat{\boldsymbol{\sigma}}}^{(z+2)}_\Theta}_\Theta + \left(r + \frac{3}{2}s^{-1}\right)\mathfrak{G}_\Theta\right), \\ \mathbf{f}^2_\Theta\left(\underline{\hat{\boldsymbol{\sigma}}}^{(z+2)}_\Theta\right) &:= \frac{1}{2}s^{-1}\mathfrak{G}_\Theta(4\mathbf{u}^{(z+1)}_\Theta - \mathbf{u}^{(z)}_\Theta) - \mathfrak{d}^{\underline{\hat{\boldsymbol{\sigma}}}^{(z+2)}_\Theta}_\Theta. \end{aligned} \tag{9.49}$$

For various reasons, there are some difficulties associated with the numerical solution of the non-linear Black-Scholes-Barenblatt equation. In order to obtain a Newton like method for solving the discrete BSB equation, we need an analogous formulation of the Jacobian since the non-linear operator with jumping volatility is not differentiable in the classical sense. This problem has already been solved by [BMZ09, Sme12, Sme15]. The principal requirement of the operator which fits into the framework of HJB equations is slant differentiability. This concept is introduced in Subsection 9.4.1.
Another problem arises in the numerical approximation of the solution and its partial derivatives with cubic B-splines since a numerical approximation can result in oscillations in the partial derivatives. To explain this fact let us first consider the volatility function in (9.41) together with a cubic B-spline discretization. The volatility jumps at those knots $\theta_{\wp^*}$, where the sign of the second derivative in one time step changes. It should be clear that these knots $\theta_{\wp^*}$ are only an approximation of the zeros of the second derivative and the approximation of the second derivative are not exactly zero at these zero points. Thus, this inaccurate approximation of the zeros has a direct consequence on the volatility function and leads to discontinuities in the partial derivatives of the numerical solution. In order to stabilize these approximations for cubic B-splines, when the volatility function in (9.41) is used, the following two aspects are necessary to obtain an accurate discretization:

i) The knots, where the volatility jumps from $\sigma^2_{\min}$ to $\sigma^2_{\max}$ or vice versa are repeated $k-1$ times.

ii) Moreover, one has to repeat those knots $k-1$ times, where the initial condition is not differentiable.

As mentioned before, the volatility function in (9.39) only needs a stable approximation of the solution in the $H^1(I)$-norm. Thus, an unstable approximation of the second derivatives has no direct influence of the approximation of the volatility function. Therefore, it is possible to find a numerical solution of the BSB equation without inserting new knots. But due to the approximation of the volatility function optimal convergence rates for higher order B-splines cannot be expected.
In order to compute an accurate numerical solution of the BSB equation with optimal convergence, we develop a semismooth Newton algorithm within a knot insertion step at the zeros of the second derivative in Subsection 9.4.2.

9.4. Semismooth Newton method

In this section we introduce a semismooth Newton method within a knot insertion step to improve the approximation of the solution and its derivatives with higher order B-splines. To simplify the notation we omit the time steps and make no distinction between the BDF1 and BDF2 scheme. Thus, we consider the following discrete BSB equation in the worst case scenario: Find $\mathbf{u}_\Theta \in \mathbb{R}^{n_\Theta}$ such that

$$F(\mathbf{u}_\Theta) := C_\Theta(\hat{\underline{\boldsymbol{\sigma}}})\mathbf{u}_\Theta - \mathbf{f}_\Theta(\hat{\underline{\boldsymbol{\sigma}}}) = 0 \tag{9.50}$$

with $C_\Theta(\hat{\underline{\boldsymbol{\sigma}}}), \mathbf{f}_\Theta(\hat{\underline{\boldsymbol{\sigma}}})$ as defined in (9.47) for the BDF1 scheme or (9.49) for the BDF2 scheme. In the literature, the following equivalent formulation can frequently be found

$$F(\mathbf{u}_\Theta) := \max_{\underline{\boldsymbol{\sigma}} \in \Sigma^{n_\Theta}} \{C_\Theta(\underline{\boldsymbol{\sigma}})\mathbf{u}_\Theta - \mathbf{f}_\Theta(\underline{\boldsymbol{\sigma}})\} = 0. \tag{9.51}$$

The change from a minimum to a maximum results from comparing the sign change in (9.38) and (9.48).
In order to obtain a Newton like method for solving the discrete BSB equation, we need a analog formulation of the Jacobian, since the operators from (9.46) and (9.48) are not differentiable in the classical sense. Semismoothness of real valued functions in the context of nonsmooth nonconvex constrained optimization was introduced in [Mif77]. Moreover, the authors in [Qi93, QS93] established the concept of semismoothness for mappings between finite dimensional spaces and generalized the Newton method to semismooth equations. In the context of HJB-equations with a finite difference or a linear finite element discretization, the concept of a semismooth Newton method was used in [BMZ09, Hei10, Sme12]. In particular, assuming some monotonicity properties on the discretization matrix, that is $C_\Theta(\underline{\boldsymbol{\sigma}})^{-1} \geq 0$ componentwise, the authors in [BMZ09] have proved that the HJB equation in the discrete setting is slantly differentiable. This monotonicity assumption has been used with success for a finite difference discretization or a discretization with linear basis function (see [JS13, BMZ09]), but is not applicable in general for our derived discrete BSB equation. Another result for HJB equations in the discrete setting can be found in [Sme12]. There, the author constructs a general method to find a slant derivative by not requiring any monotonicity properties of the operator. We will use the last-mentioned result in the next subsection.

9.4.1. Slant derivative

First, we introduce the definition of a *slant derivative.*

Definition 9.8. *(Slant derivative)*
Let $U \subset \mathbb{R}^N$ be open. A function $F : U \mapsto \mathbb{R}^N$ is slantly differentiable in U if there exists $G(\cdot) : U \mapsto M(N, \mathbb{R})$ such that for every $\mathbf{x} \in U$

$$\lim_{\mathbf{h} \to 0} \frac{1}{\|\mathbf{h}\|} \|F(\mathbf{x} + \mathbf{h}) - F(\mathbf{x}) - G(\mathbf{x} + \mathbf{h})\mathbf{h}\| = 0. \tag{9.52}$$

G is called a slant derivative of F in U.

Consider the definition of classical derivatives $\lim_{h\to 0} 1/|h|\,|F(x+h) - F(x) - G(x)h| = 0$ for $x \in \mathbb{R}$, we see that all classical differentiable function are slant differentiable. A difference to classical derivatives is that a slant differentiable function is not necessarily unique. As a result, there is no general method to find a slant derivative. In the context of discrete HJB equations the author in [Sme12] has introduced a general method to find a slant derivative for a linear and continuous map $\underline{\boldsymbol{\sigma}} \mapsto C_\Theta(\underline{\boldsymbol{\sigma}})$ and compact space Σ. The next Corollary provides a slant derivative for the discrete BSB equation.

Corollary 9.9. *(Slant derivative for the discrete BSB equation)*
Let $C_\Theta(\underline{\boldsymbol{\sigma}}) \in \mathbb{R}^{n_\Theta \times n_\Theta}$ and $\mathbf{f}_\Theta(\underline{\boldsymbol{\sigma}}) \in \mathbb{R}^{n_\Theta}$ with $\underline{\boldsymbol{\sigma}} \in \Sigma^{n_\Theta}$ as defined in (9.47) for the BDF1 scheme or (9.49) for the BDF2 scheme. Then

$$(G(\mathbf{u}_\Theta))_{i,j} := (C_\Theta(\hat{\underline{\boldsymbol{\sigma}}}_i))_{i,j} \quad \text{with } \hat{\underline{\boldsymbol{\sigma}}}_i \text{ choosen from (9.39)} \tag{9.53}$$

is a slant derivative of $F(\mathbf{u}_\Theta)$ defined in (9.51).

Proof. Since $\underline{\boldsymbol{\sigma}} \mapsto C_\Theta(\underline{\boldsymbol{\sigma}})$ and $\underline{\boldsymbol{\sigma}} \mapsto \mathbf{f}_\Theta(\underline{\boldsymbol{\sigma}})$ are continuous and $C_\Theta(\underline{\boldsymbol{\sigma}})$ is linear, the result follows directly from [Sme12, p.46, Corollary 3.7]. □

9.4.2. Semismooth Newton method within a knot insertion step

Within the framework of HJB equations a well-known method to solve (9.50) is the so called Howard's algorithm. It can be shown that Howard's algorithm is equivalent to a semismooth Newton method. The algorithm has been developed by [Bel57] and [How60] to solve steady infinite-horizon Markovian dynamic programming problems. The algorithm can be found in the context of discrete HJB equations for example in [BMZ09, Sme12, JS13] and in the context of the BSB equation with a finite difference discretization in [Hei10] as well. The semismooth Newton method to solve $F(\mathbf{u}_\Theta) = 0$ with $F(\mathbf{u}_\Theta) : \mathbb{R}^{n_\Theta} \to \mathbb{R}^{n_\Theta}$ and invertible slant derivative $G_\Theta \in \mathbb{R}^{n_\Theta \times n_\Theta}$ is as follows:

Algorithm 9.1 Semismooth Newton method

1) Initialize $\mathbf{u}_\Theta^0$ and $\varsigma := 0$.
2) Compute $\hat{\underline{\boldsymbol{\sigma}}}_\Theta^\varsigma := \hat{\underline{\boldsymbol{\sigma}}}_\Theta(\mathbf{u}_\Theta^\varsigma)$ (9.39).
3) Compute $F(\mathbf{u}_\Theta^\varsigma)$ and it's slant derivative $G_\Theta(\hat{\underline{\boldsymbol{\sigma}}}_\Theta^\varsigma)$.
4) Solve $G_\Theta(\hat{\underline{\boldsymbol{\sigma}}}_\Theta^\varsigma)\left(\mathbf{u}_\Theta^{\varsigma+1} - \mathbf{u}_\Theta^\varsigma\right) = -F^p(\mathbf{u}_\Theta^\varsigma)$
5) If $\|\mathbf{u}_\Theta^{\varsigma+1} - \mathbf{u}_\Theta^\varsigma\| \leq$tol stop otherwise return to step 2).

Note, that this algorithm only works for the volatility function in (9.39), where no stable approximation of the second derivatives is needed.
From [BMZ09, Sme12] it is known that if $F(\cdot)$ is slantly differentiable, the slant derivative is invertible and its inverse is bounded so the semismooth Newton method converges superlinearly.
As mentioned in Subsection 9.3 the approximation of the volatility function for the discrete BSB equation leads to discontinuities in the partial derivatives of the numerical

solution. In order to stabilize the B-spline approximation, we reduce the smoothness of the approximation by repeating knots at those points. Therefore, let $\mathbf{u}_\Theta^\varsigma \in \mathbb{R}^{n_\Theta}$ the given Newton iterate or B-spline coefficient vector in the ςth iteration step of the Newton method. Let further $u_\Theta^\varsigma \in \mathcal{V}_\Theta \subset \mathbb{S}_{k,\Theta}$ the corresponding spline solution of the BSB equation in the worst case scenario. We define

$$I_\eta^+ := \left\{ \theta_i \in \Theta | - \int_{\tilde{\mathcal{Q}}} \left(\frac{\partial u_\Theta^\varsigma + \mathcal{H}(S)}{\partial S} \left(S^2 \tilde{N}'_{i,k} + 2S\tilde{N}_{i,k} \right) \right) \mathrm{d}S \geq 0 \text{ for all } i = i_\eta, i_\eta + 1, \dots \right\} \tag{9.54}$$

and

$$I_\eta^- := \left\{ \theta_i \in \Theta | - \int_{\tilde{\mathcal{Q}}} \left(\frac{\partial u_\Theta^\varsigma + \mathcal{H}(S)}{\partial S} \left(S^2 \tilde{N}'_{i,k} + 2S\tilde{N}_{i,k} \right) \right) \mathrm{d}S < 0 \text{ for all } i = i_\eta, i_\eta + 1, \dots \right\}. \tag{9.55}$$

Thus, I_η^+ corresponds to the knots, where the volatility function in (9.39) is given by $\underline{\hat{\boldsymbol{\sigma}}}(u_\Theta^\varsigma(\theta_i)) = \sigma_{\min}^2$, and I_η^- corresponds to the knots, where $\underline{\hat{\boldsymbol{\sigma}}}(u_\Theta^\varsigma(\theta_i)) = \sigma_{\max}^2$. In order to find the knots, where the volatility functions jumps from $\sigma_{\min}^2$ to $\sigma_{\max}^2$ or vice versa, compute all $\theta_\wp \in \Theta$ such that

$$\theta_\wp \in I_\eta^+ \text{ and } \theta_{\wp+1} \in I_{\eta+1}^- \quad \text{or} \quad \theta_\wp \in I_\eta^- \text{ and } \theta_{\wp+1} \in I_{\eta+1}^+. \tag{9.56}$$

Now, we construct a knot series $\tilde{\Theta}$, where the knots $\theta_\wp$ have multiplicity $k-1$. Therefore, let the set $\Theta_\wp \leftarrow \{\theta_\wp, \dots, \theta_\wp\}$ consists of $k-2$ knots $\theta_\wp$. If $\Theta_\wp \notin \Theta$ set

$$\tilde{\Theta} \leftarrow \Theta \cup \Theta_\wp, \tag{9.57}$$

otherwise in the knot series Θ the knot $\theta_\wp$ has already multiplicity $k-1$. Finally, the Newton iterate regarding the new knots series $\tilde{\Theta}$ is computed as follows

$$\mathbf{u}_{\tilde{\Theta}}^\varsigma = K_{\tilde{\Theta}} \mathbf{u}_\Theta^\varsigma, \tag{9.58}$$

where $K_{\tilde{\Theta}} \in \mathbb{R}^{n_{\tilde{\Theta}} \times n_\Theta}$ describes a knot insertion matrix. The construction of a knot insertion matrix is discussed in Appendix A.
In summary, the Newton algorithm for solving the discrete BSB equation in one time step within a knot insertion step to stabilize the spline approximation of the partial derivatives can be found in Algorithm 9.2.

Algorithm 9.2 Newton algorithm with knot repetition of existing knots in Θ

Input: Order of the B-spline basis k, uniform knot series Θ with possible repeating knots, initial guess $\mathbf{u}_\Theta^0 \in \mathbb{R}^{n_\Theta}$.

Output: Solution $\mathbf{u}_\Theta$ regarding a new knot series Θ.

1) Initialize $\mathbf{u}_\Theta^0$ and $\varsigma := 0$.

2a) For all $\eta = 1, 2, \ldots$ compute all $\theta_\wp \in \Theta$ such that

$$\theta_\wp \in I_\eta^+ \text{ and } \theta_{\wp+1} \in I_{\eta+1}^- \quad \text{or} \quad \theta_\wp \in I_\eta^- \text{ and } \theta_{\wp+1} \in I_{\eta+1}^+.$$

2b) Repeat each knot $\theta_\wp$ as defined above $k-2$ times and set $\Theta_\wp \leftarrow \{\theta_\wp, \ldots, \theta_\wp\}$. If $\Theta_\wp \notin \Theta$ set

$$\tilde{\Theta} \leftarrow \Theta \cup \Theta_\wp,$$

otherwise set $\tilde{\Theta} \leftarrow \Theta$.

2c) Compute the Newton iterate regarding the new knot series $\tilde{\Theta}$ by

$$\mathbf{u}_{\tilde{\Theta}}^\varsigma = K_{\tilde{\Theta}} \mathbf{u}_\Theta^\varsigma \text{ and set } \mathbf{u}_\Theta^\varsigma \leftarrow \mathbf{u}_{\tilde{\Theta}}^\varsigma, \quad \Theta \leftarrow \tilde{\Theta},$$

where $K_{\tilde{\Theta}}$ is the knot insertion matrix from (A.7).

3) Compute $\underline{\hat{\boldsymbol{\sigma}}}_\Theta^\varsigma := \underline{\hat{\boldsymbol{\sigma}}}_\Theta(u_\Theta^\varsigma)$ from equation (9.39).

4) Compute $F(\mathbf{u}_\Theta^\varsigma)$ and its slant derivative $G_\Theta(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^\varsigma)$.

5) Solve $G_\Theta(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^\varsigma)\left(\mathbf{u}_\Theta^{\varsigma+1} - \mathbf{u}_\Theta^\varsigma\right) = -F(\mathbf{u}_\Theta^\varsigma)$.

6) If $\|\mathbf{u}_\Theta^{\varsigma+1} - \mathbf{u}_\Theta^\varsigma\| \leq$tol stop otherwise return to step 2).

In order to improve the approximation of the optimal controls $\underline{\hat{\boldsymbol{\sigma}}}_\Theta$ and also the approximation of the solution and its partial derivatives for cubic B-splines, we are going to present another approach now. We compute zeros of the second derivative with a certain accuracy and add these zeros to the existing knot series Θ. In this context, it is important that the knots are not to close to each other since in this case the discretization matrix becomes singular. Therefore, a certain tolerance is specified. Let us define the zeros of the second derivative $\frac{\partial^2 u_\Theta^\varsigma}{\partial S^2}$ in the ςth Newton step as S_η. Moreover, let

$$I_\eta^{\text{opt},+} := \left\{ x \in \mathbb{R} \Big| \frac{\partial^2 u_\Theta^\varsigma}{\partial S^2}\Big|_{S=x} \geq 0 \right\} \tag{9.59}$$

define the corresponding knots where the volatility function in (9.41) is given by $\underline{\hat{\boldsymbol{\sigma}}}^{\text{opt}}(u_\Theta^\varsigma(\theta_i)) = \sigma_{\min}^2$ and let

$$I_\eta^{\text{opt},-} := \left\{ x \in \mathbb{R} \Big| \frac{\partial^2 u_\Theta^\varsigma}{\partial S^2}\Big|_{S=x} < 0 \right\} \tag{9.60}$$

define the corresponding knots where the volatility function in (9.41) is given by $\underline{\hat{\boldsymbol{\sigma}}}^{\text{opt}}(u_\Theta^\varsigma(\theta_i)) = \sigma_{\max}^2$. Algorithm 9.3 presents a semismooth Newton method for solving the discrete BSB equation with cubic B-splines in the worst case scenario, where the zeros S_η of the second derivative $\frac{\partial^2 u_\Theta^\varsigma}{\partial S^2}$ are inserted $k-1$ times.
In general, the algorithm is applicable for B-splines of order $k \geq 4$, since this enables a pointwise approximation of the second derivatives.
In order to remove unnecessary knots we apply a knot removal algorithm in each time step. A detailed description of the algorithm can be found in [PT97, Chapter 5.4] or Appendix A.

Algorithm 9.3 Newton algorithm with knot insertion of nonexisting knots in Θ

Input: Order of the B-spline basis $k \geq 4$, knot series Θ with possible repeating knots, initial guess $\mathbf{u}_\Theta^0 \in \mathbb{R}^{n_\Theta}$.

Output: Solution $\mathbf{u}_\Theta$ regarding a new knot series Θ.

1) Initialize $\mathbf{u}_\Theta^0$ and $\varsigma := 0$.

2a) Compute all zeros $S_\eta \in I$, $\eta = 1, 2, \ldots$ of the second derivative $\frac{\partial^2 u_\Theta^\varsigma}{\partial S^2}$ to a certain accuracy.

For all $\eta = 1, 2, \ldots$

2b) If $S_\eta \notin \Theta$ insert S_η with multiplicity $k-1$. Set $\tilde{\Theta}_\eta \leftarrow \{S_\eta, \ldots, S_\eta\}$.

If $S_\eta \in \Theta$ with multiplicity one repeat S_η with multiplicity $k-2$. Set $\tilde{\Theta}_\eta \leftarrow \{S_\eta, \ldots, S_\eta\}$.

Else there are knots $S_\eta \in \Theta$ with multiplicity $k-1$, set $\tilde{\Theta}_\eta \leftarrow \varnothing$.

End For

Insert all knots $\tilde{\Theta}_\eta$ to the existing knot series Θ: Set

$$\tilde{\Theta} \leftarrow \Theta \cup \tilde{\Theta}_1 \cup \tilde{\Theta}_2 \cup \ldots.$$

2c) Compute the Newton iterate regarding the new knot series $\tilde{\Theta}$ by

$$\mathbf{u}_{\tilde{\Theta}}^\varsigma = K_{\tilde{\Theta}} \mathbf{u}_\Theta^\varsigma \quad \text{and set} \quad \mathbf{u}_\Theta^\varsigma \leftarrow \mathbf{u}_{\tilde{\Theta}}^\varsigma, \quad \Theta \leftarrow \tilde{\Theta}$$

with the knot insertion matrix $K_{\tilde{\Theta}}$ from (A.7).

3) Compute $\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{\text{opt}} := \underline{\hat{\boldsymbol{\sigma}}}_\Theta^{\text{opt}}(\mathbf{u}_\Theta^\varsigma)$ from equation (9.41).

4) Compute $F(\mathbf{u}_\Theta^\varsigma)$ and its slant derivative $G_\Theta(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{\text{opt}})$.

5) Solve $G_\Theta(\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{\text{opt}}) \left(\mathbf{u}_\Theta^{\varsigma+1} - \mathbf{u}_\Theta^\varsigma\right) = -F(\mathbf{u}_\Theta^\varsigma)$.

6) If $\|\mathbf{u}_\Theta^{\varsigma+1} - \mathbf{u}_\Theta^\varsigma\| \leq$tol stop otherwise return to step 2).

9.5. Numerical results

This section presents the numerical experiments for the BSB equation with a B-spline semi-discretization. The code developed for this problem is implemented in Matlab R2018b. The aim in this chapter is to analyze the discretization error and the convergence of the algorithm described in Section 9.4.2. In the following we consider a European butterfly-spread option with parameters choosen as in Example 9.10.

Example 9.10. *(European butterfly-spread option)*
Consider a European butterfly-spread option as defined in (2.3) with strike prices $K_1 = 9$, $K_2 = 11$ *and* $K = (K_1 + K_2)/2 = 10$ *and expire date* $T = 0.25$ *on the domain* $(S_{\min}, S_{\max}) = (4, 20)$. *The risk-free interest rate is* $r = 0.1$ *and the dividend yields is* $D_0 = 0$. *As not stated otherwise, the volatility is chosen as* $\sigma_{\min} = 0.15$ *and* $\sigma_{\max} = 0.25$.

To analyze the discretization error we compute a reference solution by applying Algorithm 9.3 with $\#\mathbb{T} = 2^{12}$ number of time steps and $N = 1056$ degrees of freedom. The numerical computations are obtained by Algorithm 9.1, Algorithm 9.2, and Algorithm 9.3. For all numerical computations we require a residual below 10^{-11} and the algorithms are stopped after maximal 12 iterations. Moreover, to study the spatial discretization error the number of time steps $\#\mathbb{T}$ are chosen sufficiently large (or time step sizes s are chosen sufficiently small) for all considered examples.
In Figure 9.1 we illustrate the approximation of Gamma with a cubic B-spline discretization applying Algorithm 9.1, Algorithm 9.2 and Algorithm 9.3. We note that Algorithm 9.1 provides an approximation of Gamma with small oscillations at these points where the volatility changes, while the approximation of Gamma is stabilized for Algorithm 9.2. Algorithm 9.3 enables an approximation of Gamma without discontinuities.
In order to analyze the convergence behavior, numerical computations are performed on a sequence of meshes with $s \approx \mathcal{O}(h^2)$. Figure 9.2 contains the results of computations for a cubic B-spline discretization. It can be observed, that the discretization error is significantly reduced when an algorithm with knot insertion is used. In particular, using Algorithm 9.3 where knots are inserted at the zeros of the second derivatives provides nearly optimal convergence rates in the $L^2(I)$-norm as in the case of variational equations.
The corresponding number of iteration steps for the semismooth Newton or semismooth Newton like method are presented in Table 9.1. It can be observed that Algorithm 9.3 needs more Newton iterations until convergence than Algorithm 9.1 and Algorithm 9.2. For each algorithm the number of iterations required for convergence varies slightly under refinement.
In Figure 9.3 and Figure 9.4 we compare the $L^2(I)$-error for today's butterfly-spread option and its first derivative Delta with cubic B-splines to a discretization with linear basis functions. As one can see, the error for cubic B-splines is significantly reduced in comparison to linear basis functions. Moreover, we observe in Figure 9.3, Figure 9.4 and Figure 9.5 that the convergence rates for the option price, Delta and Gamma are

optimal for linear basis function and nearly optimal for cubic basis functions.
A graphical illustration of today's approximated volatility in the worst case scenario is presented in Figure 9.6 (left). The results in Figure 9.6 (right) confirm that the volatility converges to the volatility obtained by the reference solution for increasing number of unknowns.
A plot of the butterfly-spread option in the worst and best case scenario in comparison to the Black-Scholes equation with constant volatility is shown in Figure 9.7. Let V_Θ^- and V_Θ^+ the approximated option price in the worst case and best case scenario; let further V_Θ the approximated option price in the standard Black-Scholes model with constant volatility $\sigma = (\sigma_{\min} + \sigma_{\max})/2$. One clearly observes $V_\Theta^- \leq V_\Theta \leq V_\Theta^+$ as expected in the uncertain volatility model.

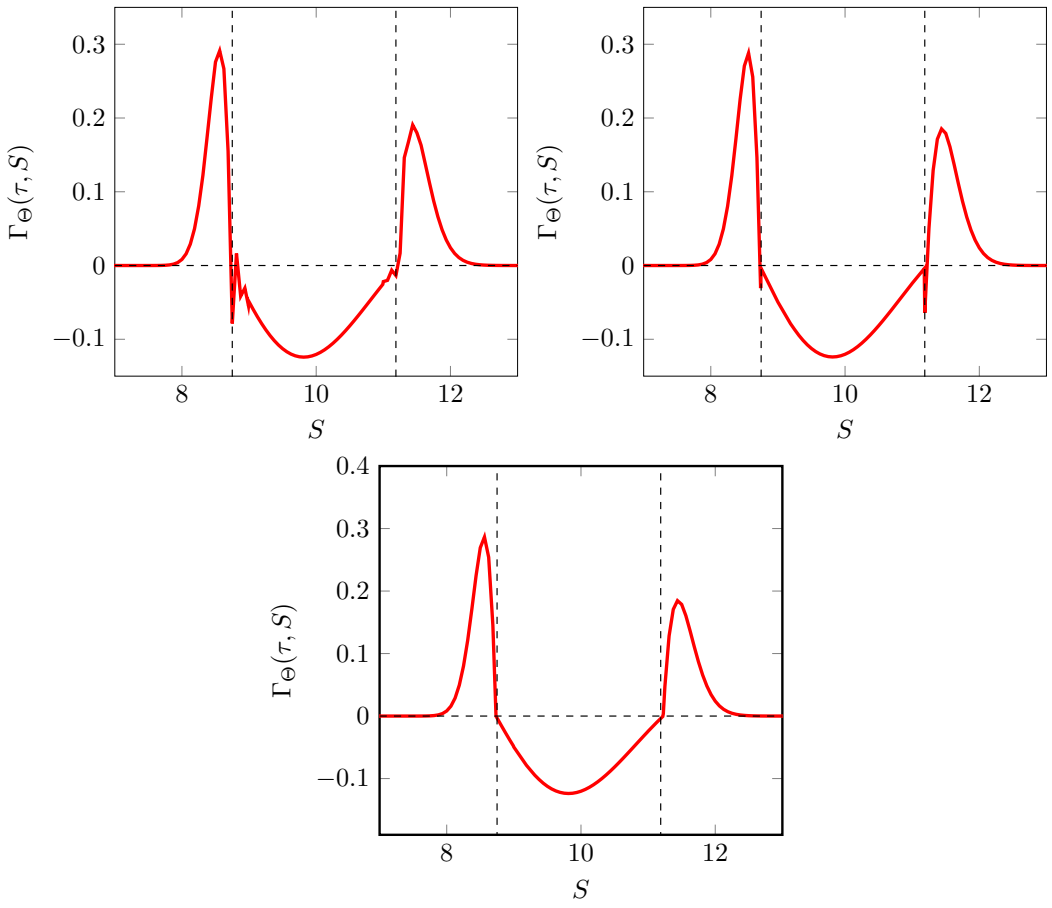

Figure 9.1.: Gamma $\Gamma_\Theta(\tau, S)$ in the worst case scenario with $\tau \approx 0.055$, $\Sigma = [0.15, 0.65]$ and cubic B-splines computed with Algorithm 9.1 (left above), Algorithm 9.2 (right above), and Algorithm 9.3 (below).

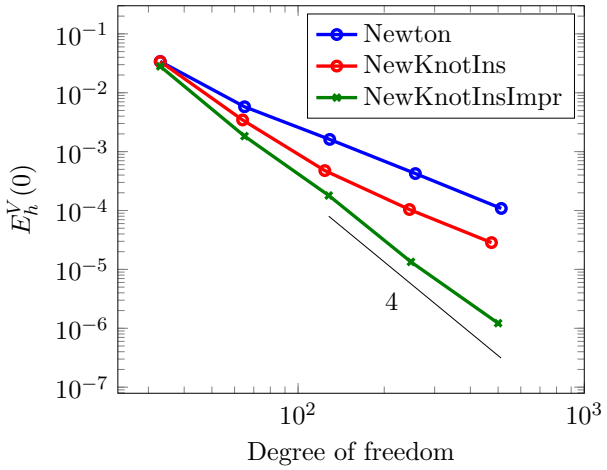

Figure 9.2.: $L^2(I)$-norm convergence of today's butterfly-spread option price in the worst case scenario computed with Algorithm 9.1 (*Newton*), Algorithm 9.2 (*NewKnotIns*) and Algorithm 9.3 (*NewKnotInsImpr*) for a cubic B-spline discretization and a BDF2 scheme with $s \approx \mathcal{O}(h^2)$.

	Newton		NewKnotIns		NewKnotInsImpr	
$\#\mathbb{T}$	N	$\overline{\#\text{it}}$	N	$\overline{\#\text{it}}$	N	$\overline{\#\text{it}}$
4	33	3.0	33	3.0	33	8.5
16	65	2.5	64	2.9	65	5.3
64	129	2.2	124	2.5	128	6.7
256	257	2.3	245	2.6	248	7.6
1,024	513	2.3	474	3.2	500	7.9

Table 9.1.: Number of Newton iterations for Algorithm 9.1 (*Newton*), Algorithm 9.2 (*NewKnotIns*) and Algorithm 9.3 (*NewKnotInsImpr*) with a cubic B-spline discretization and a BDF2 scheme for a butterfly-spread option in the worst case scenario.

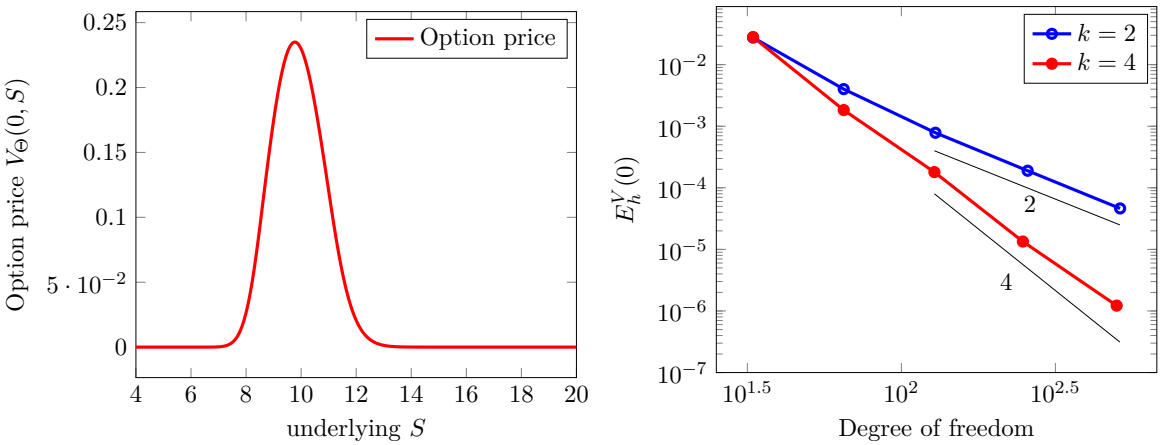

Figure 9.3.: Today's butterfly spread option price $V_\Theta(0,S)$ in the worst case scenario for cubic B-splines (left) and $L^2(I)$-error convergence for the BSB equation (right) with linear and cubic B-splines applying a BDF2-scheme with time step size $s \approx \mathcal{O}(h^2)$

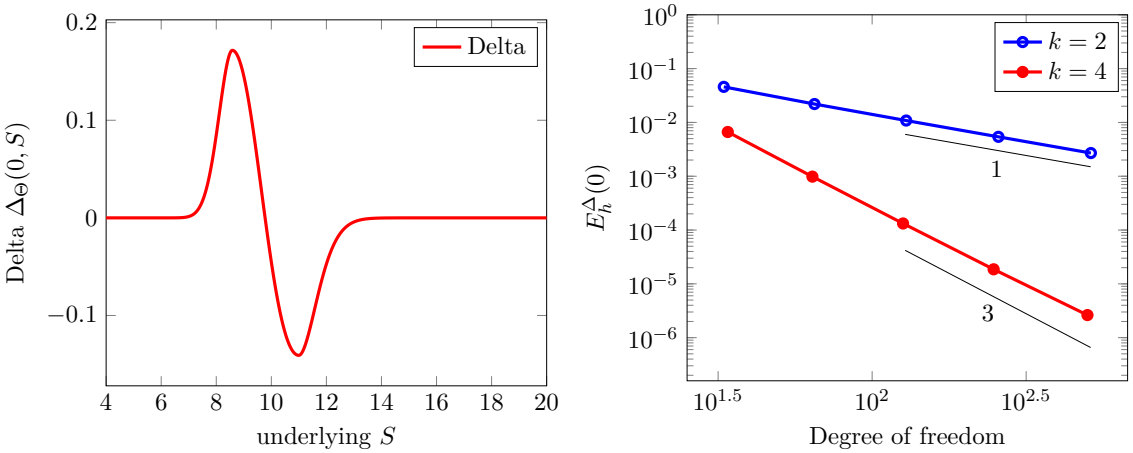

Figure 9.4.: First derivative of today's butterfly spread option price $\Delta_\Theta(0,S)$ in the worst case scenario, for cubic B-splines (left) and $L^2(I)$-error convergence of $\Delta_\Theta(0,S)$ for the BSB equation (right) with linear and cubic B-splines applying a BDF2-scheme with time step size $s \approx \mathcal{O}(h^{3/2})$.

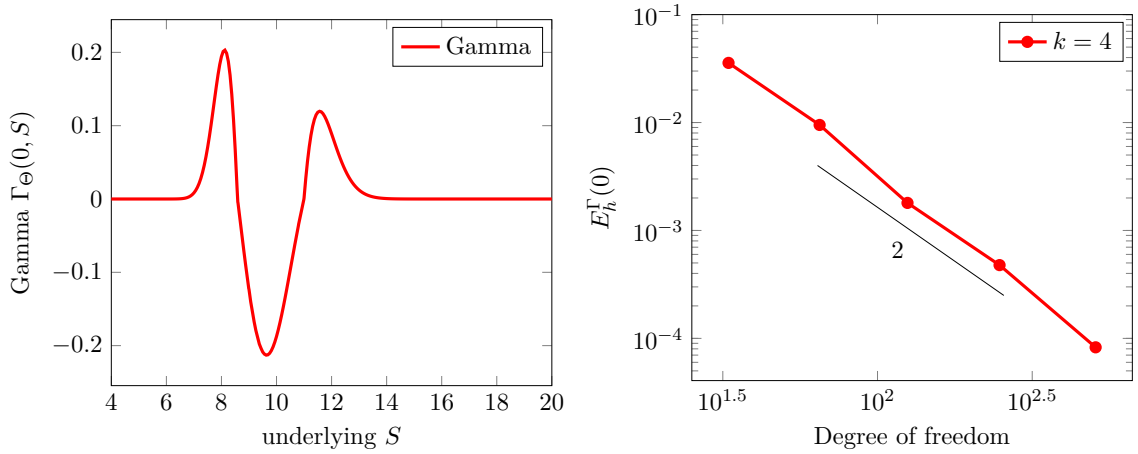

Figure 9.5.: Second derivative of today's butterfly spread option price $\Gamma_\Theta(0,S)$ in the worst case scenario, for cubic B-splines (left) and $L^2(I)$-error convergence of $\Gamma_\Theta(0,S)$ for the BSB equation (right) with linear and cubic B-splines applying a BDF2-scheme with time step size $s \approx \mathcal{O}(h)$.

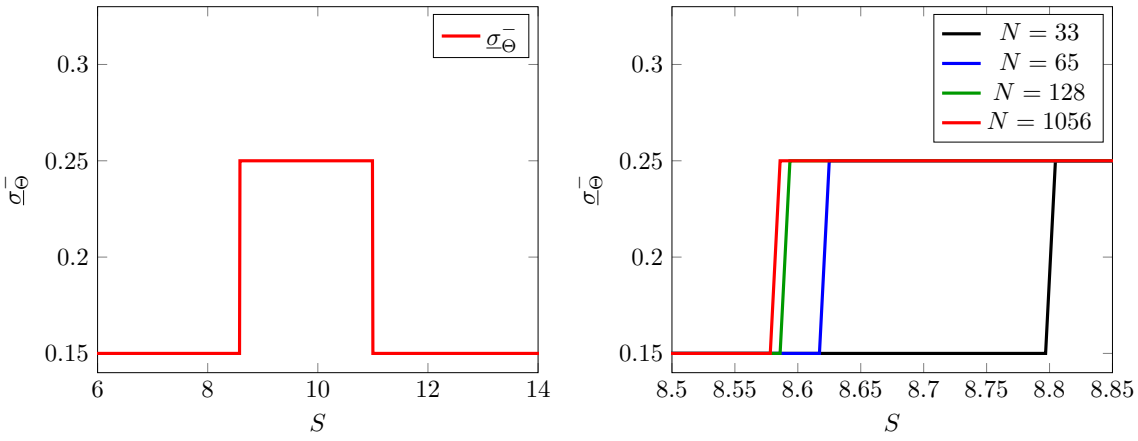

Figure 9.6.: Today's approximated volatility $\underline{\sigma}_\Theta^-$ in the worst case scenario computed with Algorithm 9.3 for cubic B-splines (left) and it's zoom in for different number of unknowns N (right) with $\sigma_{\min} = 0.15$ and $\sigma_{\max} = 0.25$ applying a BDF2 scheme with time step size $s \approx \mathcal{O}(h^2)$.

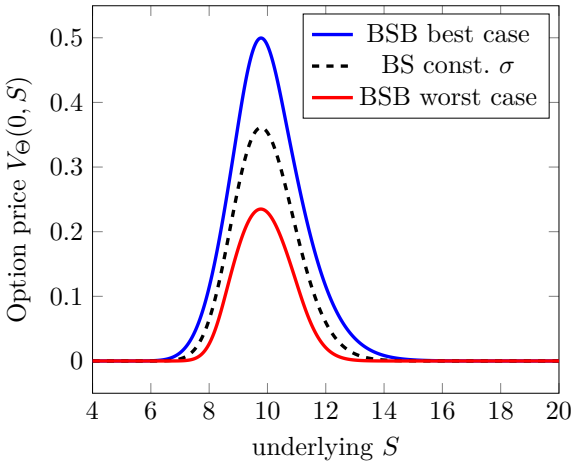

Figure 9.7.: Approximation of today's butterfly-spread option price $V_\Theta(0,S)$ with cubic B-splines in the Black-Scholes-Barenblatt model in the best and worst case scenario with $\sigma_{\min} = 0.15$ and $\sigma_{\max} = 0.25$. The dashed line corresponds to the standard Black-Scholes model with constant volatility $\sigma = 0.2$.

10. Conclusion and outlook

10.1. Conclusion

We have provided several numerical methods with (tensor product) B-splines for the pointwise highly accurate approximation of the Greeks, the partial derivatives up to order two, for the more challenging case of American option pricing problems. In particular, we have considered the American option price in Heston's and the Black-Scholes model formulated as a parabolic variational inequality. In this context we have proved the well-posedness of the Heston variational inequality derived in this thesis. One main feature of this thesis is the finding that due to the initial condition typically given as continuous piecewise linear payoff function a B-spline discretization with coinciding knots is preferable. More precisely, a cubic B-spline discretization with three coinciding knots at the points where the payoff function is not differentiable enables a pointwise highly accurate approximation of the Greeks up to order two. Several numerical experiments for linear, quadratic and cubic B-splines confirm that the optimal convergence rates, as expected for variational inequalities, are attained. Moreover, we have established a monotone multigrid method for (tensor product) B-splines with possible coinciding knots. In order to ensure the global convergence of the scheme we have constructed monotone coarse grid approximation for (tensor product) B-splines with possible coinciding knots. Corresponding numerical results have shown that the algebraic convergence rates of the MMG method are independent of the refinement level and mesh size. Moreover, it was shown that our provided MMG method facilitates a fast and highly accurate approximation of the American option price and its Greeks.

Furthermore, we have considered the pricing of an European option price with the Black-Scholes-Barenblatt equation. Based on the Cordes coefficients we have proved its well-posedness. In order to obtain a pointwise approximation of the European option price and its Greeks under the worst and best case volatility path we have discretized the BSB equation with higher order B-splines. Since a classical weak formulation in H^1 does not exist, the non-linear BSB equation is approximated in H^1 by L^1-normalized B-splines. It has turned out that the approximation of the volatility function leads to discontinuities in the partial derivatives of the numerical solution. In order to improve the approximation of the solution and its partial derivatives a semismooth Newton method within a knot insertion step was developed. Corresponding numerical results have shown that the convergence of the solution and its partial derivatives up to order two are nearly optimal in the L^2-norm as for variational equations when the volatility is approximated with desired accuracy.

10.2. Outlook

Numerical experiments for the constructed test problem for elliptic variational inequalities (see Test problem 8.1) have shown that a refinement in the neighborhood of the free boundary reduces the discretization error significantly. This clearly motivates the use of adaptive methods with hierarchical B-splines to improve the efficiency of the approach. To my knowledge adaptive methods were applied to variational inequalities in the context of option pricing problems with a linear finite element discretization (see e.g. [ZW09]) or to the Signorini problem (a special case of an elliptic variational inequality) for quadratic finite elements as well (see e.g. [Woh11]).
In this thesis we have noticed that the convergence rates of the monotone multigrid method depends on the B-spline order k. In current research, there are ideas for k-robust multigrid methods for variational equations, see e.g. [HT17]. This method is based on a stable splitting of the spline space into a subspace of interior splines which satisfy a robust inverse inequality, as well as one or several smaller subspaces which capture the boundary effects responsible for spectral outliers. Therefore they use an L^2-orthogonal B-spline splitting, which leads to the use of non positive L^2-orthogonal B-spline basis functions at the boundaries. For our constructed monotone multigrid method, which relies on the positivity of the B-spline basis functions, such L^2-orthogonal basis functions are not suitable due to the non positive basis functions. In particular, preserving the positivity for L^2-orthogonal basis function would destroy the optimal complexity of the monotone multigrid method. Thus, constructing a k-robust monotone multigrid method for obstacle problems would be interesting for future work.
A convergence theory for the Black-Scholes-Barenblatt equation for the proposed approach with a cubic B-spline discretization is still outstanding. The classical convergence analysis of Barles and Sougandis provides convergence of the solution to the viscosity solution if the method is stable, consistent and monotone. In particular, proving monotonicity for a higher order discretization is a challenging task. Another idea is to apply the discontinuous Galerkin approach for HJB equations of [Sme15] to the Black-Scholes-Barenblatt equation. This approach provides a convergence analysis, which is also applicable for higher order basis functions. To the best of my knowledge another convergence theory for HJB equation with a higher order discretization is not available. Moreover, it would be interesting to consider a cubic B-spline discretization of the Black-Scholes-Barenblatt equation in the multi-asset case introduced in [Lyo95].

A. Appendix

Knot insertion

In the following, we present an algorithm for computing the B-spline coefficients for a refined knot series from a given spline. This algorithm is needed for the knot insertion step in Algorithm 9.2 and Algorithm 9.3. It is important to note that knot insertion is only a change of the corresponding basis but the spline does not change. The results presented here are based on [LM08, PT97].
Let $\mathcal{S}(x) := \sum_{i=1}^{n} c_i N_{i,k}(x)$ define a spline of order k regarding a knot series Θ in the spline space $\mathbb{S}_{k,\Theta}$. Let further $\tilde{\Theta}$ a refined knot series of Θ, i.e. any real number occurs at least as many times in $\tilde{\Theta}$ as in Θ. Therefore one clearly has $\mathbb{S}_{k,\Theta} \subseteq \mathbb{S}_{\tilde{\Theta},k}$ (c.f. [LM08, Chapter 4.1, Lemma 4.2]). Thus, the spline $\mathcal{S}(x)$ also has a representation in the spline space $\mathbb{S}_{\tilde{\Theta},k}$, i.e. there exist B-spline coefficients $\tilde{\mathbf{c}} := (\tilde{c}_j)_{j=1}^{\tilde{n}}$ such that $\mathcal{S}(x) = \sum_{j=1}^{\tilde{n}} \tilde{c}_j N_{i,k,\tilde{\Theta}}(x)$. Moreover, each B-spline $N_{i,k,\Theta}(x) \in \mathbb{S}_{k,\Theta}$ can also be represented in the spline space $\mathbb{S}_{\tilde{\Theta},k}$, that means

$$N_{i,k,\Theta}(x) = \sum_{j=1}^{\tilde{n}} \alpha_{i,j,k} N_{j,k,\tilde{\Theta}}(x) \quad \text{for } i = 1, \ldots, n. \tag{A.1}$$

In matrix vector notation we can write

$$\mathbf{N}_{k,\Theta}^T = \mathbf{N}_{k,\tilde{\Theta}}^T K_{\tilde{\Theta}}, \tag{A.2}$$

with $\mathbf{N}_{k,\Theta}^T$ and basis transformation matrix $K_{\tilde{\Theta}} \in \mathbb{R}^{\tilde{n} \times n}$. The basis transformation matrix $K_{\tilde{\Theta}}$ is also called *knot insertion matrix* of order k from Θ to $\tilde{\Theta}$. Due to equation (A.2) the spline $\mathcal{S}(x)$ can be represented as

$$\mathcal{S}(x) = \mathbf{N}_{k,\tilde{\Theta}}^T \tilde{\mathbf{c}} = \mathbf{N}_{k,\Theta}^T \mathbf{c} = \mathbf{N}_{k,\tilde{\Theta}}^T K_{\tilde{\Theta}} \mathbf{c}. \tag{A.3}$$

Therefore, the linear independency of B-splines implies that the coefficient vectors must be related by $\tilde{\mathbf{c}} = K_{\tilde{\Theta}} \mathbf{c}$.
A powerful tool to add one knot to an existing knot series is the so called *Böhm's method*, which was developed by [Boe80]. This concept is provided in the next theorem. A proof using divided differences to determine the B-spline coefficients of the splines can be found in [LM08, Chapter 4.8, Lemma 4.19].

Theorem A.1. *(Böhm's method)*
Let $\Theta := (\theta_i)_{i=1}^{n+k}$ be a given knot vector and let $\tilde{\Theta} := (\tilde{\theta}_j)_{j=1}^{n+k+1}$ be the knot vector obtained by inserting a knot $\tilde{x}$ in Θ in the interval $[\theta_\wp, \theta_{\wp+1})$. If

$$\mathcal{S}(x) = \sum_{i=1}^{n} c_i N_{i,k,\Theta}(x) = \sum_{j=1}^{\tilde{n}} \tilde{c}_j N_{i,k,\tilde{\Theta}}(x) \tag{A.4}$$

then $\tilde{c}_j$ can be expressed in terms of c_i by the formulas

$$\tilde{c}_j = \begin{cases} b_j & \text{if } 1 \le j \le \wp - k + 1, \\ \frac{\tilde{x}-\theta_j}{\theta_{j+k-1}-\theta_j} c_j + \frac{\theta_{j+k-1}-\tilde{x}}{\theta_{j+k-1}-\theta_j} c_{j-1} & \text{if } \wp - k + 2 \le j \le \wp, \\ c_{j-1} & \text{if } \wp + 1 \le j \le n + 1. \end{cases} \tag{A.5}$$

For implementation, it is useful to rewrite (A.5) in matrix vector notation. Let us consider $\tilde{\mathbf{c}} = K_{\tilde{\Theta}}\mathbf{c}$ where $K_{\tilde{\Theta}}$ is the knot insertion matrix for inserting one knot, then (A.5) can be expressed in matrix vector notation with the matrix $K_{\tilde{\Theta}} \in \mathbb{R}^{(n+1)\times n}$ given by

$$K_{\tilde{\Theta}} := \begin{pmatrix} 1 & 0 & & & & & & & \\ & \ddots & \ddots & & & & & & \\ & & 1 & 0 & & & & & \\ & & 1-\lambda_{\wp-k+2} & \lambda_{\wp-k+2} & & & & & \\ & & & \ddots & \ddots & & & & \\ & & & & 1-\lambda_\wp & \lambda_\wp & & & \\ & & & & 0 & 1 & & & \\ & & & & & \ddots & \ddots & & \\ & & & & & & 0 & 1 \end{pmatrix}, \tag{A.6}$$

where $\lambda_j := (\tilde{x} - \theta_j)/(\theta_{j+k-1-\theta_j})$ for $\wp - k + 2 \le j \le \wp$. The knot insertion matrix is bi-diagonal and all entries are positive since $\tilde{x}$ lies in the interval $[\theta_\wp, \theta_{\wp+1})$.

The following corollary generalizes the knot insertion matrix to insert several knots. This result can be obtained by applying Böhm's method $\tilde{n} - n$ times. A detailed proof can be found in [LM08, Chapter 4.5, Lemma 4.20].

Corollary A.2. *Let $\Theta := (\theta_i)_{i=1}^{n+k}$ and $\tilde{\Theta} := (\tilde{\theta}_j)_{j=1}^{\tilde{n}+k}$ be two knot vector for splines of order k with $\Theta \subseteq \tilde{\Theta}$. Then all the entries of the knot insertion matrix $K_{\tilde{\Theta}}$ from $\mathbb{S}_{k,\Theta}$ to $\mathbb{S}_{\tilde{\Theta},k}$ are nonnegative and $K_{\tilde{\Theta}}$ can be factored as*

$$K_{\tilde{\Theta}} := K_{\tilde{\Theta},\tilde{n}-n} K_{\tilde{\Theta},\tilde{n}-n-1} \cdots K_{\tilde{\Theta},1}, \tag{A.7}$$

where $K_{\tilde{\Theta},j} \in \mathbb{R}^{(n+j)\times(n+j-1)}$ is a bi-diagonal matrix with nonnegative entries.

Knot removal

Knot removal means exactly what the name suggests, namely removing knots from an existing knot series. One particular difficulty arises from the fact that not all knots are removable since the spline can change geometrically when knots are removed. Thus it is necessary to construct a method, which determines if a knot is removable and if so, how many times. Moreover, this method must compute the new spline coefficient vector from the existing coefficient vector. The subsequent method presented here are from [PT97, Chapter 5.4], where the authors discuss knot removal in the context of Nurbs. Since Nurbs are only a generalization of B-splines, the results and methods can be easily transferred to B-splines.
Let us consider the spline associated with the given knot series Θ

$$\mathcal{S}(x) := \sum_{i=1}^{n} c_i N_{i,k,\Theta}(x). \tag{A.8}$$

Let further $\tilde{\Theta}$ denote the knot series obtained by removing the knot $\theta_\wp$ b times from Θ. Then $\theta_\wp$ is called b times *removable* if

$$\mathcal{S}(x) = \sum_{i=1}^{n-b} \tilde{c}_i N_{i,k,\tilde{\Theta}}(x). \tag{A.9}$$

Considering the properties of B-splines in Properties 3.16 it is known that a spline at the knot $\theta_\wp$ with multiplicity q has $k-q-1$ continuous derivatives at this knot. But for some positions of the coefficients even higher continuous derivatives are possible. Thus a knot with multiplicity q is b times removable if the spline has $k-q-1+b$ continuous derivatives for $1 \leq q \leq k-1$. Hence, by exploiting the formula in (3.70) for the derivatives of splines at the knot $\theta_\wp$ and its neighbouring knots the following formula is derived for removing one knot $x = \theta_\wp \neq \theta_{\wp+1}$ with multiplicity q. Let us assume that $\tilde{c}_i^{(0)} := c_i$ then the formulas for the new coefficients are

$$\tilde{c}_i^{(1)} = \frac{\tilde{c}_i^{(0)} - (1-\alpha_i)\tilde{c}_{i-1}^{(1)}}{\alpha_i} \quad \text{for } \wp - k + 1 \leq i \leq \frac{1}{2}(2\wp - k - q) \tag{A.10}$$

$$\tilde{c}_j^{(1)} = \frac{\tilde{c}_j^{(0)} - \alpha_j \tilde{c}_{j+1}^{(1)}}{(1-\alpha_j)} \quad \text{for } \frac{1}{2}(2\wp - k - q + 3) \leq j \leq \wp - q \tag{A.11}$$

with

$$\alpha_\mathfrak{s} = \frac{x - \theta_\mathfrak{s}}{\theta_{\mathfrak{s}+k} - \theta_\mathfrak{s}} \text{ for } \mathfrak{s} = i, j. \tag{A.12}$$

A detailed derivation of the formulas can be found in [PT97, Chapter 5.4]. For the implementation in Matlab we have modified the Nurbs-code from [SCdFV] for B-splines.

B. Code documentation

This chapter provides a brief overview of the Matlab package, which was implemented in connection with this thesis. As not stated otherwise the functions are implemented by myself.
We start by presenting the code developed for the numerical solution for the Black-Scholes and Heston variational inequality with a B-spline discretization of order k. The implementation allows to discretize the Black-Scholes or Heston-Variational inequality for arbitrary (tensor product) B-spline order k. Due to the low regularity of the solution at the free boundary, a linear ($k = 2$), quadratic ($k = 3$) or cubic ($k = 4$) B-spline discretization is recommended. In particular in the context of option pricing problems the most relevant discretization is the cubic B-spline discretization to facilitate a pointwise approximation of the Greeks up to order two.
Although we concentrate on specific payoff functions in this thesis, the Black-Scholes or Heston variational inequality can be discretized for any once weak differentiable payoff function by simple changing the function in Matlab to another payoff function. Clearly, in this context the user has to take into account the truncation effects in the respective model.
The implementation of the B-splines relies on its recursive definition, the Neville-like scheme in Theorem 3.18 for its point evaluation and its expression as derivatives from Corollary 3.17. This functions were implemented by [Mol16]. In order to compute the Greeks, the derivatives of the spline solution are implemented in this thesis according to Corollary 3.19. All these functions works for arbitrary B-spline order k and possible non-uniform sequences of knots with coinciding internal knots as well.
Since B-splines are piecewise polynomials of order k on $[\theta_\mathfrak{r}, \theta_{\mathfrak{r}+1}]$ the entries of the discretization matrices can be computed exactly with the Gauss-Legendre quadrature rule for sufficiently high order. Thus, the computation of the integrals over the diff'th derivative of the B-splines and some polynomial $a(x)$ are performed as follows

$$B_{i,j} = \int_{\mathcal{Q}} a(x) N_{j,k}^{(\mathsf{diff}_1)}(x) N_{i,k}^{(\mathsf{diff}_2)}(x) \mathrm{d}x \tag{B.1}$$

$$= \sum_{\mathfrak{r}=\max(i,j)}^{\min(i,j)+k-1} \int_{\theta_\mathfrak{r}}^{\theta_{\mathfrak{r}+1}} a(x) N_{j,k}^{(\mathsf{diff}_1)}(x) N_{i,k}^{(\mathsf{diff}_2)}(x) \mathrm{d}x. \tag{B.2}$$

Moreover, the Gauss points are given as the roots of the Legendre polynomials on the interval $[-1,1]$. As a consequence transformation from $[\theta_{\mathfrak{r}},\theta_{\mathfrak{r}+1}]$ to the interval $[-1,1]$ with the variable $\tilde{x}^{\mathfrak{r}} := (\theta_{\mathfrak{r}+1}+\theta_{\mathfrak{r}})/2 + (\theta_{\mathfrak{r}+1}-\theta_{\mathfrak{r}})/2\tilde{x}$ and $c^{\mathfrak{r}} := (\theta_{\mathfrak{r}+1}-\theta_{\mathfrak{r}})/2$ yields

$$\begin{aligned} B_{i,j} &= \sum_{\mathfrak{r}=\max(i,j)}^{\min(i,j)+k-1} \int_{-1}^{1} a(\tilde{x}^{\mathfrak{r}}) N_{j,k}^{(\mathsf{diff}_1)}(\tilde{x}^{\mathfrak{r}}) N_{i,k}^{(\mathsf{diff}_2)}(\tilde{x}^{\mathfrak{r}}) c^{\mathfrak{r}} \, \mathrm{d}\tilde{x} \\ &= \sum_{\mathfrak{r}=\max(i,j)}^{\min(i,j)+k-1} \sum_{l=1}^{\tilde{\mathfrak{n}}} \tilde{w}_l^{\mathfrak{r}} \; a(\tilde{x}_l^{\mathfrak{r}}) N_{j,k}^{(\mathsf{diff}_1)}(\tilde{x}_l^{\mathfrak{r}}) N_{i,k}^{(\mathsf{diff}_2)}(\tilde{x}_l^{\mathfrak{r}}), \end{aligned} \tag{B.3}$$

where $\tilde{w}_l^{\mathfrak{r}} := w_l c^{\mathfrak{r}}$ are the Gauss-Legendre weights and $\tilde{x}_l^{\mathfrak{r}}$ the corresponding Gauss-Legendre points on the interval $[\theta_{\mathfrak{r}},\theta_{\mathfrak{r}+1}]$. For computing `N` Gauss-Legendre points and weights on an interval $[a,b]$ the Matlab function `lgwt[N,a,b]` from [vW] is used.
Due to the tensor product structure, the discretization of the bilinear forms in higher dimensions can be expressed by the Kronecker product of the one dimensional discretization matrices. In this thesis, the discretization matrices and right hand side for the Heston variational inequality are implemented in terms of Kronecker products according to the formulas in (5.43) and (5.45).
In order to solve the discrete variational inequalities arising from a (tensor product) B-spline discretization an implementation of the MMG method in Matlab is proposed. This function works for arbitrary B-spline order $k = 2,3,4$ and a uniform knot series with possible internal coinciding knots as well. The corresponding restriction and prolongation operators are build up automatically for a uniform knot series with $k-1$ coinciding knots at $\theta_{\wp}$. Since for option pricing problems payoff functions are typically given as piecewise continuous functions with discontinuity in the partial derivative at the strike price K, this facilitates the numerical computation of the American option price and its Greeks in the Black-Scholes or Heston model for payoff functions with one strike price K. Moreover, restriction and prolongation operators are proposed for a payoff function with three discontinuities at $\theta_{\wp_1}, \theta_{\wp_2}$ and $\theta_{\wp_3}$, this, for example, enables the numerical computation of the American butterfly-spread option price.
For the a priori smoothing steps in the MMG method an implementation of the PSOR method from Section 7.1 is needed. Due to the projection step in the PSOR approach, the iteration procedure can not be vectorized, thus the execution in Matlab is very slow. As a consequence the PSOR method is implemented as a C Mex-function, which provides a gateway to Matlab. This code implemented in C runs much faster than the corresponding code in Matlab.
Now, the functionality of the most relevant functions are specified. To simplify the handling, the code for the Black-Scholes and Heston variational are located in separate directories. Note, that the specific implementation for the respected problem also optimized the run time of several operations. We start by the documentation of the B-spline code for the Black-Scholes variational inequality.

B-spline code for the Black-Scholes variational inequality:

- `BS_main`: In this script Example 8.2, Example 8.3, Example 8.6 and Example 8.4 are implemented. The user can specify the following input parameters:
  - Refinement level `j`, lowest level of the MMG method `lev_end`-1,
  - B-spline order `k`, only the parameters `k`=2,3,4 make sense,
  - number of time steps `timestep`,
  - number of a priori and a posteriori smoothing steps `eta` as 2×1 `array`,
  - type of the payoff function `payoff`; possible choices are `'put'` with parameters as in Example 8.2, `'butterfly'` with parameters as in Example 8.4, `'call1'` with parameters as in Example 8.3, `'call2'` with parameters as in Example 8.6,
  - choice of the solution algorithm `method`; possible choices are `'MMG'` (monotone multigrid method), `'MMG_Nest'` (monotone Multigrid with initial value obtained by Nested iteration) and `'PGS'` (projective Gauss-Seidel method),
  - type of the time stepping method `choice`, possible choices are `'euler'` (implicit Euler method), `'ran'` (Rannacher timestepping method), `'crank'` (Crank-Nicolson method)
  - repetition of knots `repetition`, `'yes'` corresponds to a uniform knot series where $k-1$ knots are repeated at the strike price/s, `'no'` corresponds to a uniform knot series without internal coinciding knots,
  - tolerance `tol_f` for the stopping criteria as anonymous function (`@` function).

  Output: $L_2(I)$-error of today's American option price, its first derivative Delta and its second derivative Gamma, number of iteration steps needed until convergence with the PGS or MMG method, averaged algebraic convergence rate and averaged cpu time needed to solve the complementarity problem in one time step, plot of today's approximated American option price, its Delta and Gamma.
  It is also possible for the user to change the parameters from the Black-Scholes model, but then the error regarding the reference solution is no longer valid.

- `produces_system`: Function to assemble the discretization matrices and right hand side for the Black-Scholes model. The integrals in (4.37) are computed exactly by applying the Gauss-Legendre quadrature rule according to (B.3). The integrals for the right hand side in (4.38) are obtained by a simple modification of (B.3).
  Input:
  - number of B-spline basis functions within the first and last B-spline `N`, B-spline order `k` and knot series `theta`,

– strike price `K` for put and call options or strike prices `[K1,K,K2]` for a butterfly-spread option, volatility `sigma`, interest rate `r`, dividend yields `d`, payoff function and it's first derivative as anonymous functions,

– time step size `s` and weighted average for the time stepping method `varpi`, possible choices are `0.5` (Crank-Nicolson method) and `1.0` (implicit Euler method).

Output: discretization matrices and ride hand side for the implicit Euler and Crank-Nicolson method.

- `psor_euler_crank`, `psor_ran`: In this functions the PSOR method is used to solve the fully discrete complementarity problem applying an implicit Euler, Crank-Nicolson or Rannacher timestepping method.
- `MMG_ran`, `MMG_euler_crank`: In this functions the implicit Euler, Crank-Nicolson or Rannacher timestepping method is applied to solve the linear complementarity problem with the MMG algorithm in each time step.
- `produce_restriction`: Helpfunction to initialize the restriction operator for linear, quadratic and cubic B-splines with coinciding knot/s at the strike prices. The restriction operators are saved in a `cell`-array for each level. In order to accelerate calculations the restricted discretization matrices `C_cell` ($\mathcal{R}C_\ell\mathcal{P}$) are also saved in a cell array for each level.
- `MMG`: Monotone multigrid algorithm to solve the linear complementarity problem in one time step.

Input: `C_cell` restricted discretization matrices for each level implemented as cell-array and its diagonal part for each level `D_cell`. In order to accelerate the matrix vector multiplication in C one needs also the transposed restricted discretization matrices implemented as cell-array `C_cellt`, right hand side `f` on the highest level, initial value `u0`, highest level `lev`, lowest level `lev_end` , B-spline coefficient vector for the obstacle on the highest level `obst`, a priori and a posteriori smoothing steps `eta1` and `eta2`, B-spline order `k`, maximum number of iteration steps `max`, tolerance for the stopping criteria `tol` and `inter` number of coinciding knots.

Output: B-spline expansion coefficients vector `u` for the approximated solution in one time step, number of iteration steps `iter` needed until convergence according the stopping criteria $\|\mathtt{u} - \mathtt{u0}\|_{\ell_2} < \mathtt{tol}$.

- `upper_coarse_grid`: In this function the monotone upper coarse grid approximation in one dimension out of Corollary 7.4 is implemented.
  Input: obstacle `obst`, level `lev`, B-spline order `k`, number of coinciding knots `inter`.
  Output: coefficient vector `q` for the monotone upper coarse grid approximation.

- `L_2`: Helpfunction to compute the $L_2(I)$-error on the interval I. The integrals are computed with the Gauss-Legendre-quadrature rule on $[\theta_i, \theta_{i+1}]$.
  Input: Approximated B-spline Galerkin solution in one time step as anonymous function or its derivatives, B-spline order `k`, knot series `theta`, order of differentiation `diff`.
  Output: Error in the L_2-norm on the interval I.

- `Hedge_plot`: Routine for plotting the spline solution (approximated option price) and its derivatives (Greeks) in one time step. Moreover, the free boundary in one time step is approximated.
  Input: expansion coefficient vector, B-spline order `k`, knot series `theta`, strike price `K` for put and call options or strike prices `[K1,K,K2]` for a butterfly-spread option, refinement level `j`, payoff function and its derivatives as anonymous function, type of the payoff function `payoff` (string).
  Output: Plot of the American option price in one time step, its Delta and Gamma and corresponding data saved in a `dat`-file, approximated free boundary `Sf_h` in one time step.

Tensor product B-spline code for the Heston variational inequality:

- `Heston_main`: In this script Example 8.5 is implemented. The user can specify the following input parameters:
  - B-spline order `k`,
  - refinement level `j` with multiindices `[jx,jv]` in the x- and v-direction,
  - number of time steps,
  - type of the time stepping method `t_choice`, possible choices are `'euler'` (implicit Euler method), `'ran'` (Rannacher timestepping method), `'crank'` (Crank-Nicolson method),

- choice of the solution algorithm `method`; possible choices are `'MMG'` (monotone multigrid method), `'MMG_Nest'` (monotone Multigrid with initial value obtained by Nested iteration) and `'PGS'` (projective Gauss-Seidel method),
- lowest level of the MMG method `lev_end-1`,
- number of a priori and a posteriori smoothing steps `eta` as 2×1 `array`,
- repetition of knots `repetition`, `'yes'` corresponds to a uniform knot series where $k-1$ knots are repeated at the strike price, `'no'` corresponds to a uniform knot series without internal coinciding knots,
- tolerance `tol_f` of the stopping criteria as anonymous function.

Output: $L^2(\Omega)$-error of today's American put option price, its Delta, Vega and Gamma, averaged algebraic convergence rate and averaged cpu time needed to solve the complementarity problem in one time step, plot of today's approximated option price, its Delta, Vega and Gamma.
It is also possible for the user to change the parameters from the Heston model, but then the error regarding the reference solution is no longer valid.

- `produces_system`: Function to assemble the discretization matrices and right hand side for the Heston model in terms of Kronecker products of the discretization matrices in the x- and v-direction according to the formulas in (5.43) and (5.45).
Input:
  - Number of B-spline basis function `N` and `M` in the x- and v-direction within the first and last B-spline,
  - B-spline order `k` and the corresponding uniform knot series `theta_x` and `theta_v` with possible internal coinciding knots,
  - choice of the boundary condition `bound_v0`; `bound_v0=0` corresponds to the set of boundary conditions in (2.93) (not recommended since it leads to unstable approximation of Vega, see Section 8.2.1) and `bound_v0=1` corresponds to the set of boundary conditions in (2.94),
  - parameters for the Heston model `rho,xi,gamma,r` and `kappa`,
  - payoff function and its derivative as anonymous function,
  - time step size `s` and weighted average `varpi` for the time discretization.

  Output: Discretization matrix and discrete right hand side for the implicit Euler and Crank-Nicolson method.

- `produces_matrices_x`, `produces_matrices_v`, `produces_vector_x` and `produces_vector_v`: Helpfunctions to assemble the discretization matrices in the x- and v-direction from (5.49). The entries are calculated according to (B.3). The integrals for the right hand side in (5.51) are computed by a simple modification of (B.3).

- `psor_euler_crank`, `psor_ran`: In this functions the PSOR method is used to solve the fully discrete complementarity problem applying an implicit Euler, Crank-Nicolson or Rannacher timestepping method.
- `MMG_ran`, `MMG_euler_crank`: In this functions the implicit Euler, Crank-Nicolson or Rannacher timestepping method is applied to solve the linear complementarity problem with the MMG algorithm in each time step.

- `produce_restriction`: Helpfunction to initialize the restriction operator for linear, quadratic and cubic tensor product B-splines with possible coinciding knots at the strike price. The restriction operators are implemented as Kronecker products and are saved in a `cell`-array for each level. In order to accelerate calculations the restricted discretization matrices `C_cell` ($\mathcal{R}C_\ell\mathcal{P}$) are also saved in a cell array for each level.
- `MMG`: Monotone multigrid algorithm to solve the linear complementarity problem in one time step.
  Input:
  - restricted discretization matrices `C_cell` for each level implemented as cell-array and its diagonal part for each level `D_cell`, in order to accelerate the matrix vector multiplication in C one needs also the transposed restricted discretization matrices implemented as cell-array `C_cellt`,
  - right hand side `f` on the highest level,
  - lowest level `lev_end-1`, B-spline order `k` and number of coinciding knots `inter`,
  - tensor product B-spline coefficient vector for the obstacle `obst` on the highest level in lexicographical order $\texttt{obst}_{i_x,i_v}$, initial value `u0`,
  - a priori and a posteriori smoothing steps `eta1` and `eta2`,
  - maximum number of iteration steps `max`, stopping criteria `tol`,
  - choice of the boundary condition `bound_v0`.

  Output: Tensor product B-spline expansion coefficients vector `u` for the approximated solution (in lexicographical order) in one time step, number of iteration steps needed until convergence according the stopping criteria $\|\texttt{u} - \texttt{u0}\|_{\ell_2} < \texttt{tol}$.
- `coarse_grid2d`: In this function the monotone upper coarse grid approximation in two dimensions for a uniform knot series with possible coinciding knots in the x-direction is implemented, see Corollary 7.4.
  Input: expansion coefficients of the obstacle function `obst` in lexicographical order, level in the x- and v-direction `lev_x` and `lev_v`, B-spline order `k`, number of coinciding knots `inter`, choice of the boundary condition `bound_v0`.
  Output: coefficient vector `q` in lexicographical order for the monotone upper coarse grid approximation.

- `L_2_2d`: Helpfunction to compute the $L_2(\Omega)$-error on the two dimensional domain Ω. The integrals are computed with the Gauss-Legendre-quadrature.
  Input: Approximated tensor product B-spline solution in one time step or its derivatives as anonymous function, B-spline order `k`, knot series `theta_x` and `theta_v`, order of differentiation in the x-direction `diff_x` and in the v-direction `diff_v`.
  Output: Error in the $L_2(\Omega)$-norm.

- `Hedge_plot`: Routine for plotting the tensor product spline solution (approximated option price) and its derivatives (Delta, Vega and Gamma). Moreover, the free boundary in one time step is approximated.
  Input: expansion coefficient vector `u` in lexicographical order, knot series `theta_x` and `theta_v`, B-spline order `k`, strike price `K`, payoff function and its derivative as anonymous function.
  Output: Plot of the approximated option price, Delta, Vega, Gamma and the approximated free boundary in one time step saved in a `pdf`-file.

Now, a short overview of the B-spline code for the Black-Scholes-Barenblatt equation is given. In particular, the implementations of Algorithm 9.1, Algorithm 9.2 and Algorithm 9.3 are described. The standard semismooth Newton method is implemented for arbitrary B-spline order $k = 2, 3, 4$ but due to the poor approximation of the volatility function defined in (9.39) increasing the order of the B-spline basis functions does not improve the global discretization error. The semismooth Newton method with knot repetition of existing knots works also for B-splines of order $k = 2, 3, 4$. But, for the same reason as before increasing the order improves only (very) slightly the global discretization error. In Algorithm 9.3 the strong form of the approximated volatility function is needed, to compute the roots of the second derivative. Thus, the corresponding implementation in Matlab is restricted to a cubic ($k = 4$) B-spline discretization. In the following, the functionality of the most relevant functions are specified. The code is implemented for the Black-Scholes-Barenblatt equation for an European butterfly-spread option in the worst case scenario. The corresponding option price in the best case scenario can be obtained by changing the minimal value for the volatility `sigma_min` to the maximal value for the volatility `sigma_max` and vice versa.

B-spline Code for the Black-Scholes-Barenblatt equation

- `BSB_Algo1`: Standard semismooth Newton algorithm as stated in Algorithm 9.1 to solve the BSB equation with the approximated volatility function in (9.39). In this script Example 9.10 is implemented.
  Input: refinement level `j`, B-spline order `k` (for reasonable choices see the description above).
  Output: $L^2(I)$-error of today's option price, its Delta and Gamma, averaged number of Newton iteration steps per time step, plot of today's option price, its Delta and Gamma.

- `volatility1`: Function to assemble the volatility function (weak form) in (9.39). This function also computes the locations of volatility changes for a butterfly-spread option.

- `BSB_Algo2`: Semismooth Newton Algorithm with knot repetition of existing knots as stated in Algorithm 9.2.
  Input: refinement level `j`, B-spline order `k` (for reasonable choices see the description above) to solve the BSB equation with the approximated volatility function in (9.39).
  Output: $L^2(I)$-error of today's butterfly-spread option price, its Delta and Gamma, averaged number of Newton iteration steps per time step, plot of today's option price, its Delta and Gamma.

- `BSB_Algo3`: Semismooth Newton Algorithm with knot insertion at the roots of the second derivative as stated in Algorithm 9.3.
  Input: refinement level `j`.
  Output: $L^2(I)$-error of today's butterfly-spread option price, its Delta and Gamma, averaged number of Newton iteration steps per time step, plot of today's option price, its Delta and Gamma.

- `volatility2`: Function to assemble the volatility function (strong form) in (9.41). This function also returns the roots of the second derivative for a butterfly-spread option.

- `roots_sec`: Helpfunction to approximate the roots of the second derivative by the Matlab function `fzero`. For this function an initial interval for the bisection method has to be specified. To find the two roots for the butterfly spread option in the middle the initial interval has to be taken carefully. Therefore, the matlab function `fmin_search` is used to find the two maxima and the minimum of Gamma, which serve as good initial values to find the two relevant roots of the second derivative. Note, that also `fminsearch` needs reasonable initial values, which vary for different model parameters.

- `prepare_ref`: Helpfunction to compute the corresponding knot insertion matrix with Böhm's method. In particular, this function checks if there are `k-1` knots `z1` or `z2` at the `mue1+1`'th or `mue2+1`'th position. If there are no knots the function returns the identity matrix, otherwise it computes the corresponding knot insertion matrix to insert so many knots `z1` and `z2` as necessary for having `k-1` coinciding knots `z1` and `z2` in the knot series. This function works for quadratic and cubic B-splines. It also returns the corresponding new knot series and the new approximated volatility function.

- `RemoveCurveKnot`: Helpunction to remove unnecessary knots from the existing knot series for B-splines. A detailed explanation can be found in [PT97, Chapter 5.4] or Appendix A. For the implementation in Matlab we have slightly modified the Nurbs-code from [SCdFV] for B-splines.

General functions

- `Nev`: Evaluates a d-dimensional spline at the point `x=[x1,..,xd]` according to a multidimensional version of the Neville-like scheme in Theorem 3.18. This function was implemented by [Mol16].

Input: Multidimensional vectorized expansion coefficient `c` in lexicographical order, extended sequence of knots in each direction in a `cell` array, B-spline order `k=[k1,..,kd]` in each direction, evaluation point `x=[x1,..,xd]`.
Output: Point evaluation `Val` of the corresponding multidimensional spline at the point `x=[x1,..,xd]`.

- `Ndiff`: Helpfunction to evaluate the `diff`'th derivative of the B-spline $N_{i_j,k}(x_j)$ in one coordinate at the point x_j, implemented by [Mol16].

  Input: knot series in each direction in a `cell` array, B-spline order `k`, evaluation point, order of differentiation `diff`.

  Output: Point evaluation `Val` of the `diff`'th derivative of the B-spline $N_{i_j,k}(x_j)$.

- `psor.c`: The function can be compiled in Matlab2018b by the command `mex -largeArrayDims psor.c`. This function solves a linear complementarity problem via the PSOR method described in Section 7.1. In order to speed up the matrix vector multiplication and to permit the transposition of the discretization matrix in each time step of the time stepping method the user has to specify the discretization matrix in transposed form. Moreover, the function works only for matrices in `sparse` format. Note, that this algorithm is a standard method to solve LCP's and many versions of implementations are available.

  Input:

  – transposed matrix `Ct` of the discretization matrix C in `sparse` format,

  – diagonal part of the `N`×`N` discretization matrix in a `N`×1 double array and the discrete right hand side in a `N`×1 double array,

  – (tensor product) B-spline expansion coefficients of the obstacle in lexicographical order, initial iterate `u0` in a `N`×1 double array,

  – relaxation parameter `om` for the PSOR method, `om=1` corresponds to the projective Gauss-Seidel method,

  – maximal number of iteration steps `iter_max` as integer number, tolerance for the stopping criteria `tol` as double number.

  Output: B-spline expansion coefficients for the approximated solution `u` in one time step (or solution of the LCP), number of iteration steps needed until convergence according the stopping criteria $\|\mathtt{u} - \mathtt{u0}\|_{\ell_2} < \mathtt{tol}$.

C. Symbols

Notations for (tensor-product) B-splines:

$\Theta := \{\theta_i\}_{i=1}^{n+k}$	ordered sequence of knots
$N_{i,k}(x)$	B-spline of order $k \in \mathbb{N}$
$\mathbb{S}_{k,\Theta}$	space of splines spanned by the basis functions $N_{i,k}(x)$
$\mathcal{S}(x)$	spline
$\mathbb{S}^d_{k,\Theta}$	d-dimensional tensor-product spline space of order k
$\mathcal{S}(x) := S(x_1, \ldots, x_d)$	d-dimensional tensor-product spline
h	uniform grid size

General Notations for elliptic and parabolic variational inequalities:

$\lesssim$, $\gtrsim$	less/greater or equal to except for a positive constant independent of all parameters, i.e., $u \lesssim v \Leftrightarrow$ there exist a constant $c > 0$ such that $u \leq cv$ and analogous for $\gtrsim$
$\frac{\partial u}{\partial x}$ or $\partial_x u$	partial derivative of u with respect to x
$\Omega \subset \mathbb{R}^d, d = 1, 2$	spatial domain
$\partial\Omega$	boundary of the spatial domain
$C^{\mathfrak{m}}(\Omega)$.....	space of $\mathfrak{m}$'th continuously differentiable function on Ω
$\|\cdot\|_X$	norm on X
$(\cdot,\cdot)_X$	inner product on X
$H := L^2(\Omega)$	space of square integrable functions on Ω
$(\cdot,\cdot)$	inner product on $L^2(\Omega)$
$W^{\mathfrak{m},p}(\Omega)$	Sobolev space of order $\mathfrak{m} \in \mathbb{N}$ on Ω in the Lebesgue space $L^p(\Omega)$ with $p \in \mathbb{N}$
$H^{\mathfrak{m}}(\Omega)$	Sobolev space of order $\mathfrak{m} \in \mathbb{N}$ on Ω in the Lebesgue space $L^2(\Omega)$
$\mathcal{V}$	Sobolev space of order one on Ω
$H_0^{\mathfrak{m}}(\Omega)$	Sobolev space of order $\mathfrak{m} \in \mathbb{N}$ on domain Ω with zero boundary condition on $\partial\Omega$
$\mathcal{V}^0 \subset \mathcal{V}$	Sobolev space of order one on domain Ω with zero boundary condition on $\Upsilon \subseteq \partial\Omega$

$\mathcal{V}^*$	Dual space of $\mathcal{V}^0$
$\langle\cdot,\cdot\rangle_{\mathcal{V}^*,\mathcal{V}}$	Duality pairing between $\mathcal{V}^*$ and $\mathcal{V}$
$W^{\mathfrak{s},p}(\Omega)$	Sobolev-Slobodeckij space with non-integer $\mathfrak{s} > 0$ on Ω in the Lebesgue space $L^p(\Omega)$
$H^{\mathfrak{s}}(\Omega)$	Sobolev-Slobodeckij space with non-integer $\mathfrak{s} > 0$ on Ω in the Lebesgue space $L^2(\Omega)$
$W(0,T,X)$	Bochner space of real vector-valued functions with $y(t) : (0,T) \to X$
$\mathcal{A} : \mathcal{V} \to \mathcal{V}^*$	spatial linear operator
$a(\cdot,\cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$.	spatial bilinear form on $\mathcal{V} \times \mathcal{V}$
$y(x)$ or $y(\tau,x)$	solution $y(x) \in \mathcal{V}$ of an elliptic variational inequality or solution $y(\tau,x) : (0,T) \to \mathcal{V}$ of a parabolic variational inequality with inhomogeneous Dirichlet boundary condition on $\Upsilon \subseteq \partial\Omega$
$\psi := \psi(x) : \Omega \to \mathbb{R}$.	Obstacle function, and also the initial condition for parabolic problems, depending only on spatial parameter $x \in \Omega$
$\mathcal{K} \subset \mathcal{V}$	closed convex set of functions in $\mathcal{V}$ with inhomogeneous Dirichlet boundary condition on $\Upsilon \subseteq \partial\Omega$ (see (3.8))
$\mathcal{K}^0 \subset \mathcal{V}^0$	closed convex set of functions in $\mathcal{V}^0$ with homogeneous Dirichlet boundary condition on $\Upsilon \subseteq \partial\Omega$ (see (3.10))
$u(x)$ or $u(\tau,x)$	solution $u(x) \in \mathcal{K}^0 \subset \mathcal{V}^0$ of an elliptic variational inequality or solution $u(\tau,x) : (0,T) \to \mathcal{V}^0$ with $u(\tau,x) \in \mathcal{K}^0$ for a.e. $\tau \in (0,T)$ of a parabolic variational inequality with homogeneous Dirichlet boundary condition on $\Upsilon \subseteq \partial\Omega$
$\tilde{\psi}(x) : \Omega \to \mathbb{R}$	obstacle function, and also the initial condition for parabolic problems after transformation to homogeneous Dirichlet boundary condition on $\Upsilon \subseteq \partial\Omega$
$\varepsilon > 0$	penalty parameter
$u^\varepsilon(\tau)$	solution $u^\varepsilon(\tau) \in \mathcal{V}^0$ of the penalty problem as stated in (3.27)
$\mathfrak{Q}u^\varepsilon(\tau)$	penalty function $\mathfrak{Q}u^\varepsilon(\tau) := \min(u^\varepsilon(\tau) - \widetilde{\psi}, 0)$
$\hat{u}(\tau)$	integral $\hat{u}(\tau) := \int_0^\tau u(\mathfrak{r})\mathrm{d}\mathfrak{r}$
$\delta_s(u(\tau))$	finite difference quotient for $\partial_\tau u(\tau)$ with time step size $s > 0$

General notations for options:

T	expiration date
t	time

$\tau := T - t$	time until expiry
S	underlying price
K, K_1, K_2	strike prices
$x = \log(S/K)$	log-transformed underlying price
$\mathcal{H}_{\mathrm{C}}(S)$	payoff for a call option
$\mathcal{H}_{\mathrm{P}}(S)$	payoff for a put option
$\mathcal{H}_{\mathrm{BS}}(S)$	payoff for a butterfly-spread option
$g(x)$	log-transformed payoff for a put option
$W(t)$	Wiener process
r	risk free interest rate

Notations for option pricing with the Black-Scholes model

Continuous Problem:

$\sigma > 0$	constant volatility
μ	constant drift
$D_0 > 0$	constant dividend yields
$V(t, S)$	American option price
$\mathcal{L}^B V(t, S)$	spatial operator of the Black-Scholes equation in strong form as defined in (2.11)
$I := (0, S_{\max})$	spatial interval
$S_f(t) : [0, T] \to \mathbb{R}^+$	early exercise price or free boundary
$\mathcal{V}$	weighted Sobolev space of once weak differentiable function on $I \subset \mathbb{R}$ as defined in (2.70)
$\mathcal{K} \subset \mathcal{V}$	convex set in $\mathcal{V}$
$\mathcal{V}^0 \subset \mathcal{V}$	weighted Sobolev space on $I \subset \mathbb{R}$ with homogeneous Dirichlet condition on ∂I as defined in (2.72)
$\mathcal{K}^0 \subset \mathcal{V}^0$	convex set in $\mathcal{V}^0$
$a^B(\cdot, \cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$	bilinear form/weak form of the spatial Black-Scholes operator as defined in (2.78)
$y(\tau) := V(T - t, S)$	time reversed American option price ($\widehat{=}$ solution of the variational inequality in (2.77))
$y^\varepsilon(\tau)$	solution of the penalty problem in (4.16)
$\mathfrak{Q} y^\varepsilon(\tau)$	penalty function
$u(\tau) := y(\tau) - \mathcal{H}$	solution of the variational inequality with homogeneous boundary condition in (2.82)
$\tilde{u}_{\mathcal{H}}$	linear extension of $\mathcal{H}(0)$ and $\mathcal{H}(S_{\max})$ as defined in (4.17)

$\tilde{u}^{\varepsilon}(\tau) := y^{\varepsilon}(\tau) - \tilde{u}_{\mathcal{H}}$	solution of the penalty problem in (4.18) with homogeneous boundary condition on ∂I

B-spline Galerkin semi-discretization:

$0 < s < 1$	time step size
$\tau^{(z)} \in [0,T]$	mesh point
$\#\mathbb{T}$	number of mesh points on $[0,T]$
$u^{(z)} := u(\tau^{(z)}, S)$	semi-discrete solution in time step z with homogeneous boundary condition on ∂I
$\varpi \in \{0, 0.5, 1\}$	weighted average: $\varpi = 0$ explicit Euler method, $\varpi = 0.5$ Crank-Nicolson method and $\varpi = 1$ implicit Euler method
$\tilde{a}^{B}(\cdot,\cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$	bilinear form after semi-discretization as defined in (4.30)
$\tilde{f}^{(z-1)}$	right hand side after semi-discretization as defined in (4.31)
$\Theta := \{\theta_i\}_{i=1}^{n+k}$	extended sequence of knots for the B-spline basis functions
$\theta_{\wp} := K$	$k-1$ times repeated knot at strike price K
$0 < h < 1$	uniform grid size
$n \in \mathbb{N}$	number of the B-spline basis function
$\mathcal{V}_h^0 \subset \mathcal{V}^0$	finite dimensional spline space with homogeneous boundary condition on ∂I
$\mathcal{I}$	Index set for the B-spline basis function in $\mathcal{V}_h^0$
$\mathcal{K}_h^0 \subset \mathcal{K}^0$	discrete convex set
$u_h^{(z)} \in \mathcal{K}_h^0$	fully discrete spline solution in time step z
$\mathbf{u}^{(z)} \in \mathbb{R}^{\#\mathcal{I}}$	coefficients vector for the spline solution in time step z
$\mathfrak{A}, \mathfrak{B}, \mathfrak{G} \in \mathbb{R}^{\#\mathcal{I}\times\#\mathcal{I}}$	B-spline discretization matrices
$C \in \mathbb{R}^{\#\mathcal{I}\times\#\mathcal{I}}$	final B-spline discretization matrix for the Black-Scholes operator
$\mathbf{f}^{(z-1)} \in \mathbb{R}^{\#\mathcal{I}}$	fully discrete right hand side in vector notation
$\mathbf{K}$	fully discrete convex set, which ensures that the coefficients of the spline solution are non negative.
$\mathfrak{S}_{S_f^h}^{(z)}$	union of the supports of B-spline basis functions whose support intersects $S_f^h(t^{(z)})$
$\Delta_h(t^{(z)}, S)$	approximation of the Greek Delta at time $t^{(z)}$
$\Gamma_h(t^{(z)}, S)$	approximation of the Greek Gamma at time $t^{(z)}$

Notations for option pricing with Heston's model:

Continuous Problem:

$v \in \mathbb{R}_+$	variance

$W_1(t), W_2(t)$	two different Wiener processes correlated by constant correlation $\|\rho\| < 1$
μ	constant drift
$\|\rho\| < 1$	constant correlation
$\kappa > 0$	mean reversion rate
$\gamma > 0$	mean reversion level
$\xi > 0$	volatility of the CIR-process
$\lambda \geq 0$	market price of volatility risk
$\Omega_{\mathcal{L}} \subset \mathbb{R}^+ \times \mathbb{R}^+$	spatial domain $\Omega := (S_{\min}, S_{\max}) \times (v_{\min}, v_{\max})$
$V(t, S, v)$	American option price
$S_f(t, v) : [0, T] \times \mathbb{R}_+ \to \mathbb{R}_+$	optimal exercise price or free boundary
$\mathcal{L}^H V(t, S, v)$	spatial operator of the Heston equation in strong form as defined in (2.23)
$\Omega \subset \mathbb{R} \times \mathbb{R}^+$	spatial domain $\Omega := (x_{\min}, x_{\max}) \times (v_{\min}, v_{\max})$ with log-transformed underlying $x := \log(S/K)$
$\partial\Omega := \cup_{i=1}^4 \Upsilon_i$	boundary of the spatial domain
$\Upsilon_v \subset \partial\Omega$	parts of the boundary, where the Neumann-boundary conditions hold, see (2.110) or (2.111)
$\Upsilon := \partial\Omega \setminus \Upsilon_v$	parts of the boundary, where the Dirichlet boundary conditions hold
$y(\tau) := y(\tau, x, v)$	time reversed and log-transformed American option price as defined in (2.101) ($\widehat{=}$ solution of the variational inequality)
$u(\tau) := y(\tau) - g$	solution of the variational inequality with homogeneous boundary condition on Υ
$A \in \mathbb{R}^{2\times 2}$	coefficients matrix for the diffusion term as defined in (2.103)
$\mathbf{b} \in \mathbb{R}^2$	coefficients vector for the convection term as defined in (2.103)
$\mathcal{Z}^H y(\tau)$	time reversed and log-transformed spatial operator in strong form
$x_f(\tau, v) : [0, T] \times \mathbb{R}^+ \to \mathbb{R}$.	log-transformed and time reversed free boundary
$\mathcal{V}$	Sobolev space of once weak differentiable function as defined in (2.113)
$\mathcal{V}^0 \subset \mathcal{V}$	Sobolev space of once weak differentiable function with homogeneous Dirichlet boundary condition on Υ, see (2.115)
$\mathcal{K} \subset \mathcal{V}$	closed convex set in $\mathcal{V}$ as defined in (2.117)
$\mathcal{K}^0 \subset \mathcal{V}^0$	closed convex set in $\mathcal{V}^0$ as defined in (2.130)

$\tilde{A} \in \mathbb{R}^{2\times 2}$	coefficients matrix for the diffusion term after integration by parts, see (2.124)
$\tilde{\mathbf{b}} \in \mathbb{R}^2$	coefficients vector for the convection term after integration by parts, see (2.126)
$a^H(\cdot,\cdot) : \mathcal{V} \times \mathcal{V}$	bilinear form for the Heston operator $\mathcal{Z}^H$, see (2.128)
f	right hand side after homogenization, see (2.131)
$y^\varepsilon(\tau) := y^\varepsilon(\tau, x, v)$	solution of the penalty problem in (5.17)
$\mathfrak{Q}y^\varepsilon(\tau)$	penalty function

B-spline Galerkin semi-discretization:

$y^{(z)} := y(\tau^{(z)}, x, v)$	semi-discrete solution in time step z
$\tilde{a}^H(\cdot,\cdot) : \mathcal{V} \times \mathcal{V}$	bilinear form after semi-discretization, see (5.22)
$\tilde{\mathcal{A}}^H : \mathcal{V} \to \mathcal{V}^*$	linear operator associated with the bilinear form $\tilde{a}^H(w, \varphi)$
$f_{\text{inh}}^{(z-1)}$	right hand side after semi-discretization for the problem with inhomogeneous Dirichlet boundary condition, see (5.23)
$u^{(z)} := u(\tau^{(z)}, x, v)$	semi-discrete solution in time step z with homogeneous boundary condition on Υ
$f^{(z-1)}$	right hand side after semi-discretization for the problem with homogeneous Dirichlet boundary condition, see (5.33)
$\Theta_x := \{\theta_i^{(x)}\}_{i=1}^{n+k}$	extended sequence of knots in the $x-$direction
$\Theta_v := \{\theta_i^{(v)}\}_{i=1}^{n+k}$	extended sequence of knots in the $v-$direction
$\Theta := \Theta_x \otimes \Theta_v$	two dimensional extended sequence of knots
h_{ℓ_x}	uniform grid size for the knot series Θ_x
h_{ℓ_v}	uniform grid size for the knot series Θ_v
$\mathcal{V}_h^0 \subset \mathcal{V}^0$	finite dimensional tensor-product spline space
$\mathcal{I}_D$	index set for the tensor-product B-spline basis function for the set of boundary conditions in (2.110)
$\mathcal{I}_N$	index set for the tensor-product B-spline basis function for the set of boundary conditions in (2.111)
$\mathcal{I}$	general index set, which is given by $\mathcal{I}_D$ or $\mathcal{I}_N$
$\#\mathcal{I} := NM$	number of the tensor-product B-spline basis function in $\mathcal{V}_h^0$
$\mathcal{K}_h^0 \subset \mathcal{K}^0$	discrete convex set
$u_h^{(z)} \in \mathcal{K}_h^0$	fully discrete tensor-product spline solution in time step z
$\mathbf{u}^{(z)} \in \mathbb{R}^{\#\mathcal{I}}$	coefficients vector for the spline solution in timestep z
$C \in \mathbb{R}^{NM\times NM}$	tensor-product B-spline discretization matrix for the Heston operator, see (5.43)
$\mathbf{f}^{(z-1)} \in \mathbb{R}^{NM}$	fully discrete right side in vector notation

$\Delta_h(t^{(z)}, S, v)$	approximation of the Greek Delta at time $t^{(z)}$
$\Gamma_h(t^{(z)}, S)$	approximation of the Greek Gamma at time $t^{(z)}$
$\nu_h(t^{(z)}, S, v)$	approximation of the Greek Vega at time $t^{(z)}$

Notations for monotone multigrid method:

ℓ	level
$L := \ell_{\max}$	highest level
$\mathcal{V}_\ell$	(tensor-product) spline space on level ℓ
$\mathcal{K}_\ell^0$	convex set on level ℓ
$\mathcal{A}_\ell$	discrete form of the linear operator $\mathcal{A}$ on level ℓ
f_ℓ	discrete from of the right side on level ℓ
u_ℓ	discrete solution on level ℓ
$\tilde{\psi}_\ell$	discrete obstacle function on level ℓ
$\dim(\mathcal{V}_\ell) := \#\mathcal{I}_\ell$	number of B-spline basis function on level ℓ
$C_\ell \in \mathbb{R}^{\#\mathcal{I}_\ell \times \#\mathcal{I}_\ell}$	discretization matrix on level ℓ
$\mathbf{f}_\ell \in \mathbb{R}^{\#\mathcal{I}_\ell}$	right side in vector notation on level ℓ
$\mathbf{u}_\ell$	B-spline coefficient vector for the solution u_ℓ
$\tilde{\boldsymbol{\psi}}_\ell \in \mathbb{R}^{\#\mathcal{I}_\ell}$	B-spline coefficients vector for the obstacle function $\tilde{\psi}_\ell$
$\mathcal{R}, \tilde{r} : \mathcal{V}_\ell \to \mathcal{V}_{\ell-1}$	restriction operator
$\mathcal{P} : \mathcal{V}_{\ell-1} \to \mathcal{V}_\ell$	prolongation operator
$\mathbb{S}_{k,\Theta}^d$	d-dimensional tensor-product spline space of order k regarding the knot sequence Θ
$\Theta = \{\otimes_{j=1}^d \Theta^{(j)}\}$	d-dimensional open knot sequences with uniform grid size h_j
$\tilde{\Theta} = \{\otimes_{j=1}^d \tilde{\Theta}^{(j)}\}$	d-dimensional open knot sequences with uniform grid size H_j
h_j	uniform grid size on a finer grid $\Theta^{(j)}$
$H_j := 2h_j$	uniform grid size on a coarser grid $\tilde{\Theta}^{(j)}$
$n_j \in \mathbb{N}$	number of B-splines on a finer grid
$\tilde{n}_j \in \mathbb{N}$	number of B-splines on a coarser grid
J_{rep}	index set for the dimensions, where internal knots are repeated
$\overline{J_{\text{rep}}}$	index set for the dimensions, where no internal knots are repeated
$N_{i_j,k}(x_j)$	B-spline regarding the knot sequence $\Theta^{(j)}$
$\tilde{N}_{i_j,k}(x_j)$	B-spline regarding the knot sequence $\tilde{\Theta}^{(j)}$
a_q	refinement coefficients as defined in (7.24)
Φ^L, Φ^R	matrices containing the boundary adapted refinement coefficients, see (7.25) and (7.26)

Φ^I	matrices containing the refinement coefficients regarding the B-splines where internal knots are repeated $k-1$ times, see (7.31)
$\boldsymbol{\mathcal{R}}_k^{(j)} \in \mathbb{R}^{\tilde{n}_j \times n_j}$	restriction operator in the j'th dimension in matrix notation for B-splines of order k
$\boldsymbol{\mathcal{P}}_k^{(j)} \in \mathbb{R}^{n_j \times \tilde{n}_j}$	prolongation operator in the j'th dimension in matrix notation for B-splines of order k
$L_k = \tilde{\mathbf{r}}^T \widetilde{\mathbf{N}}_k \in \mathbb{S}_{k,\tilde{\Theta}}^d$	quasioptimal monotone coarse grid approximation to the lower obstacle $S := \mathbf{c}^T \mathbf{N}_k \in \mathbb{S}_{k,\Theta}^d$, see (7.45)

Notations for the Black-Scholes-Barenblatt equation:

$V(t,S)$	option price under the worst or best-case volatility path
$\sigma \in \Sigma$	volatility in the range $\Sigma := [\sigma_{\min}, \sigma_{\max}]$
$\mathcal{L}^\sigma$	spatial differential operator for the BSB-equation
$\underline{\sigma}^-[V_{SS}]$	volatility function under the worst-case volatility path
$\overline{\sigma}^+[V_{SS}]$	volatility function under the best-case volatility path
$y(\tau,S) := V(T-t,S)$	time reversed option price
$I := (0, S_{\max})$	spatial interval
$I := (S_{\min}, S_{\max})$	truncated spatial interval
$I_x := (x_{\min}, x_{\max})$ ...	truncated spatial interval for the log-transformed option price
$\mathcal{W}$	Sobolev space $\mathcal{W} := H^2(I_x) \cap H_0^1(I_x)$
$H(0,T;I_x)$	Bochner space $H(0,T;I_x) := L^2(0,T;\mathcal{W}) \cap H^1(I_x; L^2(I_x))$
$F[\cdot]$	fully non-linear operator $F : H(0,T;I_x) \to L^2(0,T;I_x)$
$\lambda > 0, \omega > 0$	parameter for the Cordes condition
$\gamma(\cdot)$	strictly positive function $\gamma(\cdot) : \Sigma \to \mathbb{R}_{>0}$
L_λ, L_ω	operators needed to establish well-posedness as defined in (9.14)
$\langle \cdot, \cdot \rangle_{L^2 \times \mathcal{W}}$	duality pairing as defined in (9.15)
$\lvert \cdot \rvert_{H^2(I_x),\lambda}$	norm in $H^2(I_x)$ as defined in (9.17)
A	operator for the BSB equation formulated as a mapping from $H(0,T,I_x) \to H(0,T,I_x)^*$, see (9.22)
$y(\tau^{(z)}, S)$	semi-discrete solution in time step z
$\mathcal{V}_\Theta$	spatial discrete solution space with B-splines regarding the knot series Θ
n_Θ	number of B-spline basis function
$\mathcal{A}^{\underline{\sigma}}$	discrete operator for the BSB-equation that maps from $\mathcal{V}^0$ to $\mathbb{R}^{n_\Theta}$, see (9.34)

$u_\Theta^{(z)}$	fully discrete solution with homogeneous Dirichlet-boundary condition
$\underline{\hat{\boldsymbol{\sigma}}}_\Theta(u_\Theta^{(\tilde{z})})$	fully discrete volatility function in weak form as defined in (9.39)
$\underline{\hat{\boldsymbol{\sigma}}}_\Theta^{\text{opt}}\left(u_\Theta^{(\tilde{z})}\right)$	fully discrete volatility function in strong form as defined in (9.41)
$F^1(\cdot), F^2(\cdot)$	fully discrete non-linear operator for the BDF1 or BDF2 scheme
$C_\Theta^1, C_\Theta^2 \in \mathbb{R}^{n_\Theta \times n_\Theta}$ ...	discretization matrices for the BDF1 or BDF2 scheme
$\mathbf{f}_\Theta^1, \mathbf{f}_\Theta^2 \in \mathbb{R}^{n_\Theta}$	discrete right side for the BDF1 or BDF2 scheme
$G_\Theta \in \mathbb{R}^{n_\Theta \times n_\Theta}$	slant derivative of the fully discrete non-linear operator $F(\cdot)$
I_η^+, I_η^-	intervals for volatility changes depending on the discrete weak derivatives, see (9.54) and (9.55)
$K_{\tilde{\Theta}} \in \mathbb{R}^{n_{\tilde{\Theta}} \times n_\Theta}$	knot insertion matrix as defined in (A.6)
$I_\eta^{\text{opt},+}, I_\eta^{\text{opt},-}$	intervals for volatility changes depending on the approximated second derivatives, see (9.59) and (9.60)

Bibliography

[ABM09] H. Albrecher, A. Binder, and P. Mayer. *Einführung in die Finanzmathematik, Mathematik Kompakt.* Birkhäuser, 2009.

[Ach05] Y. Achdou. An inverse problem for a parabolic variational inequality in volatility calibration with American Options. *SIAM J. Control Optim.*, 43(5):1583–1615, 2005.

[AF03] R. Adams and J. Fournier. *Sobolev Spaces.* Elsevier Science, 2nd edition, 2003.

[ALP95] M. Avellaneda, A. Levy, and A. Paras. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2(2):73–88, 1995.

[AMST07] H. Albrecher, P. Mayer, W. Schoutens, and J. Tistaert. The little Heston trap. *Wilmott Magazine*, pages 83–92, 2007.

[AP05] Y. Achdou and O. Pironneau. *Computational Methods for Option pricing.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 4th edition, 2005.

[AWW01] T. Apel, G. Winkler, and U. Wystup. Valuation of options in Heston's stochastic volatility model using finite element methods. *Foreign Exchange Risk, Risk Publications*, London,2001.

[Bai89] C. Baiocchi. Discretization of evolution variational inequalities. *In Partial differential equations and the calculus of variations*, Vol. 1, Birkhäuser:59–92, 1989.

[BC83] A. Brandt and C.W. Cryer. Multigrid algorithms for the solution of linear complementarity problems arising from free boundary problems. *SIAM J. Sci. Stat. Comput.*, 4(4):655–684, 1983.

[Bel57] R. Bellman. Dynamic programming. *Princeton University Press, Princeton, NJ*, 1957.

[BHR77] F. Brezzi, W. W. Hager, and P.A. Raviart. Error estimates for the finite element solution of variational inequalities. *Numer. Math.*, 28:431–443, 1977.

[BHSW15] O. Burkovska, B. Haasdonk, J. Salomon, and B. Wohlmuth. Reduced basis methods for pricing options with the Black-Scholes and Heston models. *SIAM Journal Financial Math.*, 6:685–712, 2015.

[BL82] A. Bensoussan and J.L. Lions. *Applications of Variational Inequalities in Stochastic Control*. North-Holland, Amsterdam-New York-Oxford, 1982.

[BMZ09] O. Bokanowski, S. Maroso, and H. Zidani. Some convergence results for Howard's algorithm. *SIAM J. Numer. Anal.*, 47:3001 – 3026, 2009.

[Boe80] W. Boehm. Inserting new knots into B-spline curves. *Computer-Aided Design*, 12(4):199 – 201, 1980.

[Bos15] S. Boschert. *Bewertung Amerikanischer Optionen mit stochastischer Volatilität: Multigridverfahren basiered auf Tensorprodukt-B-splines höherer Ordnung*. Masterarbeit, Universität zu Köln, Mathematisches Institut, 2015.

[Bré71] H. Brézis. Nouveaux théorèms de régularité pour les problèmes unilatèraux. *Rencontre entre physiciens théoriciens et mathématiciens*, Strasbourg 12, 1971.

[BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81:637–654, 1973.

[BS91] G. Barles and P. Souganidis. Convergence of approximation schemes for fully nonlinear second-order equations. *Asymptotic Anal.*, 4:271–283, 1991.

[Bud13] A. Budke. *Finite Difference Methods for the Non-linear Black-Scholes-Barenblatt Equation*. PhD thesis, Universität zu Köln, Mathematisches Institut, 2013.

[Bur16] O. Burkovska. *Reduced Basis Methods for Option Pricing and Calibration*. PhD Thesis, Technische Universität München, Fakultät für Mathematik, 2016.

[CH19] Durkbin Cho and Rafael Vázquez Hernández. Bpx preconditioners for isogeometric analysis using analysis-suitable T-splines. *ArXiv*, abs/1907.01790, 2019.

[Che11] W.-T. Chen. *An extensive exploration of three key quantitative approaches for pricing various financial derivatives*. PhD Thesis, University of Wollongong, School of Mathematics and Applied Statistics, 2011.

[CP99] N. Clarke and K. Parott. Multigrid for american option pricing with stochastic volatility. *Applied Mathematical Finance*, 6(3):177–195, 1999.

[Cry71] C. W. Cryer. The solution of a quadratic programming problem using systematic overrelaxation. *SIAM J. Control*, 9:385–392, 1971.

[dB01] C. de Boor. *A practical Guide to Splines*. Springer, 2001.

[DB06] P. Doktor and A. Ž. Brno. The density of infinitely differentiable functions in Sobolev spaces with mixed boundary conditions. *Appl. Math.*, 51:517–547, 2006.

[DL] R. Dautov and A.V. Lapin. Error estimates for lagrange-galerkin approximation of american options valuation. *SIAM J. Numer. Anal.*

[DL19a] R. Dautov and A.V. Lapin. Sharp error estimate for implicit finite element element scheme for American put option. *Russ. J. Numer. Anal. Math. Modelling*, 34:1–11, 2019.

[DL19b] R. Dautov and A.V. Lapin. Three new weak formulations of the problem of american call options valuation. *J. Phys. Conf. Ser.*, 1158:1–8, 2019.

[DR08] W. Dahmen and A. Reusken. *Numerik für Ingenieure und Naturwissenschaftler*. Springer, 2th edition, 2008.

[Eva98] L. C. Evans. *Partial Differential equations*. American Mathematical Society, 1998.

[Fal74] R. S. Falk. Error estimates for the approximation of a class of variational inequalities. *Mathematics of Computation.*, 28(128):963–971, 1974.

[FLMN11] L. Feng, V. Linetsky, J. L. Morales, and J. Nocedal. On the solution of complementarity problems arising in american options pricing. *Optim. Methods Softw.*, 26.4-5:813–825, 2011.

[FP03] P.A. Forsyth and D. M. Pooley. Numerical convergence properties of option pricing pdes with uncertain volatility. *IMA Journal of Numerical Analysis*, 23:241 – 267, 2003.

[FV02] P.A. Forsyth and K.R. Vetzal. Quadratic convergence for valuing american options using a penalty method. *SIAM Journal on Scientific Computing*, 23(6), 2002.

[Gat06] J. Gatheral. *The volatility surface: A practitioner's Guide*. John Wiley & Sons, 2006.

[GC06] M. B. Giles and R. Carter. Convergence analysis of crank-nicolson and rannacher time-marching. *Journal of Computational Finance*, 9(4):89–112, 2006.

[GK09] C. Gräser and R. Kornhuber. Multigrid methods for obstacle problems. *Journal of Computational Mathematics*, 27(1):1–44, 2009.

[GKT13] K.P.S. Gahalaut, J.K. Kraus, and S.K. Tomar. Multigrid methods for isogeometric discretization. *Comput. Methods Appl. Mech. Engrg.*, 253:412 – 425, 2013.

[Glo71] R. Glowinski. La méthode de rélaxation. *Rend. Mat.*, 14:1–56, 1971.

[Glo84] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Springer-Verlag, 1th edition, 1984.

[Hei10] P. Heider. Numerical methods for non-linear Black-Scholes equations. *Applied Mathematical Finance*, 17(1):59 – 81, 2010.

[Hes93] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Finan. Stud.*, 6(2):327–343, 1993.

[HK07] M. Holtz and A. Kunoth. B-spline-based monotone multigrid methods. *Society for Industrial and Applied Mathematics, SIAM J. Numer. Anal.*, 45(3):1175–1199, 2007.

[HM83] W. Hackbusch and H.-D. Mittelmann. On multigrid methods for variational inequalities. *Numer. Math.*, 42:65–76, 1983.

[HMS19] C. Hofreither, L. Mitter, and H. Speleers. Local multigrid solvers for adaptive isogeometric analysis in hierarchical spline spaces. *RICAM-Report 2019-34,* `https://www.ricam.oeaw.ac.at/files/reports/19/rep19-34.pdf`, 2019.

[Hol04] M. Holtz. *Konstruktion B-Spline-basierter monotoner Mehrgitterverfahren zur Bewertung Amerikanischer Optionen.* Diplomarbeit, Rheinische Friedrich-Wilhelms-Universität Bonn, Mathematisch-Naturwissenschaftliche Fakultät, 2004.

[Hop87] R. Hoppe. Multigrid algorithms for variational inequalities. *SIAM J. Numer. Anal.*, 24:1046–1065, 1987.

[How60] R.A. Howard. Dynamic Programming and Markov Processes. *The MIT Press, Cambridge, MA*, 1960.

[HRSW13] H. Hilber, O. Reichmann, C. Schwab, and C. Winter. *Computational Methods for quantitative Finance.* Springer, 2013.

[HT17] C. Hofreither and S. Takacs. Robust multigrid for isogeometric analysis based on stable splittings of spline spaces. *SIAM J. on Numer. Anal.*, 55(4):2004–2024, 2017.

[Hul09] J.C. Hull. *Options, Futures and Other Derivatives.* Pearson Education, 7th edition, 2009.

[IK06] K. Ito and K. Kunisch. Parabolic variational inequalities: The lagrange multiplier approach. *Math. Pures Appl.*, 85:415–449, 2006.

[IT09] S. Ikonen and J. Toivanen. Operator splitting methods for american options with stochastic volatility. *Numer. Math.*, 113(2):299–324, 2009.

[JL13] M. Jung and U. Langer. *Methode der finiten Elemente für Ingenieure.* Springer Vieweg, 2nd edition, 2013.

[JS13] M. Jensen and I. Smears. On the convergence of finite element methods for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.*, 51(1):137 – 162, 2013.

[KLM+17] A. Kunoth, T. Lyche, C. Manni, G. Sangalli, S. Serra-Capizzano, and Hendrik Speleers. *Splines and PDEs: From Approximation Theory to Numerical Linear Algebra.* Springer, 2017.

[Kor94] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities I. *Numer. Math.*, 69:167–184, 1994.

[Kor96] R. Kornhuber. A posteriori error estimates for elliptic variational inequalities. *Comput. Math. Appl.*, 31:49–60, 1996.

[KS80] D. Kinderlehrer and G. Stampacchia. An Introduction to Variational Inequalities and their Applications. *Pure and Applied Mathematics, Academic Press Inc.*, 88, New York,1980.

[KS91] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer Verlag, New York, vol. 113 of grad. texts in mathematics edition, 1991.

[KS00] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and their Applications.* SIAM, 2000.

[KSW12] A. Kunoth, C. Schneider, and K. Wiechers. Multiscale methods for the valuation of american options with stochastic Volatility. *International Journal of Computer Mathematics*, 89(2):1145–1163, 2012.

[KTK17] M.K. Kadalbajoo, L.K. Tripathi, and A. Kumar. An error analysis of a finite element method with IMEX-time semi-discretzations for some partial integro-differential inequalities arising in the pricing of American options. *SIAM J. Numerical Analysis*, 55(3):869–891, 2017.

[KY94] R. Kornhuber and H. Yserentant. Multilevel methods for elliptic problems on domains not resolved by the coarse grid. *Contemporary Mathematics*, 180:49–60, 1994.

[LM08] T. Lyche and K. Mørken. *Spline methods draft.* Department of Informatics, Centre of Mathematics for Applications, University of Oslo, 2008.

[LS67] J. L. Lions and G. Stampacchia. Variational inequalities. *Communications on Pure and Applied Mathematics*, 20(3):493–519, 1967.

[Lyo95] T. J. Lyons. Uncertain volatility and the risk-free synthesis of derivatives. *Applied Mathematical Finance*, 2(2):117–133, 1995.

[Man84] J. Mandel. A multilevel iterative method for symmetric, positive definite linear complementarity problems. *Appl. Math. Optim.*, 11:77–95, 1984.

[Mau06] K. Mautner. *Numerical treatment of the Black-Scholes Variational Inequality in Computational Finance.* PhD thesis, Mathematisch-Naturwissenschaftliche Fakultät, Humboldt-Universität zu Berlin, 2006.

[Mem12] S. Memon. Finite element method for american option pricing: A penalty approach. *IJNAM*, 3(3):345–370, 2012.

[Mer73] R. Merton. The theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4:141–183, 1973.

[Mif77] R. Mifflin. Semismooth and semiconvex functions in constrained optimization. *SIAM J. Control Optim.*, 15(6):959–972, 1977.

[Mol16] C. Mollet. *Parabolic PDEs in space-time formulations: Stability for Petrov-Galerkin Discretizations with B-Splines and Existence of Moments for Problems with Random Coeffcients.* PhD thesis, Universität zu Köln, Mathematisches Institut, 2016.

[Oos03] C.W. Oosterlee. On multigrid for linear complementary problems with application to American-style options. *Electronics Transactions on Numerical Analysis*, 15:165–185, 2003.

[Pan84] J. S. Pang. Necessary and sufficient conditions for the convergence of iterative methods for the linear complementarity problem. *Journal of optimization theory and applications*, 42(1):1–17, 1984.

[PT97] L. Piegl and W. Tiller. *The Nurbs Book.* Springer, 2nd edition, 1997.

[Qi93] L. Qi. Convergence analysis of some algorithms for solving nonsmooth equations. *Oper. Res.*, 18:227–244, 1993.

[QS93] L. Qi and J. Sun. A nonsmooth version of newton's method. *Math. Program.*, 58:353–367, 1993.

[Ran84] R. Rannacher. Finite element solution of diffusion problems with irregular data. *Numerische Mathematik*, 43: pp. 309–327, 1984.

[Rei04] C. Reisinger. *Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben.* PhD Thesis, Universität Heidelberg, Mathematisches Institut, 2004.

[Rou13] F. D. Rouah. *The Heston model and its extension in Matlab and C#.* John Wiley & Sons, 2013.

[SCdFV] M. Spink, D. Claxton, C. de Falco, and R. Vázquez. The NURBS toolbox. `http://octave.sourceforge.net/nurbs/index.html`, July 2020.

[Sch07] L. L. Schumaker. *Spline functions: Basic theory.* Cambridge University Press, 3rd edition, 2007.

[Sen15] E. Sen. *Bewertung europäischer Optionen für die Black-Scholes-Barenblatt-Gleichung in Variationsformulierung mittels B-Splines höherer Ordnung.* Masterarbeit, Universität zu Köln, Mathematisches Institut, 2015.

[Sey09] R. Seydel. *Tools for Computational Finance.* Springer, 4th edition, 2009.

[Sey12] R. Seydel. *Tools for Computational Finance.* Springer, 5th edition, 2012.

[Sme12] I. Smears. Hamilton-Jacobi-Bellman Equations; Analysis and numerical analysis. *Report*, `https://fourier.dur.ac.uk/Ug/projects/highlights/PR4/Smears_HJB_report.pdf`, 2012.

[Sme15] I. Smears. *Discontinuous Galerkin Finite Element approximation of Hamilton-Jacobi-Bellman equations with Cordes coefficients.* PhD thesis, Worcester College, University of Oxford, 2015.

[SS16] I. Smears and E. Süli. Discontinuous Galerkin finite element methods for time-dependent Hamilton-Jacobi-Bellman equations with Cordes coefficients. *Numer. Math.*, 133:141–176, 2016.

[TLG81] R. Trémolières, J.-L. Lions, and R. Glowinski. *Numerical Analysis of Variational Inequalities.* Elsevier North–Holland, 8th edition, 1981.

[Tri78] H. Triebel. *Spaces of Besov-Hardy-Sobolev Type.* Teubner, 1978.

[Var01] T. Vargiolu. Existence, uniqueness and smoothness for the Black-Scholes-Barenblatt equation. *Technical Report 5, Universita di Padova, Department of Pure and Applied Mathematics*, `https://www.math.unipd.it/~vargiolu/BSB.pdf`, 2001.

[vW] G. v. Winckel. Legendre-gauss quadrature weights and nodes. `https://www.mathworks.com/matlabcentral/fileexchange/4540-legendre-gauss-quadrature-weights-and-nodes`, MATLAB Central File Exchange, July 2020.

[WDH93] P. Wilmott, J. Dewynne, and S. Howison. Option Pricing: Mathematical Models and Computation. *Oxford Financial Press*, 1993.

[Wei14] M. Weigler. *Bewertung Amerikanischer Optionen mit stochastischer Volatilität: Finite-Elemente-Ansatz basierend auf Tensorprodukt-B-Splines höherer Ordnung.* Masterarbeit, Universität zu Köln, Mathematisches Institut, 2014.

[Wil01] P. Wilmott. Paul Wilmott introduces quantitative finance. *John Wiley & Sons*, 2001.

[Woh11] B. Wohlmuth. Variationally consistent discretization schemes and numerical algorithms for contact problems. *Acta Numerica*, 20:569–734, 2011.

[Yse86] H. Yserentant. On the multi-level splitting of finite element spaces. *Numer. Math.*, 49:379–412, 1986.

[Yse90] H. Yserentant. Two preconditioners based on the multi-level splitting of finite element spaces. *Numer. Math.*, 58:163–184, 1990.

[Yse93] H. Yserentant. Old and new convergence proofs for multigrid methods. *Acta Numerica*, 2:285–326, 1993.

[ZC11] S.-P. Zhu and W.-T Chen. A predictor-corrector scheme based on adi method for pricing american puts with stochastic volatility. *Comput. Math. Appl.*, 62(2):1–26, 2011.

[ZFV98] R. Zvan, P.A. Forsyth, and K.R. Vetzal. Penalty methods for american options with stochastic volatlity. *Journal of Comp. and Appl. Math.*, 91(2):199–218, 1998.

[Zha07] C.-S. Zhang. *Adaptive Finite Element Methods for Variational Inequalities: Theory and Applications in Finance*. PhD thesis, Faculty of the Graduate School of the University of Maryland, 2007.

[ZW09] K. Zhang and S. Wang. A computational scheme for uncertain volatility model in option pricing. *Elsevier, Applied Numerical Mathematics*, 59:1754 – 1767, 2009.

Eidesstatliche Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, 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 unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist, sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Frau Prof. Dr. Angela Kunoth betreut worden.

Teilpublikationen

[1] S. Boschert and A. Kunoth. B-spline based methods for variational inequalities and a special HJB equation with non-smooth initial data. In *Oberwolfach Report No. 33/2019*, Mathematical Foundations of Isogeometric Analysis, pp. 7-10, 2019.

Köln, den 10.8.2020

Sandra Boschert