Black–Scholes Type Equations: Mathematical Analysis, Parameter Identification & Numerical Solution

Dissertation
zur Erlangung des Grades
Doktor der Naturwissenschaften

am Fachbereich Physik, Mathematik und Informatik
der Johannes Gutenberg–Universität Mainz

Bertram Düring
geboren in Berlin

Mainz, Juli 2005

Abstract

In this work we are concerned with the analysis and numerical solution of Black–Scholes type equations arising in the modeling of incomplete financial markets and an inverse problem of determining the local volatility function in a generalized Black–Scholes model from observed option prices.

In the first chapter a fully nonlinear Black–Scholes equation which models transaction costs arising in option pricing is discretized by a new high order compact scheme. The compact scheme is proved to be unconditionally stable and non–oscillatory and is very efficient compared to classical schemes. Moreover, it is shown that the finite difference solution converges locally uniformly to the unique viscosity solution of the continuous equation.

In the next chapter we turn to the calibration problem of computing local volatility functions from market data in a generalized Black–Scholes setting. We follow an optimal control approach in a Lagrangian framework. We show the existence of a global solution and study first– and second–order optimality conditions. Furthermore, we propose an algorithm that is based on a globalized sequential quadratic programming method and a primal–dual active set strategy, and present numerical results.

In the last chapter we consider a quasilinear parabolic equation with quadratic gradient terms, which arises in the modeling of an optimal portfolio in incomplete markets. The existence of weak solutions is shown by considering a sequence of approximate solutions. The main difficulty of the proof is to infer the *strong* convergence of the sequence in H^1. Furthermore, we prove the uniqueness of weak solutions under a smallness condition on the derivatives of the covariance matrices with respect to the solution, but *without* additional regularity assumptions on the solution. The results are illustrated by a numerical example.

Contents

Chapter 1

Introduction

The first part of this introduction is thought of as a motivation for the following. The reader interested in the scientific results may skip the first section of this introduction and turn directly to the second section which features a short overview on some basic notions and results.

1.1 Mathematics and finance

1.1.1 Financial markets and mathematical research

In mathematical research the study of financial markets has drawn rising attention. The modeling approaches encompass methods and techniques from many different mathematical disciplines: Stochastical and statistical methods, deterministic and stochastic partial differential equations (PDE), methods of applied functional analysis and others. In this work, we take a PDE point of view. Hence, we will try to limit the stochastic calculus to a minimum which is needed to understand the foundations and derivations of the mathematical models considered. Then we investigate analytically and numerically in–depth the resulting partial differential equations.

1.1.2 Mathematical modelling

What is to be understood by mathematical modelling? The process of mathematical modeling usually involves the following steps:

(i) Proper specification of the real problem,

(ii) Conversion into a mathematical formulation,

(iii) Analysis of the mathematical problem,

(iv) Numerical (or analytical, if possible) solution,

(v) Interpretation of the results.

If the results are unsatisfactory, the mathematical model needs to be refined and the cycle is repeated.

Consider a typical example, where we can identify the different steps of the above modelling scheme: In the context of financial mathematics, a common problem is the valuation of particular financial contracts, so–called *financial derivatives* (i). The 'classical' model has been developed by Black and Scholes and independently by Merton. Their mathematical formulation is the Black–Scholes partial differential equation (ii). Its analysis shows that it can be transformed into the heat equation (iii), which admits an analytical solution (iv), resulting in the famous *Black–Scholes formula* which has been used by practitioners for thirty years now (v).

In this work, the focus is laid on the steps (iii) and (iv), the mathematical analysis and numerical solution, which are always complemented by short discussions of the results. Steps (i) and (ii), the specification and mathematical formulation of the problems, are presented in the introduction to give the reader a basis for understanding the sources of the mathematical problems treated later. For details on the derivation and technical points we refer to the original works.

1.2 Basic notions and results

In this section we will introduce some basic notions as well as fundamental assumptions of the modelling, followed by a very short overview on the classical models for two common problems in financial mathematics, option pricing and portfolio optimization.

1.2.1 Markets, derivatives, and options

A *market* is a place where buyers and sellers make transactions, directly or via intermediaries and a *financial market* is a market for the exchange of capital and credit. An *investment* is an item of value purchased for income or capital appreciation and hence an *investor* is someone making investments. A *security* is an investment, other than an insurance policy or fixed annuity, issued by a corporation or other organization which offers evidence of debt or equity.

The *financial market* can be divided into the *money market*, where short–term debt securities, such as banker's acceptances and treasury bills with a maturity of one year or less and often 30 days or less are traded and the *capital market*, where debt or equity securities are traded.

A *financial derivative* is a financial instrument whose characteristics and value depend upon the characteristics and value of an *underlying* security, typically a commodity, bond, equity or currency. Derivatives are also known as *contingent claims* since their pay–offs are contingent upon the underlying. Examples of derivatives include *futures* and *options*. Advanced investors purchase or sell derivatives to manage the risk associated with the underlying security, to protect against fluctuations in value, or to profit from periods of inactivity or decline.

A *future* is a contract that requires delivery of a commodity, bond, currency, or stock index, at a specified price, on a specified future date. Unlike options, futures convey an obligation to buy. The derivatives that we will be concerned with in this work are options.

Definition 1.1 (European Call/Put option). *A* European Call (Put) option *is a financial derivative that certifies the holder's* right *— but not obligation — to buy (for a call option) or sell (for a put option) a specific amount of an underlying security, for a fixed price E* (exercise price)*, at a fixed future time T* (maturity or expiry)*.*

An *American option* is an option which in contrast to the European option can be exercised at any time between the purchase date and the expiration date. Since an option securitizes a right it has a certain *option value* or *option price*. This value, denoted by $V\big(S(t), t\big)$, depends on the price $S(t)$ of the underlying and the time t. The value of a call option at maturity time T depends on the price of the underlying at this time. There are two cases:

- The underlying's price at maturity is higher than the exercise price, i.e. $S(T) > E$. The call option is exercised, the holder buys the underlying at price E and sells it immediately, realizing a profit of $S(T) - E$.

- The underlying's price at maturity is equal or lower than the exercise price, i.e. $S(T) \leq E$. In this case, the call option is not exercised and expires worthless.

Thus, the call option value at maturity, the so–called *pay–off* is given by $V\big(S(T), T\big) = \max\big(0, S(T) - E\big) =: \big(S(T) - E\big)^+$. Similar arguments show that the pay–off for the put option is given by $V\big(S(T), T\big) = \max\big(0, E - S(T)\big) =: \big(E - S(T)\big)^+$. Generally speaking, the holder of a call option speculates on rising prices, the holder of a put option on declining prices of the underlying. Options cannot only be used as a speculative investment but also as an 'insurance'. For example, if an investor holds a number of shares of stock, he can insure himself against falling stock prices by buying a put option on the specific stock. In this case, the option value can be understood as an insurance premium. Usually the underlying is not delivered physically at maturity but rather the pay–off value is payed in cash (*cash settlement*). In particular, this holds true for options where the underlying cannot be delivered, e.g. options on indices.

Next, we come to a key assumption in modelling the market, the absence of so–called *arbitrage* opportunities. Note that this term appears in different variations in the literature. Here, we give only a formal definition and discuss the justification of this assumption. The ability to make an instantaneous riskless profit is called *arbitrage*. The words *instantaneous* and *riskless* play an important role here. Putting money in a bank account yields a riskless profit (assuming that there is no bank failure), but not instantaneously. Investing into securities can lead to instantaneous profit, but this investment is exposed to a certain risk. Usually, the market is assumed to be arbitrage–free, i.e. no arbitrage possibilities exist.

This model assumption can be motivated as follows. In very liquid markets with frequent trading, for example international stock markets, there are many investors looking for arbitrage opportunities. If they spot an arbitrage possibility they will take advantage of it by buying or selling securities. This leads to a price movement in the security which removes the arbitrage possibility. Therefore in a limit process with trading frequencies tending to infinity, the no–arbitrage assumption seems to be plausible. In less liquid markets, e.g. markets for defaultable bonds, this assumption may be violated. In this work — conforming with the larger part of the literature — we will generally assume that the market is arbitrage–free. Now, with the basic notions recalled, we are able to take a glance at the classical Black–Scholes model for pricing options.

1.2.2 Black–Scholes equation and formula

As a topic of mathematical works, the question of how to price options has not arisen just recently. Already in 1900, Bachelier [**5**] proposed an option pricing model, which had the great disadvantage that it lead to negative option prices. A satisfying answer was given in 1973 by Fischer Black and Myron Scholes [**20**] as well as, independently, by Robert Merton [**116**]. In the following we give a comprehensive overview on their model which leads to the so–called Black–Scholes equation and formula.

First, consider a bond $B(t)$ paying a fixed, riskless interest at rate $r \geq 0$. In a continuous–time market its value evolves according to $B(t) = B(0)e^{rt}$ or

$$\ln B(t) = \ln B(0) + rt. \tag{1.1}$$

The 'classic' way to model a stock price $S(t)$ is to assume that it follows a geometric Brownian motion

$$\ln S(t) = \ln S(0) + \Big(\mu - \frac{1}{2}\sigma^2\Big)t + \sigma W(t), \tag{1.2}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ are the so–called *mean return rate* (or drift) and *volatility* , respectively. The Brownian motion (or Wiener process) $W(t)$ is normally distributed with zero mean and variance t, i.e. $W(t) \sim \mathcal{N}(0,t)$. S_t is lognormally distributed, i.e. $\ln S(t)$ is normally distributed, with expectation value $\mathrm{E}\big(S(t)\big) = S(0)e^{\mu t}$ and variance $\mathrm{Var}\big(S(t)\big) = S(0)^2 e^{2\mu t}(e^{\sigma^2 t} - 1)$. The idea of this ansatz is that in the mean the stock price follows a certain trend and the actual stock price fluctuates randomly around this mean.

Definition 1.2. *Let $(\Omega, \mathcal{F}, P)$ be a probability space with σ–algebra $\mathcal{F}$, probability measure P, and let $\mathcal{F}_t$ be a family of sub–σ–algebras of $\mathcal{F}$ with $\mathcal{F}_s \subset \mathcal{F}_t$, for $s < t$ and $s,t \in [0,\infty)$, a so–called filtration. A continuous* stochastic process *is a family of random variables $X : \Omega \times [0,\infty) \to \mathbb{R}$, where $t \mapsto X(\omega, t)$ is continuous for all $\omega \in \Omega$ and $X(\cdot, t)$ is $\mathcal{F}_t$–measurable for all $t \in [0,\infty)$. For each $t \in [0,\infty)$, $X(t) = X(\cdot, t)$ is a random variable.*

The set $\mathcal{F}_t$ models the information available up to time t. An example for such a stochastic process is the Brownian motion $W(t)$ [**109**]. For the derivation of the Black–Scholes equation we need to introduce the following (formal) definition.

Definition 1.3. *Let $X(t)$ be a stochastic process, $W(t)$ the Brownian motion, and a, b (sufficiently smooth) functions. An integral equation of the form*

$$X(t) = X(0) + \int_0^t a\big(X(s), s\big)\, ds + \int_0^t b\big(X(s), s\big)\, dW(s) \tag{1.3}$$

is called stochastic differential equation of Itô. *Symbolically, we write*

$$dX(t) = a\big(X(t), t\big)\, dt + b\big(X(t), t\big)\, dW(t).$$

A solution $X(t)$ of (1.3) is called Itô process. *In (1.3), the second integral is (formally) defined by*

$$\int_0^t Y(s)\, dW(s) := \lim_{n\to\infty} \sum_{k=0}^{n-1} Y(t_k)\big(W(t_{k+1}) - W(t_k)\big),$$

where $Y(s)$ is a stochastic process and $0 = t_0 < t_1 < \cdots < t_n = t$ are partitions of $[0, t]$ with $\max\{|t_{i+1} - t_i| : i = 0, \ldots, n-1\} \to 0\,(n \to \infty)$.

For a proper definition which goes beyond the scope of this introduction and results on the well–posedness of stochastic differential equations we refer to [**2, 120, 131**].

A basic tool is *Itô's lemma*, which we will state in the following.

Lemma 1.4 (Itô). *Let $X(t)$ be an Itô process and $f \in C^2(\mathbb{R} \times [0, \infty))$. Then the stochastic process $f(t) = f(X(t), t)$ is an Itô process and it holds:*

$$df = \big[f_t + a f_x + \tfrac{1}{2} b^2 f_{xx}\big]\, dt + b f_x\, dW, \tag{1.4}$$

omitting the arguments for the sake of clarity. Subscript indices are indicating the partial derivatives, e.g. $f_x = \frac{\partial f}{\partial x}$.

Formally, equation (1.4) can be derived by a Taylor expansion of f up to the order dt, using the 'rule of thumb' $dW = \mathcal{O}(\sqrt{dt})$ and neglecting higher order terms [**69, 145**]. A complete proof is given in [**2, 96, 120**]. The lemma can be interpreted as a 'chain rule' for Itô processes, with a correction term $\frac{1}{2} b^2 f_{XX}$.

To model the financial market we assume the following:

- the stock price S_t follows a geometric Brownian motion (1.2),

- no dividends are payed on the underlying stock,
- no transaction costs, taxes, ask–bid–spreads, ... occur,
- the market is arbitrage–free, liquid and time–continuous,
- short–selling is allowed.

We give now a very short and formal derivation of the Black–Scholes partial differential equation. For details we refer to [**20, 96**]. Applying Itô's Lemma to (1.2) yields

$$dS(t) = \mu S(t)\, dt + \sigma S(t)\, dW(t).$$

Assume that the option value is a (sufficiently smooth) function of the stock price S and time t, i.e. $V = V(S, t)$. Consider now a portfolio $Y(t)$ consisting of a certain amount $c_1(t)$ of a bond $B(t)$, a certain number $c_2(t)$ of shares of the stock $S(t)$ and one option sold at time $t = 0$, namely

$$Y(t) = c_1(t)B(t) + c_2(t)S(t) - V(S, t). \tag{1.5}$$

The goal is to find a trading strategy $\big(c_1(t), c_2(t)\big)$ in bonds and stocks such that the portfolio is riskless, i.e. its value is not subject to random fluctuations, and self–financing, i.e. it holds

$$dY(t) = c_1(t)dB(t) + c_2(t)dS(t) - dV(S, t).$$

Applying Itô's Lemma to $V(S, t)$ we obtain

$$\begin{aligned} dY(t) &= c_1(t)dB(t) + c_2(t)dS(t) - dV(S, t) \\ &= \big[c_1(t)rB(t) + c_2(t)\mu S(t) - (V_t + \mu S V_S + \tfrac{1}{2}\sigma^2 S^2 V_{SS})\big]dt \\ &\qquad\qquad + \big[c_2(t)\sigma S - \sigma S V_S\big]dW. \end{aligned} \tag{1.6}$$

By choosing $c_2(t) = V_S(S, t)$ randomness is removed and the portfolio is now riskless. Thus, since the market is assumed to be arbitrage–free, it has to evolve like a riskless bond, i.e. $dY(t) = rY(t)\, dt$. From this equation and (1.6) follows the *Black–Scholes equation*

$$V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0. \tag{1.7}$$

This backward parabolic equation has to be solved for $S > 0$ and $t \in (0, T)$. To guaranty unique solutions, one has to impose final conditions at $t = T$,

$$\begin{aligned} V(S, T) &= (S - E)^+ \quad \text{for a European call option,} \\ V(S, T) &= (E - S)^+ \quad \text{for a European put option,} \end{aligned}$$

(see Section 1.2.1) and 'boundary' conditions [**145**] at $S = 0$ and as $S \to \infty$,

$$\begin{aligned} &V(0,t) = 0, && V(S,t) \sim S \ (S \to \infty) && \text{for a European call option,} \\ &V(0,t) = Ee^{-r(T-t)}, && V(S,t) \to 0 \ (S \to \infty) && \text{for a European put option,} \end{aligned}$$

where $V(S,t) \sim S$ has to be understood in the sense $V(S,t)/S \to 1$ as $S \to \infty$ uniformly for each $t \in [0,T]$.

With the above final and boundary conditions, the *Black–Scholes equation* can be solved analytically by transforming it to the heat equation using an adequate variable transformation [**20, 69, 145**]. This yields the famous *Black–Scholes formula*:

$$\begin{aligned} V(S,t) &= S\Phi(d_1) - Ee^{-r(T-t)}\Phi(d_2), && \text{for a European call option,} \\ V(S,t) &= Ee^{-r(T-t)}\Phi(-d_2) - S\Phi(-d_1), && \text{for a European put option,} \end{aligned}$$

with $S > 0$, $t \in [0,T)$, and the standard normal distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{s^2}{2}}\, ds, \quad x \in \mathbb{R},$$

and

$$d_{1,2} = \frac{\ln\left(\frac{S}{E}\right) + \left(r \pm \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}.$$

The parameters E, T are given by the particular design of the option, the interest rate r is easy to observe. The only 'real' parameter that is not observable is the volatility σ of the underlying. It can be estimated from historical stock prices (*historic volatility*) or computed from option prices observed in the market by inverting the Black–Scholes formula (*implied volatility*).

1.2.3 Generalizations

There are several possibilities to extend the above results to more general situations without leaving the general framework. The Black–Scholes equation can be easily extended to options on multiple underlyings, to underlyings paying dividends, and to volatilities and interest rates being (deterministic) functions of time. We will not pursue this any further here and refer to [**69, 145**].

Although the Black–Scholes model has been widely accepted and used in academics and by practitioners, it has also obtained rising criticism during the years. One reason is the following. The essential model parameter, which is not directly observable, is the volatility σ. It is often determined by computing the so–called *implied volatility* out of observed option prices by inverting the Black–Scholes formula. A widely observed phenomenon — the so–called *smile/skew effect* — is that these computed volatilities are not constant, in contradiction to the Black–Scholes model assumptions. This leads to a natural generalization of the Black–Scholes model replacing the constant volatility σ in the model by a local volatility function $\sigma = \sigma(E, T)$, where E denotes the exercise price. It arises the question of how to determine this volatility function from option prices observed in markets, such that the generalized Black–Scholes model replicates the market prices. This problem is often referred to as the *calibration problem.* A numerical approach to deal with it via an optimal control approach is conducted in Chapter 3.

Further generalizations can be made to weaken the model assumptions. The interest rate and the volatility are no longer assumed to be known (deterministic) functions but are modeled stochastically. Well–known is the model of Cox–Ingersoll–Ross [**37**] for the stochastic interest rate, which was extended by Hull and White, who also proposed a stochastic volatility model [**80**]. Dividends can also be modeled stochastically. For a good overview we refer to [**45**].

The stochastic processes in the Black–Scholes model are continuous. This does not allow to model price jumps which are observable in real markets. So–called *Jump–diffusion* models use Poisson processes to reflect this behavior [**117, 144**].

Another part of the criticism focuses on the assumption that the underlying is lognormally distributed. Empirical results show that in actual markets high gains and losses are more probable than predicted by this assumption. In the literature, this phenomenon is often referred to as ‘fat tails’. To cope with these observations, alternative distributions have been proposed, e.g. the log–gamma distribution [**74**], fractional Brownian motions [**135**], Lévy processes [**32**], and hyperbolic distributions [**50**]. An interesting, alternative approach is to start from alternative Radon–Nikodym densities to price the options rather than from the underlying's distribution, since they can be inferred from empirical data [**49, 86**].

In practice, transaction costs arise when trading securities. Although they are generally small for institutional investors, recent studies of their influence show that they lead to a notable increase in the option price. This is one possible explanation for the widely observed phenomenon that option prices observed in real markets are generally significantly higher than predicted by the Black–Scholes formula. The problem of option pricing with transaction costs from a numerical viewpoint is the focus of Chapter 2.

1.2.4 Portfolio Optimization

A *portfolio* is a collection of investments all owned by the same investor. These investments often include *stocks*, which are investments in corporations and *bonds*, which are investments in debt that are designed to earn interest. An investor typically faces the following questions. How to allocate stocks and bonds to a portfolio in order to maximize his or her wealth and minimize the risk associated with it? Is there an "optimal" strategy? To find answers to these questions is one classical problem of mathematical finance — portfolio optimization.

The first mathematical answer was given by Markowitz in 1952 [**113**]. He considers a one–period market model, where securities are traded only once, at the beginning of the period. If maximal wealth of the portfolio is the only goal, this leads to an investment of the complete initial capital in the stock with maximal expected return. Since this stock may be a highly risky investment, the investor exposes himself or herself to a high risk. Minimizing the risk of the portfolio is therefore a natural second goal of the optimization process. Markowitz' idea is based on either prescribing a minimal bound for the expected return and selecting from all possible portfolios that one with minimal variance of the return or to set an upper bound for the variance of the portfolio's expected return and select from the remaining portfolios the one with maximal expected return. Mathematically, this approach involves solving a linear optimization problem with quadratic constraint or solving a quadratic optimization problem. It is rather easy to implement and hence widely used in practice. Its great disadvantage is that stock price dynamics are not part of the model since trading occurs only at initial time; the model is static. The extension to multiple periods results in a quickly growing complexity, which renders it untractable and makes it necessary to consider time–continuous models. They have been introduced by Merton [**114, 115**] and we will give a comprehensive overview in the following. For a detailed presentation we refer to [**118**].

Let us start by stating some definitions and assumptions. We consider a fixed time interval $[t, T]$. The market consists of a riskless bond $B(s)$ modeled by

$$B(s) = B(t)e^{r(s-t)}, \quad s \in [t, T],$$

and — for simplicity — of one risky stock $S(s)$ modeled by an Itô process

$$dS(s) = \mu S(s)\, ds + \sigma S(s)\, dW(s), \quad s \in [t, T],$$

with $\mu > r$ and $\sigma > 0$. The investor starts at time t by investing wealth $X(t) = x > 0$. His or her control is the weighting of the portfolio between these two assets, i.e. the portfolio process $H_1(s)$, which equals the fraction of wealth invested in the risky asset at time s. The rest of his wealth, $H_0(s) = 1 - H_1(s)$, is invested in the bond. Transaction costs are ignored.

The portfolio is assumed to be self–financing, i.e. the whole wealth is invested at initial time and thus, changes over time of the portfolio process are solely caused by changes of the values of the assets held in the portfolio and only hedging is possible during lifetime, but nothing is ever withdrawn from or added to the portfolio during lifetime,

$$X(s) = x + \int_t^s H_0(\tau)X(\tau)r\, d\tau + \int_t^s \frac{H_1(\tau)X(\tau)}{S(\tau)}\, dS(\tau), \quad s \in [t, T], \tag{1.8}$$

where H_i $(i = 0, 1)$ is assumed to be smooth enough for the integrals on the right hand side to be well–defined.

To measure the investor's satisfaction gained from wealth $\xi > 0$, a so–called *utility function* is introduced.

Definition 1.5. *A continuously differentiable, strictly concave function*

$$U : (0, \infty) \to \mathbb{R} \quad \textit{with} \quad U'(0) := \lim_{\xi \downarrow 0} U'(\xi) = +\infty,\ U'(\infty) := \lim_{\xi \to \infty} U'(\xi) = 0,$$

is called utility function.

Following from this definition, U is strictly monotone, which expresses *more is better than less*. The limit in zero can be interpreted as *something is a lot better than nothing*. Moreover, since U' is decreasing monotonically, the utility gain for each additional fraction of wealth decreases and vanishes in infinity, which models a *saturation effect*.

It is assumed that the investor's goal is to maximize her or his utility from final wealth. In particular, this assumes rational behavior of the investor.

Hence, our goal is to maximize the utility of the portfolio value at time T, or, mathematically, the value function is

$$u(x,t) = \max_{H_1} \mathrm{E}\big[U\big(X(T)\big)\big].$$

We pursue the following strategy. Using heuristic arguments, we derive the so–called *Hamilton–Jacobi–Bellman* equation associated with the problem, solve it, and then need to show by a verification argument that the solution obtained also solves the original problem.

By applying the *dynamic programming principle* to a short time interval Δt, we have heuristically

$$u(x,t) \approx \max_{H_1} \mathrm{E}\big[u\big(X(t+\Delta t), t+\Delta t\big)\big]. \tag{1.9}$$

To evaluate the argument of the expectation term, we use Itô's lemma and obtain

$$du = \big[u_t + [(1-H_1)r + H_1\mu]X(s)u_x + \tfrac{1}{2}H_1^2\sigma^2X^2(s)u_{xx}\big]ds + H_1\sigma X(s)u_x\,dW(s),$$

since $X(s)$ is an Itô process fulfilling

$$dX(s) = (1-H_1)rX(s)\,ds + H_1X(s)\big(\mu\,ds + \sigma\,dW(s)\big), \tag{1.10}$$

which is another way of writing equation (1.8). Using the above relation for du we have

$$\begin{aligned} u\big(X(t+\Delta t), t+\Delta t\big) - u\big(X(t),t\big) = \\ \int_t^{t+\Delta t} \big[u_t + [(1-H_1)r + H_1\mu]X(s)u_x + \tfrac{1}{2}H_1^2\sigma^2X^2(s)u_{xx}\big]ds \\ + \int_t^{t+\Delta t} H_1\sigma X(s)u_x\,dW(s), \end{aligned}$$

Applying the expected value and noting that $X(t) = x$ and that the second integral has expected value zero, we get

$$\begin{aligned} &\mathrm{E}\big[u\big(X(t+\Delta t), t+\Delta t\big)\big] - u(x,t) \\ &\qquad \approx \big[u_t + [(1-H_1)r + H_1\mu]X(t)u_x + \tfrac{1}{2}H_1^2\sigma^2X^2(t)u_{xx}\big]\Delta t. \end{aligned}$$

Collecting the parts, we have

$$u(x,t) \approx \max_{H_1}\big\{u(x,t) + \big[u_t + [(1-H_1)r + H_1\mu]xu_x + \tfrac{1}{2}H_1^2\sigma^2x^2u_{xx}\big]\Delta t\big\},$$

and formally obtain the *Hamilton–Jacobi–Bellman equation*

$$u_t + \max_{H_1}\big\{[(1-H_1)r + H_1\mu]xu_x + \tfrac{1}{2}H_1^2\sigma^2x^2u_{xx}\big\} = 0, \tag{1.11}$$

which has to be solved for $t < T$ with final condition $u(x,T) = U(x)$.

For solving (1.11), we restrict ourselves to the special choice of the utility function $U(\xi) = \xi^p,\ \xi > 0,\ p \in (0,1)$. To carry on, it is useful to make the following assumptions (which need to be checked later)

$$X(s) > 0, \quad u \text{ strictly concave in } x, \quad \text{and} \quad H_1 \in [h_1, h_2],$$

for some $h_1, h_2 \in \mathbb{R}$. We also assume $u_x > 0$ since larger initial wealth should produce larger final utility. Then the optimal control candidate is given by

$$H_1 = -\frac{(\mu - r)u_x}{\sigma^2 x u_{xx}} > 0. \tag{1.12}$$

Inserting (1.12) into (1.11) yields the following equation

$$u_t - \frac{1}{2}\frac{(\mu - r)^2u_x^2}{\sigma^2u_{xx}} + rxu_x = 0. \tag{1.13}$$

Since $u(\lambda x, t) = \max_{H_1} E[U(\lambda X(T))] = \lambda^p u(x,t)$, it seems evident to employ the separation ansatz $u(x,t) = f(t)x^p$ with $f(T) = 1$. Inserting this ansatz in (1.13), we obtain

$$f'(t) + \left(r + \frac{(\mu - r)^2}{2\sigma^2(1-p)}\right)pf(t) = 0. \tag{1.14}$$

The solution of this ordinary differential equation is

$$f(t) = \exp\left(\left[r + \frac{(\mu - r)^2}{2\sigma^2(1-p)}\right]p(T-t)\right).$$

We conclude that

$$u(x,t) = f(t)x^p$$

is the solution to (1.11) (for the chosen utility function) and the optimal control is given by

$$H_1 = \frac{(\mu - r)}{\sigma^2(1-p)}. \tag{1.15}$$

It may seem contra–intuitive at first glance that the optimal control which represents the optimal investment strategy does not depend on the current wealth of the investor. This is clarified by the fact that the power utility function signifies constant relative risk aversion, i.e. the investor's attitude toward risk is independent of his wealth.

We have to check the assumptions we made above. Obviously, u is strictly concave in x, since f is strictly positive. Since H_1 is constant, it is possible to choose h_1, h_2 above in a way that they impose no restriction. Then they do not influence the optimal solution and can be omitted completely. By variation of constants, e.g. Theorem II.42 in [**96**], the stochastic differential equation (1.10) for the wealth process $X(t)$ with initial condition $X(0) = x$ possesses the unique solution

$$X(t) = x \exp\Big(\big[(1-H_1)r + H_1\mu - \frac{1}{2}\sigma^2 H_1^2\big]t + \sigma H_1 W(t)\Big),$$

which is strictly positive.

We have solved the Hamilton–Jacobi–Bellman equation (1.11). In general, however, it is a non–trivial task to find the solution. To complete our analysis, we need to verify that u is also a solution of the optimal expected terminal wealth problem. This verification argument is based on the same main tools that we used above, Itô's formula and the fact that the expected values of the stochastic integrals vanish. We omit it here and refer to [**118**] for details.

Again, it it easy to generalize the situation to multiple stocks and returns, interest rates and volatilities being deterministic functions of time. The situation is more complicated if the market is *incomplete*, e.g. if non–tradable investments exist in the investor's portfolio. Examples for such non–tradable investments are credit risks of a bank or an employee's personal income. The problem of optimizing a portfolio's expected utility from terminal wealth in an incomplete market has been studied in [**97, 107**]. The well–posedness of a partial differential equation arising in the modelling of optimal portfolios in incomplete markets is studied in Chapter 4.

1.3 Nonlinear Black–Scholes type equations

In an idealized financial market the value of an option can be determined by the classical Black–Scholes theory. The unrealistic assumptions that are made in the derivation of this model prove to be too restrictive in practice. Therefore, in recent years different models have been proposed to weaken one or more of these assumptions. These models result in strongly or fully nonlinear, possibly degenerate, parabolic diffusion–convection equations for the option value V as a function of the underlying security S and the time t

$$V_t + \frac{1}{2}\hat{\sigma}(S,t,V_S,V_{SS})^2 S^2 V_{SS} + rSV_S - rV = f(S,t,V_S), \quad S > 0, t \in (0,T),$$

with final and boundary conditions. Here $r \geq 0$ denotes the riskless interest rate, the nonlinear function $\hat{\sigma}$ is the volatility, and the nonlinear function f models different effects. In the case of multiple underlying assets the partial derivatives have to be interpreted as the gradient and the Hessian, respectively. Equations of this form arise quite frequently in mathematical finance, not only in option pricing, for example in the modeling of transaction costs [**11, 43**], optimal portfolios in incomplete markets [**107**], feedback effects due to large traders [**60, 133**] and inverse problems (implied volatility) [**17**]. The mathematical analysis and the numerical solution of equations of this type is the focus of this work.

Usually, the source of such nonlinearities is so–called *market incompleteness*. A market is called *complete* if each contingent claim, for example the pay–off of an option, can be matched by the final wealth of an admissible trading strategy. Mathematically, a market is defined to be complete if and only if there exists a unique equivalent martingale measure [**71**]. On the other hand, a market is called *incomplete* if not every contingent claim is reachable. Typical reasons for market incompleteness are trading restrictions, additional randomness in the model (e.g. stochastic volatility) or the presence of transaction costs.

1.3.1 Market incompleteness: Transaction costs

In the Black–Scholes model for the pricing of options the influence of transaction costs was neglected and it was possible to construct a riskless portfolio that perfectly replicates the option pay–off. If transaction costs are taken into account perfect replication of the contingent claim is no longer possible, and it has been shown in [**134**] that further restrictions have to be imposed in the model.

One possibility is to restrict trading to discrete points in time. In 1992, Boyle and Vorst [**33**] derived from a binomial model an option price that takes into account transaction costs and that is equal to a Black–Scholes price but with a modified volatility of the form

$$\hat{\sigma} = \sigma(1 + cA)^{1/2}, \quad A = \frac{\alpha}{\sigma\sqrt{\triangle t}},$$

with $c = 1$. Here, α is the proportional transaction cost, $\triangle t$ the transaction period, and σ is the original volatility constant of the underlying. Leland [**108**] computed the constant $c = (2/\pi)^{1/2}$. Kusuoka [**98**] then showed that the 'optimal' c depends on the risk structure of the market. Parás and Avellaneda [**122**] derived the modified volatility

$$\hat{\sigma} = \sigma(1 + A\,\mathrm{sign}(V_{SS}))^{1/2} \tag{1.16}$$

from a binomial model using the algorithm of Bensaid et al. [**14**]. Here, V is the option price, S the price of the underlying asset, and V_{SS} denotes the second derivative of V with respect to S (the 'Gamma'). In particular, the option price does not need to be convex.

Another popular approach is to introduce preferences by assuming that the investor's behavior is characterized by a given utility function. In [**78**] it has been shown that the option price can be obtained as the initial cash increment which offsets the difference between the maximum utility of terminal wealth when there is no option liability and when there is such a liability. Davis et al. [**43**] extended this approach to markets with transaction costs (see also [**143**]). It has the disadvantage that the option price depends on the special choice of the utility function but Constantinides and Zariphopoulou [**36**] obtained universal bounds independent of the utility function. Using this utility maximization approach, the following model has been proposed by Barles and Soner [**11**]. In the following we recall the key points of this model.

1.3.2 The transaction cost model of Barles and Soner

Consider a portfolio of a bond and a stock, which evolves according to

$$dS(s) = \mu S(s)\,ds + \sigma S(s)\,dW(s), \quad s \in [t, T],$$

with constant mean return rate $\mu \in \mathbb{R}$ and constant volatility $\sigma > 0$. For the sake of brevity, we set the interest rate r to zero for the moment. Denote by $X(s), Y(s)$ the processes of bond holdings and shares owned and a trading strategy $(L(s), M(s))$ as a pair of nondecreasing processes with $L(t) = M(t) = 0$, which are interpreted as the cumulative transfers, measured in shares of stock, from bond market to stock and vice versa. Let $\alpha \in (0,1)$ be a proportional transaction cost. With initial values x, y, the corresponding portfolio $X(s) = X(s; t, x, y, L(\cdot), M(\cdot))$ and $Y(s) = Y(s; t, x, y, L(\cdot), M(\cdot))$ is assumed to evolve according to

$$X(s) = x - \int_t^s S(\tau)(1+\alpha)\, dL(\tau) + \int_t^s S(\tau)(1-\alpha)\, dM(\tau), \tag{1.17}$$

$$Y(s) = y + L(s) - M(s), \tag{1.18}$$

for $s \in [t, T]$. In (1.17) the first integral represents buying of shares of stock at a price increased by the proportional transaction cost, the second integral represents selling stock where the stock price is reduced by the proportional transaction cost.

Following the utility maximization approach of Hodges and Neuberger [**78**], the price of a European call option can be obtained as the initial cash increment which offsets the difference between the maximum utility of terminal wealth when there is no option liability and when there is such a liability, i.e. we consider two optimization problems. We choose the exponential utility function

$$U(\xi) = 1 - e^{-\gamma\xi}, \quad \xi \in \mathbb{R},$$

with risk aversion factor $\gamma > 0$. The first optimization problem when we have no option liability has the value function

$$V_1\big(x, y, S(t), t\big) := \sup_{L(\cdot), M(\cdot)} \mathrm{E}\big[U\big(X(T) + Y(T)S(T)\big)\big], \tag{1.19}$$

in the second we suppose we have sold N European call options and the value function is given by

$$V_2\big(x, y, S(t), t\big) := \sup_{L(\cdot), M(\cdot)} \mathrm{E}\big[U\big(X(T) + Y(T)S(T) - N\big(S(T) - E\big)^+\big)\big]. \tag{1.20}$$

According to [**78**] the price of each option is equal to the maximal solution Λ of the algebraic equation

$$V_2\big(x + N\Lambda, y, S(t), t\big) = V_1\big(x, y, S(t), t\big),$$

i.e. the option price Λ equals the increment of initial capital at time t that is needed to cope with the option liabilities arising at T. Hence, the option price is a function of the initial data $x, y, S(t), t$ and γ, N, that is $\Lambda = \Lambda\big(x, y, S(t), t; \gamma, N\big)$. By linearity of the equations, it holds

$$\Lambda\big(Nx, Ny, S(t), t; \gamma, N\big) = \Lambda\big(Nx, Ny, S(t), t; \gamma N, 1\big),$$

i.e. selling N options with risk aversion factor of γ yields the same price as selling one option with risk aversion factor of γN. This scaling argument leads to performing an asymptotic analysis as γN tends to infinity. Hence, consider $U(\xi) = 1 - e^{-\gamma N \xi}$. Set $\varepsilon = 1/(\gamma N)$ and $U_\varepsilon(\xi) = 1 - e^{-\xi/\varepsilon}$, $\xi \in \mathbb{R}$. Then, the two optimization problems take the following form

$$v_1\big(x, y, S(t), t\big) = 1 - \inf_{L(\cdot), M(\cdot)} \mathrm{E}\Big[\exp\Big(-\frac{1}{\varepsilon}\big[X(T) + Y(T)S(T)\big]\Big)\Big],$$
$$v_2\big(x, y, S(t), t\big) = 1 - \inf_{L(\cdot), M(\cdot)} \mathrm{E}\Big[\exp\Big(-\frac{1}{\varepsilon}\big[X(T) + Y(T)S(T) - \big(S(T) - E\big)^+\big]\Big)\Big].$$

Now, define $z_{1,2} : \mathbb{R} \times (0, \infty) \times (0, T) \to \mathbb{R}$ by

$$v_1\big(x, y, S(t), t\big) = 1 - \exp\Big(-\frac{1}{\varepsilon}\big[x + yS(t) - z_1\big(y, S(t), t\big)\big]\Big),$$
$$v_2\big(x, y, S(t), t\big) = 1 - \exp\Big(-\frac{1}{\varepsilon}\big[x + yS(t) - z_2\big(y, S(t), t\big)\big]\Big).$$

Then, it holds

$$z_1\big(y, S(t), T\big) = 0, \quad z_2\big(y, S(t), T\big) = \big(S(T) - E\big)^+. \tag{1.21}$$

and the option price Λ is given by

$$\Lambda\Big(x, y, S(t), t; \frac{1}{\varepsilon}, 1\Big) = z_2\big(y, S(t), t\big) - z_1\big(y, S(t), t\big).$$

The value functions v_1, v_2 are solutions of an associated dynamic programming equation (not given here), which leads to a dynamic programming equation for z_1, z_2. The coefficients in this equation and the terminal data (1.21) are independent of the variable x and hence z_1, z_2 are independent of x as well [**43**].

Supposing that $\alpha = a\sqrt{\varepsilon}$ for some constant $a > 0$, Barles and Soner now perform an asymptotic analysis and prove that

$$z_1 \to 0 \quad \text{and} \quad z_2 \to V \quad \text{as } \alpha, \varepsilon \to 0 \text{ with } a = \frac{\alpha}{\sqrt{\varepsilon}} \text{ constant},$$

where V is the unique (viscosity) solution of the nonlinear Black–Scholes equation

$$V_t + \tfrac{1}{2}\hat{\sigma}(V_{SS})^2 S^2 V_{SS} = 0.$$

In the case of non–zero interest rate r the equation reads

$$V_t + \tfrac{1}{2}\hat{\sigma}(V_{SS})^2 S^2 V_{SS} + rSV_S - rV = 0, \tag{1.22}$$

where the nonlinear volatility $\hat{\sigma}(V_{SS})$ is given by

$$\hat{\sigma} = \sigma\Big(1 + \Psi\big[\exp\big(r(T-t)\big)a^2 S^2 V_{SS}\big]\Big)^{\frac{1}{2}}. \tag{1.23}$$

Herein, σ is the constant volatility of the underlying. The function Ψ is given as the solution of the nonlinear initial–value problem

$$\begin{aligned} \Psi'(A) &= \frac{\Psi(A) + 1}{2\sqrt{A\Psi(A)} - A}, \quad A \neq 0, \\ \Psi(0) &= 0. \end{aligned} \tag{1.24}$$

For the case of no transaction costs occuring, i.e. $a = 0$, we recover the standard Black–Scholes equation.

Equation (1.22) is solved for the price $S \geq 0$ of the underlying asset and time $T \geq t \geq 0$, i.e. backward in time. The terminal condition is

$$V(S,T) = \max(0, S - E), \quad S \geq 0. \tag{1.25}$$

The ‘boundary’ conditions for $T \geq t \geq 0$ are as follows:

$$V(0,t) = 0, \qquad V(S,t) \sim S - Ee^{r(t-T)} \quad (S \to \infty). \tag{1.26}$$

The last condition has to be understood in the sense

$$\lim_{S\to\infty} \frac{V(S,t)}{S - Ee^{r(t-T)}} = 1,$$

uniformly for $T \geq t \geq 0$.

In [**11**] the existence of a unique viscosity solution to (1.22) is proved and some numerical results using an explicit finite difference approximation are given. For a general presentation of the notion of viscosity solutions see [**38**]. The authors of [**11**] compare the resulting prices to that of Leland's model [**108**] and discuss briefly the choice of the parameter a. The numerical results indicate an economically significant price difference between the standard Black–Scholes prices and the prices with transaction costs.

In Chapter 2 we study this model intensively from a numerical point of view. We compare different approaches to solve the problem by compact finite difference schemes in terms of stability and efficiency. Furthermore, we propose a new compact finite difference scheme and study in detail its properties. With the help of these results, we prove the convergence of the finite difference solution to the viscosity solution of (2.1).

1.3.3 Market incompleteness: Non–tradable assets

For a *complete* market as in Section 1.2.4 the problem of maximizing an investor's expected utility from terminal wealth was solved in [**114, 115**], deriving a nonlinear partial differential equation (Bellman equation) for the value function of the optimization problem, i.e. the utility of the optimal portfolio. The maximization of expected utility from terminal wealth in *incomplete* markets has been studied in [**97, 107**], deriving a nonlinear partial differential equation with quadratic gradient terms. In the following, we give a sketch of its derivation. The details can be found in [**107**].

Consider an arbitrage–free continuous time market model with unrestricted trading and a fixed time horizon, i.e. $t \in [0, T]$. The market consists of a riskless bond, d risky assets, and d' non–tradable state variables and hence is incomplete. The bond will not appear explicitly in the following since prices discounted to the bond are considered everywhere. Examples for such state variables are credit risks of a bank or an employee's personal income, which usually cannot be traded. Another possibility is to interpret them as information variables, e.g. the unemployment rate or an economy forecast variable. The optimization problem is to find a portfolio strategy which maximizes the expected utility from terminal wealth over the set of self–financing portfolios $\pi(t)$ with initial capital $\pi(0) = \pi_0 > 0$ and non–

negative wealth, using utility functions with constant relative risk aversion,

$$U^0(\xi) = \ln \xi, \quad U^{(p)}(\xi) = \begin{cases} \mathrm{sign}(p)\xi^p, & p < 1, \\ -x^p, & p > 1, \end{cases}$$

with $\xi > 0$ and exponent $p \in \mathbb{R}\backslash\{0, 1\}$.

The portfolio is assumed to be self–financing and the influence of transaction costs is neglected. The optimal value function v of this problem is defined by

$$v(\pi_0) = \sup_{H \in \mathcal{B}} E\Big[U^{(p)}\big(\pi(T)\big)\Big],$$

where $H \in \mathcal{B}$ is the hedging strategy representing the fractions of wealth invested into assets and $\mathcal{B}$ is a set of arbitrage–free, self–financing portfolio strategies. Solving this optimization problem with $p < 1$ is an approach to finding portfolios of optimal expected growth [**91, 92, 97**]. For $p = 2$ the problem is related to the mean variance hedging problem [**66, 105, 125**].

1.3.4 A nonlinear partial differential equation with quadratic growth of the gradient

Let $0 \leq t \leq T$ and $\hat{W} = (W, W')$ be a $d+d'$–dimensional Brownian motion on $\Omega := \{\Omega, \mathcal{F}, P\}$. Assume $\hat{S}(t) = \big(S(t), S'(t)\big)$ to be an Itô process satisfying

$$\begin{aligned} d\hat{S}(t) &= \hat{\mu}\big(\hat{S}(t), t\big)\, dt + \hat{\sigma}\big(\hat{S}(t), t\big)\, d\hat{W}(t), \\ \hat{S}(0) &= \big(S_0, S_0'\big). \end{aligned}$$

with drift $\hat{\mu} = (\mu, \mu')$ where

$$\begin{aligned} \mu\big(S(t), S'(t), t\big) &: \mathbb{R}^d \times \mathbb{R}^{d'} \times [0, T] \to \mathbb{R}^d, \\ \mu'\big(S(t), S'(t), t\big) &: \mathbb{R}^d \times \mathbb{R}^{d'} \times [0, T] \to \mathbb{R}^{d'}, \end{aligned}$$

and volatility $\hat{\sigma} = (\sigma, \sigma')$ where

$$\begin{aligned} \sigma\big(S(t), S'(t), t\big) &: \mathbb{R}^d \times \mathbb{R}^{d'} \times [0, T] \to \mathbb{R}^{d \times d}, \\ \sigma'\big(S(t), S'(t), t\big) &: \mathbb{R}^d \times \mathbb{R}^{d'} \times [0, T] \to \mathbb{R}^{d' \times d'}. \end{aligned}$$

We interpret S as stock prices and S' as non–tradable state variables. Assume the covariance matrix $\hat{C}$ to have a uniformly bounded inverse and block structure, i.e.

$$\hat{C} = \begin{pmatrix} C & 0 \\ 0 & C' \end{pmatrix}, \quad \text{with } C \in \mathbb{R}^{d\times d},\ C' \in \mathbb{R}^{d'\times d'},$$

which means that assets and state variables are uncorrelated. The coefficients of the asset price process, however, can depend on the non–tradable state variables. This induces a risk that cannot be hedged against, since trading in S' is impossible.

Following a stochastic duality approach, the existence of an optimal portfolio π is proved in [**107**]. The value of the portfolio π is given by $\pi(t) = H(t)\cdot S(t)$ for the optimal hedging strategy H. The optimal portfolio is characterized by a backward stochastic differential equation for $u = \ln(v)$, the logarithm of the value function v,

$$\begin{aligned} du(t) =& \frac{1}{2}\big[(p-1)H(t)C(t)H(t) - \nu(t)C'(t)\nu(t) - \lambda(t)C(t)\lambda(t)\big]\ dt \\ &+ \big[(p-1)H(t)C(t)\lambda(t) - pr(t)\big]\ dt \\ &+ \big[\lambda(t) - (p-1)H(t)\big]\sigma\big(S(t), S'(t), t\big)\ dW(t) \\ &+ \nu(t)\sigma'\big(S(t), S'(t), t\big)\ dW'(t), \end{aligned} \tag{1.27}$$

where

$$\begin{aligned} &C(t) = C\big(S(t), S'(t), t\big), \quad C'(t) = C'\big(S(t), S'(t), t\big), \\ &\lambda(t) = \lambda\big(S(t), S'(t), t\big) = C^{-1}\big[r\big(S(t), S'(t), t\big)S(t) - \mu\big(S(t), S'(t), t\big)\big]. \end{aligned}$$

Set $\beta\big(S(t), S'(t), t\big) = \sqrt{\lambda C \lambda}$.

Since u is expected to be a function of S, S', and time t we apply the (multidimensional) Itô formula, obtaining

$$\begin{aligned} du(t) =& \Bigg[u_t + \sum_{i=1}^{d} \mu^{(i)} u_{S^{(i)}} + \sum_{i=1}^{d'} \mu'^{(i)} u_{S'^{(i)}} + \frac{1}{2}\sum_{i,j=1}^{d} c_{ij} u_{S^{(i)}S^{(j)}} \\ &+ \frac{1}{2}\sum_{i,j=1}^{d'} c'_{ij} u_{S'^{(i)}S'^{(j)}}\Bigg]\ dt + \sum_{i=1}^{d} \sigma^{(i)} u_{S^{(i)}} dW^{(i)}(t) + \sum_{i=1}^{d'} \sigma'^{(i)} u_{S'^{(i)}} dW'^{(i)}(t), \end{aligned} \tag{1.28}$$

with the covariance matrices $C = \big(c_{ij}\big)_{i,j=1}^{d}$, $C' = \big(c'_{ij}\big)_{i,j=1}^{d'}$, where

$$c_{ij} = \rho_{ij}\sigma^{(i)}\big(S(t), S'(t), t\big)\sigma^{(j)}\big(S(t), S'(t), t\big),$$

and ρ_{ij} is the correlation of $S^{(i)}$ and $S^{(j)}$, and similarly for c'_{ij}. For a generalization of the multidimensional Itô formula to functions in Sobolev spaces, see [**52, 57**].

We derive now a deterministic partial differential equation by comparing the terms in (1.27) and (1.28). Comparing the 'stochastic parts' we find

$$H(t) = \frac{\lambda(t) - \nabla u}{p-1}, \quad \nu(t) = \nabla' u. \tag{1.29}$$

Note that for $d = 1$, $d' = 1$, and u constant with respect to the asset prices the strategy coincides with the classical Merton strategy (1.15). We use (1.29) in (1.27) and compare the right–hand sides of (1.27) and (1.28). The stochastic terms vanish and we end up with a non–linear partial differential equation with terms with quadratic growth of the gradient. We study in the following the generalized version of this equation where the covariance (diffusion) coefficients also depend on u,

$$\begin{aligned}\partial_t u - &\frac{1}{2}\sum_{i,j=1}^{d} c_{ij}(u)\,\partial_i\partial_j u - \frac{1}{2}\sum_{i,j=1}^{d'} c'_{ij}(u)\,\partial'_i\partial'_j u \\ &= \mu\cdot\nabla u + \mu'\cdot\nabla' u - q(\mu - rS)\cdot\nabla u - \frac{q}{2}\beta(u)^2 + pr \qquad (1.30\text{a}) \\ &\quad - \frac{1}{2(p-1)}(\nabla u)^\top C(u)\nabla u + \frac{1}{2}(\nabla' u)^\top C'(u)\nabla' u, \qquad \text{in } \hat\Omega\times(0,T),\end{aligned}$$

$$u(S,S',t) = u_D(S,S',t) \quad \text{on } \partial\hat\Omega\times(0,T), \tag{1.30b}$$

$$u(S,S',0) = u_0(S,S') \qquad \text{in } \hat\Omega, \tag{1.30c}$$

where $u = u(S,S',t)$ is the logarithm of the optimal value function, either $\hat\Omega = \Omega\times\Omega' \subset \mathbb{R}^d\times\mathbb{R}^{d'}$ is a bounded domain or $\hat\Omega = \mathbb{R}^d\times\mathbb{R}^{d'}$, and $T > 0$. We use the notations $\partial_t = \partial/\partial t$ and $\nabla = (\partial_1,\ldots,\partial_d)$, $\nabla' = (\partial'_1,\ldots,\partial'_{d'})$ with the partial derivatives $\partial_i = \partial/\partial S_i$, $\partial'_i = \partial/\partial S'_i$. Furthermore,

- $C = \big(c_{ij}(S,S',t,u)\big)_{i,j} : \Omega\times\Omega'\times(0,T)\times\mathbb{R}\to\mathbb{R}^{d\times d}$ and $C' = \big(c'_{ij}(S,S',t,u)\big)_{i,j} : \Omega\times\Omega'\times(0,T)\times\mathbb{R}\to\mathbb{R}^{d'\times d'}$ are the symmetric and positive definite covariance matrices of the risky assets and the non–tradable state variables, respectively;

- $\mu(S, S', t) : \Omega \times \Omega' \times (0, T) \to \mathbb{R}^d$ and $\mu'(S, S', t) : \Omega \times \Omega' \times (0, T) \to \mathbb{R}^{d'}$ are the expected returns;
- $r(S, S', t) : \Omega \times \Omega' \times (0, T) \to \mathbb{R}$ is the riskless interest rate;
- $\beta(S, S', t, u)^2 = (\mu - rS)^\top C^{-1}(\mu - rS)$ is the square of the risk premium;
- $p \notin \{0, 1\}$ is the exponent of the utility function and $q \in \mathbb{R}$ is given by $\frac{1}{p} + \frac{1}{q} = 1$.

In case of $p = 0$, which is associated with the logarithmic utility function $U^0(\xi) = \ln \xi$, the optimization problem is also known as maximizing the Kelly criterion [**65, 92, 94**]. Note that if $p = 0$, the quadratic terms in (1.30a) can be removed by an exponential transformation. If the expression $pr - q\beta^2/2$ and the initial data u_0 is constant in $\hat{\Omega} \times (0, T)$, equation (1.30a) has the solution $u(S, S', t) = (pr - q\beta^2/2)t + u_0$. To our knowledge the well–posedness of (1.30) has not been studied in the literature.

In Chapter 4 we study in–depth the well–posedness of this quasilinear partial differential equation and prove the existence and uniqueness of generalized solutions in a bounded domain and in the full–space. Furthermore, we consider a numerical example and give an interpretation of the optimal portfolio strategy.

1.4 Outline

In Chapter 2 the nonlinear Black–Scholes equation (1.22) which models transaction costs arising in option pricing is discretized semi–implicitly using high order compact finite difference schemes. A new compact scheme is derived generalizing the schemes of Rigal [**128**] and numerical results are compared to standard finite difference schemes. It turns out that the compact scheme has very satisfying stability and non–oscillatory properties and is generally more efficient than the considered classical schemes. Moreover, it is shown that the finite difference solution converges locally uniformly to the unique viscosity solution of the continuous equation. The proof is based on a careful study of the discretization matrices and on an abstract convergence result of Barles and Souganides [**12**].

In Chapter 3 we turn to the calibration problem of computing local volatility functions from market data in a generalized Black–Scholes setting as in

Section 1.2.3. Several approaches for solving this problem have been discussed in the literature. We follow an optimal control approach in a Lagrangian framework. We show the existence of a global solution and study first– and second–order optimality conditions. Furthermore, we propose an algorithm, that is based on a globalized sequential quadratic programming (SQP) method with a modified Hessian as an outer loop. In each iteration a linear–quadratic optimal control problem with box constraints is solved by a primal–dual active set strategy in an inner loop. Finally, we present some numerical results.

In Chapter 4 we consider the quasilinear parabolic equation with quadratic gradient terms (1.30), which arises in the modeling of an optimal portfolio in incomplete markets. The existence of weak solutions is shown by considering a sequence of approximate solutions. We derive uniform H^1 estimates using nonlinear test functions of the form $\sinh(\lambda u)$ for sufficiently large λ, which imply *weak* convergence of the sequence in H^1. The main difficulty is to infer the required *strong* convergence in H^1. This is proved by adapting the monotonicity method of Frehse [**59**]. Furthermore, we prove the uniqueness of weak solutions under a smallness condition on the derivatives of the covariance matrices with respect to the solution, but *without* additional regularity assumptions on the solution. The influence of the non–tradable state variables on the optimal value function is illustrated by a numerical example.

Chapter 2

Option Pricing with Transaction Costs

2.1 Introduction

In this chapter we are concerned with the numerical discretization of a nonlinear Black–Scholes equation modeling transaction costs arising in the hedging of portfolios. In the mathematical literature, many results can be found on the numerical discretization of Black–Scholes equations. The numerical approaches vary from binomial approximations (see, for instance [**104**] for American options in a stochastic framework), Monte–Carlo methods [**102**], finite–element discretizations [**58, 124**], and finite–difference approximations [**44, 139**] — however, mainly for *linear* Black–Scholes equations.

The numerical discretization of the Black–Scholes equations with the *nonlinear* volatilities (1.16) and (1.23) has been performed using explicit finite difference schemes [**11, 122**]. However, explicit schemes have the disadvantage that restrictive conditions on the discretization parameters (for instance, the ratio of the time and space step) are needed to obtain stable, convergent schemes [**138**]. Moreover, the convergence order is only one in time and two in space.

In this chapter we discretize the Black–Scholes equation with nonlinear volatility (1.23). More precisely, we study the equation

$$V_\tau + \tfrac{1}{2}\hat{\sigma}(V_{SS})^2 S^2 V_{SS} + rSV_S - rV = 0, \tag{2.1}$$

where the nonlinear volatility $\hat{\sigma}(V_{SS})$ is given by (1.23). This equation is solved for the price $S \geq 0$ of the underlying asset and time $T \geq t \geq 0$, i.e. backward in time. The problem is complemented by the terminal and boundary conditions (1.25),(1.26).

Our main goal is to obtain *efficient* and *precise* schemes, i.e. we wish to derive numerical schemes whose order is superior to standard schemes for second–order equations (like the explicit Euler, the semi–implicit Euler, the Crank–Nicolson and the Leap–Frog Du Fort–Frankel schemes) and whose computing time is comparable to that of classical schemes. Since the equation is of second order, usually a three–point approximation is used [**44**] and a scheme which is (consistent) of order one in time and two in space is obtained (except the Leap–Frog Du Fort–Frankel scheme; see below). To obtain higher order schemes (of order 2 in time and 4 in space), one possibility is to use more spatial points but this complicates the approximation of the boundary conditions and results in a discretization matrix with larger band width. In this chapter, we present an alternative using high–order compact schemes which need three points in space only. These compact schemes are tested on the equation (2.1) as a *first test*. The schemes will be used for more complex situations (multiple dimensions, path–dependency) in the future.

We apply some standard schemes and the compact schemes R3A and R3B derived by Rigal [**128**] to the initial–boundary value problem (2.1)–(1.26) (Section 2.3). The nonlinearity is treated explicitly i.e., the final scheme is semi–implicit. Furthermore, we construct a *new* three–point compact scheme R3C, which generalizes the scheme R3B of Rigal (Section 2.4).

In the nonlinear case, the information on the stability of the schemes becomes very limited. Clearly, linear stability is a necessary condition for the stability of nonlinear problems but it is not sufficient. Particular methods were developed for nonlinear problems. Stetter and Keller studied the nonlinear stability and local stability [**93, 136**], which have been used successfully for nonlinear ordinary differential equations. Ben–Yu [**16**] proposed a generalized stability of difference schemes which has been applied widely to numerical solutions of many nonlinear partial differential equations. Here, we do not use this technique which can be very complicated.

Instead, for the standard schemes and for the compact schemes of Rigal we give results for the linear case ($a = 0$) and validate them in the nonlinear case by numerical studies. Furthermore, we analyze the properties of the new compact scheme R3C and show that this scheme is unconditionally stable and non–oscillatory. The study of the properties of this scheme is based upon a thorough Fourier analysis of the Cauchy problem associated with (2.1). We resort to a local analysis with frozen values of the nonlinear coefficient to make the formulation linear. Furthermore, we prove certain properties of the discretization matrices as well as the positivity of numerical solutions (Section 2.4).

The numerical experiments show that, as expected, the l_2 error of the compact schemes is much smaller, for fixed parameters, than the error for the standard schemes (explicit Euler, semi–implicit Euler, Leap–Frog Du Fort–Frankel). The CPU time is only slightly higher for the compact schemes, for a fixed number of grid nodes (Section 2.5).

Our next goal is to prove the convergence of the numerical solution obtained by the compact scheme R3C to the unique viscosity solution to (2.1)–(1.26). In the literature, unlike for many standard finite difference schemes, there are very few results concerning the convergence of high–order compact finite difference schemes. In [**27**] compact finite difference methods for initial–boundary–value problems for mixed systems of strongly parabolic and strictly hyperbolic equations are studied. Assuming the existence of a smooth solution, a pilot function [**119, 137**] is constructed which leads to convergence results. Il'in [**81**] studies compact finite difference schemes for *linear* convection–diffusion equations and gives error estimates. Wang and Liu [**141**] propose a fourth–order scheme for the two–dimensional, incompressible Navier–Stokes equations in vorticity formulation and prove its convergence using energy estimates. The convergence of approximation schemes for fully nonlinear second order equations is studied in a general setting in [**12**]. The originality of the results of this chapter consists in the combination of high–order compact finite difference schemes and techniques for viscosity solutions. Since the numerical solution involves an approximation process for (1.24), we prove in Section 2.6.1 an analytical convergence result using the "half–relaxed limits" technique. Then, in Section 2.6.2, we show our main result, the convergence of the compact scheme R3C, based on a result of [**12**].

Finally, we present a numerical example of a European Call option with different transaction cost parameters a. As expected, the option price with positive a is higher than the Black–Scholes price ($a = 0$). The difference between the option price for $a > 0$ and the Black–Scholes price is maximal (in absolute value) near the strike price E, but changes with time. In particular, far from the maturity the difference is maximal at asset prices smaller than the strike price (Section 2.7).

2.2 The transformed problem

In this section we reformulate the problem (2.1), (1.23)–(1.26) using a variable transformation. In [**11**] the existence of a unique continuous viscosity solution V to this problem has been shown.

To overcome a possible degeneration at $S = 0$ and to obtain a forward parabolic problem, we use the variable transformations

$$x(S) = \ln\left(\frac{S}{E}\right), \quad \tilde{t}(t) = \frac{1}{2}\sigma^2(T-t), \quad u = \exp(-x)\frac{V}{E}.$$

Equation (2.1) is hereby transformed into (dropping the tilde)

$$u_t - (1 + \Psi\left[\exp(Kt+x)a^2E(u_{xx}+u_x)\right])(u_{xx}+u_x) - Ku_x = 0, \tag{2.2}$$

with

$$x \in \mathbb{R}, \quad 0 \leq t \leq \tilde{T} = \sigma^2 T/2, \quad K = \frac{2r}{\sigma^2}.$$

The problem is completed by the following initial and boundary conditions:

$$\begin{aligned} u(x,0) &= u_0(x) = \max(1-\exp(-x),0), \\ u(x,t) &= 0\ (x \to -\infty), \\ u(x,t) &\sim 1\ (x \to +\infty). \end{aligned}$$

For the computation we replace $\mathbb{R}$ by $[-R,R]$ with $R > 0$. For simplicity, we consider a uniform grid $Z = \{x_i \in [-R,R] : x_i = ih,\ i = -N,...,N\}$ consisting of $2N+1$ grid points, with $R = Nh$ and with space step h and time step k. Let U_i^n denote the approximate solution to (2.2) in x_i at the time $t_n = nk$ and set $U^n = (U_i^n)_{i=1}^{2N+1}$ and $U = (U^n)_{n=1}^{M}$. The boundary conditions on the limited grid are treated as follows.

$$u(x,0) = \max(1-\exp(-x),0), \tag{2.3}$$
$$u(-R,t) = 0, \tag{2.4}$$
$$u(R,t) = 1-\exp(-R-Kt). \tag{2.5}$$

The latter condition corresponds to the asymptotic value of the exact solution to the equation for $a = 0$. More precisely, the solution of (2.2) satisfies (see (1.26))

$$u(x,t) \sim 1-\exp(-x-Kt) \qquad \text{as } x \to \infty.$$

Approximately, we expect to have $u(R,t) \approx 1 - \exp(-x_N - Kt)$ for sufficiently large $R > 0$. The nonlinear correction of the volatility in (2.1) is a function of the second derivative, so we assume that the influence of the nonlinearity at the boundary can be neglected for large R. For a justification of this assumption and a detailed investigation of this matter we refer to [**146**]. The error caused by boundary conditions imposed on an artificial boundary for a class of Black–Scholes equations has been studied rigorously in [**89**]. Note that the use of a Neumann boundary condition will give similar results, since the error inflicted by the boundary data is restricted to a boundary layer [**7**].

2.3 Finite difference schemes

In this section we recall standard finite difference schemes and the compact schemes derived by Rigal [**128**]. All schemes we consider here use two time levels — except the Leap-Frog Du Fort-Frankel scheme which uses three time levels. In the space variable x we use a compact stencil requiring only three consecutive points in time level $n+1$. The schemes — except the Leap-Frog Du Fort-Frankel scheme — can be written in the following form

$$A^n U^{n+1} = B^n U^n, \tag{2.6}$$

with the discretization matrices A^n and B^n

$$A^n = [a_{-1}, a_0, a_1], \quad B^n = [b_{-2}, b_{-1}, b_0, b_1, b_2].$$

The matrix A^n is tridiagonal, therefore the resulting linear systems can be solved very efficiently linearly in time using a special form of the Gaussian elimination known as the Thomas algorithm [**34**]. We suppose that

$$\sum_{i=-1}^{1} a_i = \sum_{i=-2}^{2} b_i = 1,$$

which is satisfied by any consistent scheme after normalization of the coefficients.

The nonlinearity is introduced explicitly in all the schemes. In the following let s_i^n denote the explicitly discretized nonlinear volatility correction

$$s_i^n = \Psi\left[\exp(Knk + x_i) a^2 E \left(\frac{U_{i-2}^n - 2U_i^n + U_{i+2}^n}{4h^2} + \frac{U_{i+1}^n - U_{i-1}^n}{2h}\right)\right]. \tag{2.7}$$

We use this form of the discretization of the second derivative because it gives a smoother approximation of u_{xx}. With a standard central difference the schemes become instable for small values of h unless k is very small. The problem lies in the initial condition u_0 since it is not differentiable at $x = 0$. We use spline interpolation of high order to carefully smooth the initial data, but only in combination with the five–point approximation (2.7) we obtain acceptable results.

We use a Dormand–Prince–4–5 Runge–Kutta scheme to solve the ordinary differential equation (1.24) and a cubic spline interpolation to obtain the values of Ψ for arbitrary arguments (cf. Figure 2.1).

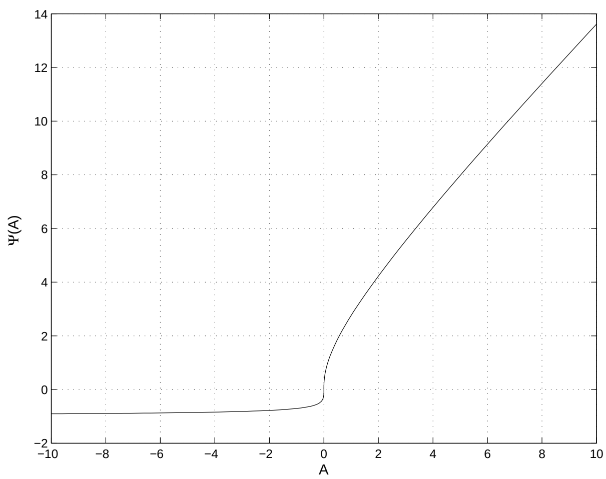

Figure 2.1: Solution of the ODE (1.24).

In the following let

$$\lambda = -(1+K), \quad \alpha = \frac{\lambda h}{2}, \quad r = \frac{k}{h^2}, \quad \mu = \frac{k}{h} \tag{2.8}$$

denote the linear part of the coefficient of the convection term in (2.2), the so–called cell Reynolds number, the parabolic mesh ratio and the hyperbolic mesh ratio, respectively. We say a scheme is of order (m, n) if it is formally consistent of order m in time and of order n in space or, more precisely, the truncation error is of order $\mathcal{O}(k^m + h^n)$.

2.3.1 Classical schemes

In the following we recall some classical finite difference schemes and their properties corresponding to the linear case, i.e. $a = 0$ or $s_i^n = 0$. We verify these properties for the nonlinear case $a > 0$ by the numerical studies in Section 2.5.

Explicit scheme (FTCS)

The Forward–Time Central–Space explicit scheme (FTCS) is given by

$$\begin{aligned}
&a_{-1} = 0 && a_0 = 1, && a_1 = 0,\\
&b_{-1} = r - \frac{\mu}{2}(s_i^n - \lambda), && b_0 = 1 - 2r - \frac{r}{2}s_i^n, && b_1 = r + \frac{\mu}{2}(s_i^n - \lambda),\\
&&& b_{-2} = b_2 = \frac{r}{4}s_i^n,
\end{aligned}$$

It is of order (1,2), with a very restrictive stability condition. In the linear case it reads [**138**]

$$r \leq \frac{1}{2}. \tag{2.9}$$

To avoid oscillations, the following condition must be satisfied

$$|\alpha| \leq 1. \tag{2.10}$$

Semi-implicit scheme (BTCS)

The semi–implicit scheme, using Backward–Time Central–Space differencing (BTCS) for the linear part and explicit treatment of the nonlinearity, is given by

$$\begin{aligned}
&a_{-1} = \frac{\lambda}{2}\mu - r, \quad a_0 = 1 + 2r, \quad a_1 = -\frac{\lambda}{2}\mu - r,\\
&b_{-2} = \frac{r}{4}s_i^n, \quad b_{-1} = -\frac{1}{2}\mu s_i^n, \quad b_0 = 1 - \frac{r}{2}s_i^n, \quad b_1 = \frac{1}{2}\mu s_i^n, \quad b_2 = \frac{r}{4}s_i^n.
\end{aligned}$$

It is unconditionally stable and of order (1,2). It is non-oscillatory if (2.10) is satisfied [**138**].

Crank–Nicolson (CN)

The Crank–Nicolson scheme with explicit treatment of the nonlinearity is given by

$$\begin{aligned} a_{-1} &= \left(-\frac{r}{2}+\frac{\mu}{4}\right) s_i^n - \frac{r}{2} - \frac{\lambda}{4}\mu, & b_{-1} &= \left(\frac{r}{2}-\frac{\mu}{4}\right) s_i^n + \frac{r}{2} + \frac{\lambda}{4}\mu, \\ a_0 &= 1 + r(1+s_i^n), & b_0 &= 1 - r(1+s_i^n), \\ a_1 &= \left(-\frac{r}{2}-\frac{\mu}{4}\right) s_i^n - \frac{r}{2} + \frac{\lambda}{4}\mu, & b_1 &= \left(\frac{r}{2}+\frac{\mu}{4}\right) s_i^n + \frac{r}{2} - \frac{\lambda}{4}\mu \end{aligned}$$

and $b_{-2} = b_2 = 0$. It is unconditionally stable and of order (2,2).

Leap–Frog Du Fort–Frankel scheme (LFDF)

The Leap–Frog Du Fort–Frankel scheme is an explicit three–time–level scheme. It is given by

$$\begin{aligned} \frac{U_i^{n+1} - U_i^{n-1}}{2k} &= \frac{U_{i-1}^n - (U_i^{n-1} + U_i^{n+1}) + U_{i+1}^n}{h^2} + (1+K)\frac{U_{i+1}^n - U_{i-1}^n}{2h} \\ &\quad + s_i^n \left(\frac{U_{i-2}^n - 2U_i^n + U_{i+2}^n}{4h^2} + \frac{U_{i+1}^n - U_{i-1}^n}{2h} \right). \end{aligned}$$

Due to the nature of three–time–level schemes we need an additional method to compute the numerical solution at the first time step. In our numerical test in Section 2.5 the compact scheme R3B (see below) is used for this purpose.

The scheme is stable in the linear case [**129**] if

$$r < \frac{1}{2\,|\alpha|} \tag{2.11}$$

and it is of order (2,2). It is non–oscillatory if condition (2.10) is valid [**129**].

2.3.2 Compact schemes of higher order

The following two schemes were introduced by Rigal [**128**] for linear convection–diffusion problems. We apply them to problem (2.2). They are both compact two–level schemes of order (2,4) in the linear case. The nonlinearity is treated semi–implicitly as in the previous subsections.

R3A scheme

The scheme R3A is given by

$$
\begin{aligned}
a_{-1} &= \left(\frac{1}{12} - \frac{r}{2}\right)(1+\alpha) - \frac{\alpha^2 r}{6} + \frac{\alpha^2 r^2}{3}, \\
a_0 &= \frac{5}{6} + r + \frac{\alpha^2 r}{3} - \frac{2\alpha^2 r^2}{3}, \\
a_1 &= \left(\frac{1}{12} - \frac{r}{2}\right)(1-\alpha) - \frac{\alpha^2 r}{6} + \frac{\alpha^2 r^2}{3}, \\
b_{-2} &= \frac{r}{4} s_i^n, \\
b_{-1} &= \left(\frac{1}{12} + \frac{r}{2}\right)(1+\alpha) + \frac{\alpha^2 r}{6} + \frac{\alpha^2 r^2}{3} - \frac{1}{2}\mu s_i^n, \\
b_0 &= \frac{5}{6} - r - \frac{\alpha^2 r}{3} - \frac{2\alpha^2 r^2}{3} - \frac{r}{2} s_i^n, \\
b_1 &= \left(\frac{1}{12} + \frac{r}{2}\right)(1-\alpha) - \frac{\alpha^2 r}{6} + \frac{\alpha^2 r^2}{3} + \frac{1}{2}\mu s_i^n, \\
b_2 &= \frac{r}{4} s_i^n.
\end{aligned}
$$

This scheme is stable in the linear case $s_i^n = 0$ [**128**] if

$$
r \leq \frac{1}{\sqrt{2}\,|\alpha|}. \tag{2.12}
$$

It is non–oscillatory for arbitrary values of α.

R3B scheme

The scheme R3B is given by

$$
\begin{aligned}
a_{-1} &= \left(\frac{1}{12} - \frac{r}{2}\right)(1+\alpha) - \frac{\alpha^2 r}{6} + \frac{\alpha^3 r^2}{3} - \frac{2\alpha^4 r^3}{3}, \\
a_0 &= \frac{5}{6} + r + \frac{\alpha^2 r}{3} + \frac{4\alpha^4 r^3}{3}, \\
a_1 &= \left(\frac{1}{12} - \frac{r}{2}\right)(1-\alpha) - \frac{\alpha^2 r}{6} - \frac{\alpha^3 r^2}{3} - \frac{2\alpha^4 r^3}{3}, \\
b_{-2} &= \frac{r}{4} s_i^n,
\end{aligned}
$$

$$
\begin{aligned}
b_{-1} &= \left(\frac{1}{12} + \frac{r}{2}\right)(1+\alpha) + \frac{\alpha^2 r}{6} + \frac{\alpha^3 r^2}{3} + \frac{2\alpha^4 r^3}{3} - \left(\frac{r}{4} + \frac{1}{2}\mu\right) s_i^n, \\
b_0 &= \frac{5}{6} - r - \frac{\alpha^2 r}{3} - \frac{4\alpha^4 r^3}{3} - 2 r s_i^n, \\
b_1 &= \left(\frac{1}{12} - \frac{r}{2}\right)(1+\alpha) + \frac{\alpha^2 r}{6} - \frac{\alpha^3 r^2}{3} + \frac{2\alpha^4 r^3}{3} - \left(\frac{r}{4} - \frac{1}{2}\mu\right) s_i^n, \\
b_2 &= \frac{r}{4} s_i^n.
\end{aligned}
$$

It is unconditionally stable and non–oscillatory in the linear case $s_i^n = 0$ [**128**].

2.4 R3C scheme

In the nonlinear case $a > 0$ it seems to be quite difficult to prove the stability of the schemes presented above. The reason lies in the fact that, using *five space points* at the present time level, the study of the amplification factor involves certain cubic polynomials. We would need to study their positivity properties for all possible values of the nonlinear coefficients in (2.2). It turns out that it is not possible to show positivity of these polynomials for arbitrary coefficients.

Therefore we construct now a (semi–implicit) two–level *three–point* compact scheme which is of high order and can be proved to be stable. We expect that these theoretical benefits will also make the scheme superior in numerical tests. We will use the modified equation technique [**142**] to construct the scheme.

2.4.1 Construction of the scheme

To obtain an efficient scheme it is important to approximate the nonlinear coefficients in (2.2) explicitly, i.e. at the time level n. Otherwise one would need to perform a nonlinear iteration in each time step which is quite time–consuming.

With

$$
\begin{aligned}
&\beta = 1 + \Psi\left[\exp(Kt + x) a^2 E(\Delta_2 U^n + \Delta_0 U^n)\right], \\
&\lambda = -\beta - K,
\end{aligned}
$$

where

$$\Delta_0 U_j^n = \frac{U_{j+1}^n - U_{j-1}^n}{2h}, \quad \Delta_2 U_j^n = \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{h^2},$$

the "semi–discretized" equation (2.2) takes the following form

$$u_t = \beta u_{xx} - \lambda u_x. \tag{2.13}$$

We will now study equation (2.13) for arbitrary values $\beta > 0$. We define a general two–level three–point scheme

$$\begin{aligned} D_t U_j^n &= \beta(\tfrac{1}{2} + A_1)\Delta_2 U_j^n + \beta(\tfrac{1}{2} + A_2)\Delta_2 U_j^{n+1} - \lambda(\tfrac{1}{2} + B_1)\Delta_0 U_j^n \\ &\quad - \lambda(\tfrac{1}{2} + B_2)\Delta_0 U_j^{n+1}, \end{aligned} \tag{2.14}$$

where

$$D_t U_j^n = \frac{U_j^{n+1} - U_j^n}{k},$$

and A_i, B_i are real constants which must be chosen in such a way that the lower order terms in the truncation error are eliminated and that the scheme is stable and non–oscillatory.

With this explicit discretization of the nonlinear coefficients we study the local stability ('frozen coefficients') of the linearized equations. It is well–known [**127**] that for linear problems with variable coefficients (not in general, but for important classes of equations, namely parabolic and symmetric hyperbolic equations) local stability is necessary for overall stability and slightly strengthened local stability is also sufficient to ensure overall stability.

We recall some general results on two–level three–point schemes.

Lemma 2.1. *A two–level three–point finite difference scheme is stable (in the sense of von Neumann) if and only if the coefficients a_i, b_i satisfy*

$$(a_1 - a_{-1})^2 - (b_1 - b_{-1})^2 > a_1 + a_{-1} - b_1 - b_{-1}, \tag{2.15}$$

$$(a_1 + a_{-1})^2 - (b_1 + b_{-1})^2 > a_1 + a_{-1} - b_1 - b_{-1}. \tag{2.16}$$

Lemma 2.2. *A two–level three–point finite difference scheme is non–oscillatory if the coefficients a_i, b_i satisfy*

$$(a_1 - b_1)(a_{-1} - b_{-1}) \geq 0. \tag{2.17}$$

For the proofs we refer to [**128**, Lemmata 1 and 2].

Applying (2.14) to a sufficiently smooth solution of (2.13), we obtain the truncation error $E_u(k,h)$. Differentiating (2.13) we obtain higher order equations. Using them to eliminate the time derivatives, we may write $E_u(k,h)$ in terms of the space derivatives only

$$E_u(k,h) = \sum_{j=1}^{4} e_j \partial_x^j u + \text{higher order derivatives},$$

with

$$\begin{aligned}
e_1 &= \lambda(B_2 + B_1), && (2.18)\\
e_2 &= -\beta(A_2 + A_1) - k\lambda^2 B_2, && (2.19)\\
e_3 &= \tfrac{1}{12}\lambda(2h^2 + 2h^2 B_2 + 6\lambda^2 k^2 B_2 + k^2\lambda^2 \\
&\quad + 12k\beta A_2 + 2h^2 B_1 + 12k\beta B_2), && (2.20)\\
e_4 &= -\tfrac{1}{24}(1 + 4B_2)\lambda^4 k^3 - \tfrac{1}{4}(2A_2 + 4B_2 + 1)\beta\lambda^2 k^2 \\
&\quad - \tfrac{1}{12}(\lambda^2 h^2 + 12\beta^2 A_2 + 2\lambda^2 h^2 B_2)k - \tfrac{1}{12}(1 + A_1 + A_2)\beta h^2. && (2.21)
\end{aligned}$$

To obtain a scheme of order (2,4), A_i, B_i must be chosen such that these error terms vanish or are of order (2,4).

Solving the linear system consisting of (2.18)–(2.20) in terms of A_1, A_2, B_1 we get a class of schemes depending only on the parameter B_2. This scheme, called (2.14'), is given by (2.14) with

$$\begin{aligned}
B_1 &= -B_2, && (2.22)\\
A_1 &= -\tfrac{1}{12k\beta}(-2h^2 + 6\lambda^2 k^2 B_2 - k^2\lambda^2 - 12k\beta B_2), && (2.23)\\
A_2 &= -\tfrac{1}{12k\beta}(2h^2 + 6\lambda^2 k^2 B_2 + k^2\lambda^2 + 12k\beta B_2). && (2.24)
\end{aligned}$$

Equation (2.22) is a necessary condition to have a consistent scheme. The coefficient B_2 should be chosen in such a way that we obtain a stable and non–oscillatory scheme of order (2,4). Further we require our scheme to be *forward diffusive* (in relation with the parabolic problem being well–posed), i.e.

$$1 + A_1 + A_2 > 0. \qquad (2.25)$$

We now study the properties of the scheme (2.14'). We obtain the following result.

Theorem 2.3. *Scheme (2.14') is stable (in the sense of von Neumann) if and only if the coefficient B_2 and r, α, β satisfy*

$$(-\beta + 4r\alpha^2 B_2)(-1 + 4r^2\alpha^2 + 12\beta r B_2) > 0. \tag{2.26}$$

It is non–oscillatory if B_2 and r, α, β satisfy

$$(-\beta + 4r\alpha^2 B_2 + \alpha)(-\beta + 4r\alpha^2 B_2 - \alpha) \geq 0. \tag{2.27}$$

It is forward diffusive if and only if B_2, r, α satisfy

$$1 - 4r\alpha^2 B_2 > 0. \tag{2.28}$$

Proof. To prove stability we need to verify conditions (2.15) and (2.16). This is shown by straightforward computation. The scheme (2.14) can be written in the form (2.6) where the coefficients a_i, b_i are given by

$$\begin{aligned} a_{-1} &= -\beta(\tfrac{r}{2} + rA_2) - \tfrac{\mu}{4} - \mu\tfrac{B_2}{2}, & b_{-1} &= \beta(\tfrac{r}{2} + rA_1) + \tfrac{\mu}{4} + \mu\tfrac{B_1}{2}, \\ a_0 &= 1 + \beta(r + 2rA_2), & b_0 &= 1 - \beta(r + 2rA_1), \\ a_1 &= -\beta(\tfrac{r}{2} + rA_2) + \tfrac{\mu}{4} + \mu\tfrac{B_2}{2}, & b_1 &= \beta(\tfrac{r}{2} + rA_1) - \tfrac{\mu}{4} - \mu\tfrac{B_1}{2} \end{aligned} \tag{2.29}$$

and $b_{-2} = b_2 = 0$. With these coefficients (2.15) and (2.16) are equivalent to

$$\begin{aligned} (B_2 + B_2^2 - B_1 - B_1^2)\mu^2 + (2\beta + 2\beta A_2 + 2\beta A_1)r > 0, \\ -2r\beta(1 + A_1 + A_2)(2\beta r A_1 - 2\beta r A_2 - 1) > 0. \end{aligned}$$

Using (2.22)–(2.24) these inequalities simplify to

$$2\beta r > 0, \tag{2.30}$$

$$\frac{2}{3}r(-\beta + 4r\alpha^2 B_2)(-1 + 4r^2\alpha^2 + 12\beta r B_2) > 0, \tag{2.31}$$

respectively. Condition (2.30) is always satisfied and (2.31) yields (2.26).

For non–oscillation we have to check condition (2.17). Elementary computations using the above coefficients and substituting B_1, A_1, A_2 from (2.22)–(2.24) give the condition

$$\frac{1}{4}(-2r\beta + 2rk\lambda^2 B_2 + \mu)(-2r\beta + 2rk\lambda^2 B_2 - \mu) \geq 0.$$

Writing this condition in terms of α and r gives (2.27).

Substituting (2.23) and (2.24) into (2.25) we obtain

$$1 - k\lambda^2 B_2 > 0,$$

which is equivalent to (2.28). $\square$

We will now propose a choice of the coefficient B_2 and study the properties of the scheme obtained by this choice. By construction, $e_1 = e_2 = e_3 = 0$ for the scheme (2.14'). The error e_4 can be written as

$$e_4 = \frac{1}{12}k(-\lambda^2 h^2 + k^2\lambda^4 + 12\beta^2)B_2 - \frac{1}{12}\beta(-h^2 + 2k^2\lambda^2). \tag{2.32}$$

We must choose B_2 in such a way that e_4 is of order (2,4). The lower order part of e_4 is

$$k\beta^2 B_2 + \frac{1}{12}\beta h^2.$$

Therefore we make the ansatz

$$B_2 = -\frac{1}{12}\frac{h^2}{\beta k} + b. \tag{2.33}$$

We have to choose the constant b of order $\mathcal{O}(h^4)$ to obtain a truncation error e_4 of the same order. The obvious choice $b = 0$ is not recommended since in the linear case $\beta = 1$ or $a = 0$, this choice leads to the R3A scheme which is not unconditionally stable. We want to choose b in such a way that the conditions (2.26)–(2.28) of Theorem 2.3 are satisfied.

We define the R3C scheme by choosing

$$b = -\frac{r\alpha^2}{3\beta} = -\frac{\lambda^2 k}{12\beta},$$

which yields

$$B_2 = -\frac{1 + 4r^2\alpha^2}{12\beta r}. \tag{2.34}$$

For the (linear) case $\beta = 1$ this choice corresponds to the R3B scheme of Rigal [**128**]. The coefficients a_i, b_i are given by

$$\begin{aligned}
a_{-1} &= -\frac{12r\beta^2 - 2\beta + r\lambda^2h^2 + r^3\lambda^4h^4 + 6r\lambda h\beta - \lambda h - r^2\lambda^3h^3}{24\beta},\\
a_0 &= \frac{10\beta + 12r\beta^2 + r\lambda^2h^2 + r^3\lambda^4h^4}{12\beta},\\
a_1 &= -\frac{12r\beta^2 - 2\beta + r\lambda^2h^2 + r^3\lambda^4h^4 - 6r\lambda h\beta + \lambda h + r^2\lambda^3h^3}{24\beta},\\
b_{-1} &= \frac{12r\beta^2 + 2\beta + r\lambda^2h^2 + r^3\lambda^4h^4 + 6r\lambda h\beta + \lambda h + r^2\lambda^3h^3}{24\beta},
\end{aligned} \tag{2.35}$$

$$b_0 = -\frac{-10\beta + 12r\beta^2 + r\lambda^2h^2 + r^3\lambda^4h^4}{12\beta},$$
$$b_1 = \frac{12r\beta^2 + 2\beta + r\lambda^2h^2 + r^3\lambda^4h^4 - 6r\lambda h\beta - \lambda h - r^2\lambda^3h^3}{24\beta}.$$

Remark 2.4. In the linear case $\beta = 1$ or $a = 0$ the R3C scheme coincides with the R3B scheme.

2.4.2 Properties of the scheme

Theorem 2.5. *The R3C scheme defined above is an unconditionally stable (in the sense of von Neumann), non–oscillatory and forward diffusive scheme of order (2,4). Its truncation error is given by*

$$e_4 = -\frac{\lambda^2(k^4\lambda^4 - h^4 + 36k^2\beta^2)}{144\beta}. \tag{2.36}$$

Proof. Substituting (2.34) in (2.32) we get (2.36), i.e. the scheme is of order (2,4), since $e_1 = e_2 = e_3 = 0$.

Taking into account (2.34), conditions (2.28) and (2.26) are equivalent to

$$1 + \frac{\alpha^2(1 + 4\alpha^2r^2)}{3\beta} > 0,$$
$$\frac{4}{3}\Big(\beta + \frac{\alpha^2(1 + 4\alpha^2r^2)}{3\beta}\Big) > 0,$$

respectively. These conditions hold for all values of β.

Using (2.34) in (2.27), we find that this condition is valid if for all values of β

$$\frac{1}{9}\frac{r^2}{\beta^2}(3\beta^2 + \alpha^2 + 4\alpha^4r^2 + 3\beta\alpha)(3\beta^2 + \alpha^2 + 4\alpha^4r^2 - 3\beta\alpha) \geq 0 \tag{2.37}$$

or, equivalently, if for all values of β

$$3\beta^2 - 3\alpha\beta + \alpha^2 + 4\alpha^4r^2 \geq 0. \tag{2.38}$$

This is a quadratic polynomial in β. Its leading coefficient is positive and its discriminant is equal to $-3\alpha^2 - 48\alpha^4r^2$ which is negative for all values of α and r. Hence there are no roots and (2.38) is true for any value of β. $\square$

In the following we prove some properties of the discretization matrices needed in the convergence proof in Section 2.6.2. To simplify the presentation, we only consider the case $K = 0$ which corresponds to zero interest rate. Notice that this implies $\lambda = -\beta$. In the general case, similar conditions as in Lemma 2.6 below can be obtained, with bounds depending on h, β, and λ.

Lemma 2.6. *If $h < 2$ and*

$$\frac{1}{6\beta} \leq r < \frac{1}{2\beta}, \tag{2.39}$$

then B^n is a positive matrix (i.e. all elements are positive) and A^n is an M–matrix. More specifically, a_0, b_{-1}, b_0, b_1 are positive, a_{-1}, a_1 are negative or zero, A^n is non–singular, and $(A^n)^{-1}$ is a positive matrix.

Proof. The coefficients a_0, b_1 are always positive. It follows from (2.35) that b_0 and b_{-1} are positive if

$$10\beta - 12r\beta^2 - r\beta^2h^2 - \beta^4h^4r^3 > 0, \tag{2.40}$$
$$12r\beta + 2 + r\beta h^2 + \beta^3h^4r^3 - 6\beta hr - h - \beta^2h^3r^2 > 0, \tag{2.41}$$

respectively, and a_{-1}, a_1 are negative or zero if

$$12r\beta^2 - 2\beta + r\beta^2h^2 + \beta^4h^4r^3 - 6\beta^2hr + \beta h + \beta^3h^3r^2 \geq 0, \tag{2.42}$$
$$12r\beta^2 - 2\beta + r\beta^2h^2 + \beta^4h^4r^3 + 6\beta^2hr - \beta h - \beta^3h^3r^2 \geq 0, \tag{2.43}$$

respectively.

First, we study (2.41). Consider the polynomial $p(\beta) = h^4r^3\beta^2 - h^3r^2\beta + rh^2 - 6hr + 12r$. It is positive for all $h \neq 4$, since its leading coefficient is positive and its discriminant is $-3r^4h^4(h-4)^2$, which is negative for $h \neq 4$. Hence, $p(\beta)\beta + 2 - h > 0$ and thus (2.41) follows if $h < 2$.

We solve the equations related to (2.40), (2.42), (2.43) for r, being cubic polynomials in r. For each equation we obtain one real root and two complex roots. From the real root of the first equation we obtain the condition $r < c_0(h)/\beta$ with

$$c_0(h) = \frac{1}{3}\frac{x^{2/3} - 36 - 3h^2}{h^2x^{1/3}},$$

where $x = 135h^2 + 3\sqrt{3}\sqrt{1728 + 432h^2 + 711h^4 + h^6}$. The function c_0 is decreasing in h with $\min_{h\in[0,2]} c_0(h) = \frac{1}{2}$, which gives the upper bound in (2.39).

The real roots of the other equations result in the condition $\max(c_1(h), c_2(h))/\beta \leq r$ with

$$c_1(h) = \frac{1}{3}\frac{z^{1/3}}{h^2} - \frac{2}{3}\frac{18 + h^2 - 9h}{h^2 z^{1/3}} - \frac{1}{3h},$$
$$c_2(h) = \frac{1}{3}\frac{y^{1/3}}{h^2} - \frac{2}{3}\frac{18 + h^2 + 9h}{h^2 y^{1/3}} + \frac{1}{3h},$$

where

$$y = -54h + 10h^3 + 6\sqrt{3}\sqrt{432 + 423h^2 + 648h + 12h^4 + 126h^3 + h^6 + 2h^5},$$
$$z = 54h - 10h^3 + 6\sqrt{3}\sqrt{432 + 423h^2 - 648h + 12h^4 - 126h^3 + h^6 - 2h^5}.$$

It can be seen that the functions c_1 and c_2 both attain their maximum at $h = 0$ with $c_1(0) = c_2(0) = \frac{1}{6}$. This yields the lower bound in (2.39). Therefore, B^n is a positive matrix and A^n is an L–matrix if (2.39) holds. Since $a_0 > |a_{-1}| + |a_1|$, A^n is strictly diagonally dominant. Hence, A^n is an M–matrix which yields the claim. $\square$

In Figure 2.2 the set $\max\bigl(c_1(h), c_2(h)\bigr)/\beta \leq r \leq c_0(h)/\beta$ is shown. As a by–product of Lemma 2.6, we obtain the following corollary, which ensures the positivity of the numerical solutions.

Corollary 2.7. *Let the assumptions of Lemma 2.6 hold. Then the linear, constant coefficient R3C scheme is positive, i.e. for all $n \in \mathbb{N}$*

$$U^n \geq 0 \implies U^{n+1} \geq 0,$$

where the inequality holds for all components of the vectors.

Remark 2.8. A finite difference scheme of the form (2.6) is called positive if $(A^n)^{-1}B^n$ is a positive matrix. Unlike for many second–order schemes, the matrices A^n and B^n resulting from fourth–order schemes generally do not commute and positivity cannot be easily deduced. The positivity of the scheme holds if both matrices $(A^n)^{-1}$ and B^n are positive.

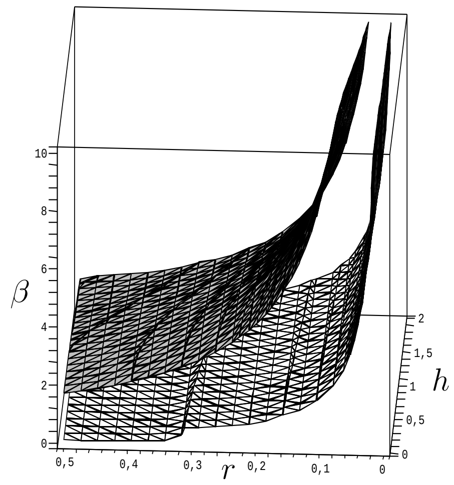

Figure 2.2: The two surfaces represent the equations $r = c_0(h)/\beta$ and $r = \max\bigl(c_1(h), c_2(h)\bigr)/\beta$.

Remark 2.9. The conditions of Lemma 2.6 are sufficient but not necessary. Frequently, such conditions are too restrictive in practice and the scheme will preserve the positivity for a larger set of discretization parameters [**128**]. We observed this also in our numerical experiments presented in the following section.

2.5 Numerical study

2.5.1 Numerical stability

To study the stability of the schemes numerically, we compute the l_2–error ε_2 of the numerical solutions for different values of α and r, where

$$\varepsilon_2 = \left(h \sum_{i=-N}^{+N} \left| U_i^{k_T} - u(x_i, T) \right|^2 \right)^{\frac{1}{2}}$$

and $T = k_T k$. In the linear case $a = 0$ we use the exact solution and in the nonlinear case $a > 0$ a solution on a very fine grid (with $N = 800$) as reference solution. The computations were done using the following parameters

$$\sigma_0 = 0.45, \quad \rho = 0.1, \quad E = 100, \quad T = 0.50625.$$

In Figure 2.3 the error ε_2 is plotted in the α–r–plane (see (2.8)) for each classical scheme and in Figure 2.4 for the compact schemes for the linear case ($a = 0$) and the nonlinear case ($a = 0.02$). Different scales were used for the error of classic and compact schemes. We notice that the schemes' behavior is similar in both cases.

- FTCS: The conditions (2.9) and (2.10) for the stability and non–oscillation can be found again numerically. For large values of α and r, the scheme is unstable and oscillations occur. The area in which the scheme produces acceptable results is very small. In the nonlinear case the stability area is even smaller.
- BTCS: The l_2–error of this scheme is large for larger values of α, giving unsatisfactory results for this region (oscillations).
- LFDF: For large values of α the error grows rapidly. The errors are slightly smaller than those of the BTCS scheme.
- CN: The error in the linear case is small compared to the other classical schemes. In the nonlinear case the error grows fast for large values of α.
- R3A: The error of this scheme shows its good properties. No oscillations occur, the stability region is very large as predicted by (2.12).
- R3B: We observe that scheme R3B gives a slightly better behavior than R3A. There are no oscillations and the scheme is unconditionally stable.
- R3C: As predicted by our theoretical results the scheme is unconditionally stable and non–oscillatory. In the nonlinear case ($a > 0$) the error is even smaller than that of R3A/R3B. In the linear case ($a = 0$) the result is identical to that of R3B (cf. Remark 2.4).

Comparing the computational results of the different schemes we can make the following observations. In the small region of the α–r–plane, where the FTCS scheme is stable, the error of all classical schemes is about the

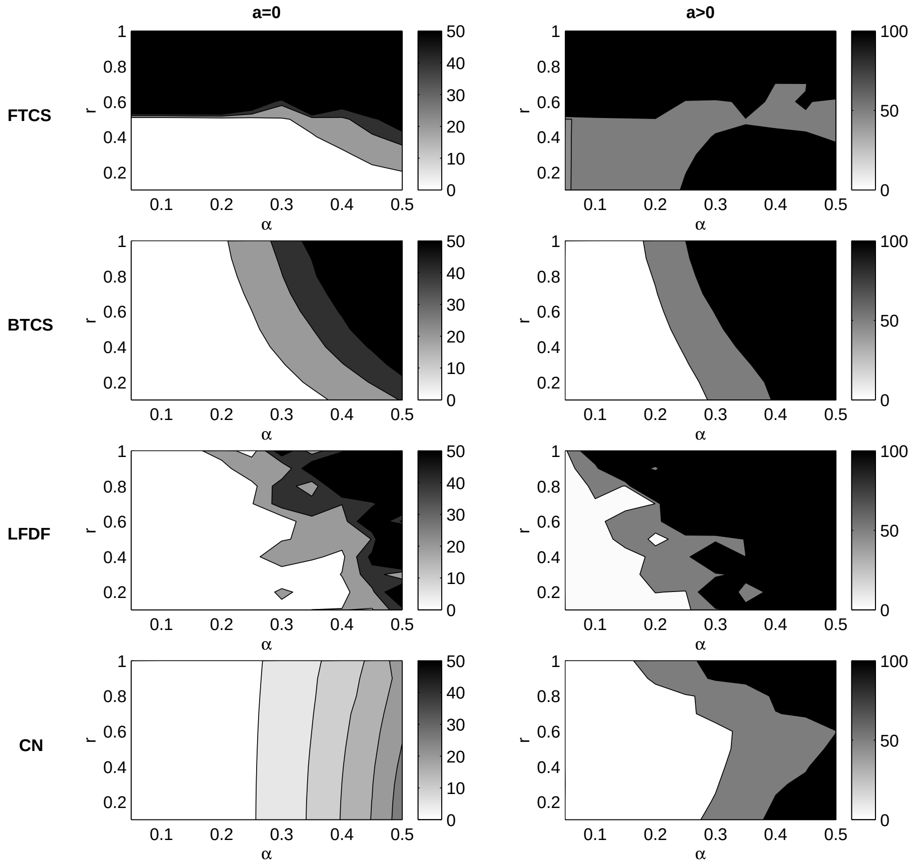

Figure 2.3: Classical schemes: l_2–error in the α–r–plane.

same. The CN scheme gives the best results of all the classical schemes. Comparing the classical (FTCS, BTCS, LFDF, CN) to the high order compact schemes, we notice the superiority of the compact schemes. They are generally significantly more accurate than the classical schemes (due to their higher order); they show no oscillations and their use is not restricted by strong stability conditions. The error difference between R3A and R3B is insignificant whereas R3C provides even better results in the nonlinear case.

2.5.2 Numerical convergence

The truncation error given by expression (2.36) represents the pointwise error in approximating the differential equation (but not necessarily the solu-

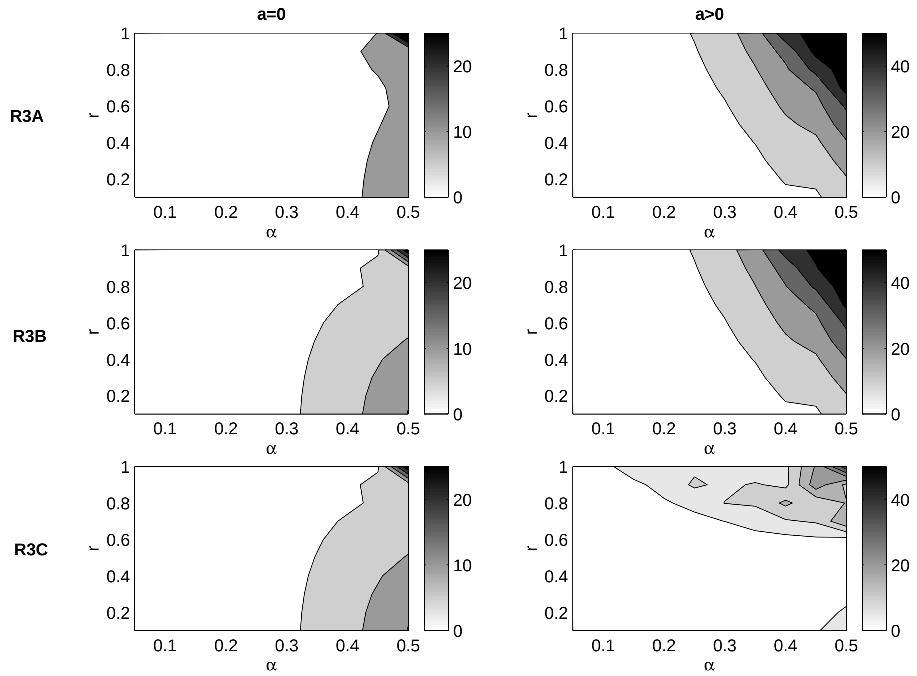

Figure 2.4: Compact schemes: l_2–error in the α–r–plane.

tion) [**138**]. We present in this section a numerical study to compute the order of convergence of the R3C scheme. Asymptotically, we expect the pointwise error to converge as

$$\varepsilon_2 = Ch^m$$

for some m and C representing a constant. This implies

$$\log(\varepsilon_2) = \log(C) + m\log(h).$$

Hence, the double–logarithmic plot ε_2 against h should be asymptotic to a straight line with slope m. This gives a method for experimentally determining the order of accuracy of the method. We refer to Figure 2.5 for the results with the parameters

$$a = 0.02, \quad \sigma_0 = 0.45, \quad \rho = 0.1, \quad E = 100, \quad T = 0.009375.$$

Table 2.1 summarizes the maximal, minimal and average numerical convergence rates. We observe that the numerical convergence rates roughly correspond to the order of the schemes.

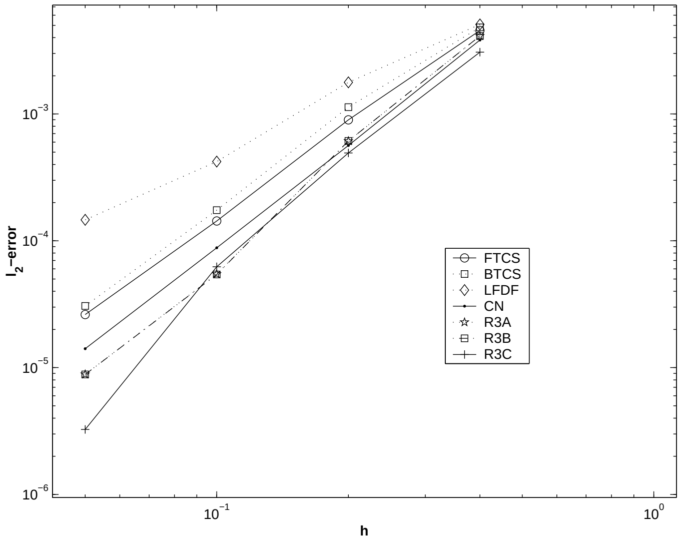

Figure 2.5: Numerical Convergence: l_2–error vs. h.

	$m_{\max}$	$m_{\min}$	m_{av}
FTCS	2.65	2.34	2.48
BTCS	2.70	2.09	2.43
LFDF	2.08	1.52	1.80
CN	2.76	2.64	2.69
R3A	3.41	2.75	3.13
R3B	3.42	2.75	3.13
R3C	4.26	2.98	3.29

Table 2.1: Convergence rates.

2.5.3 Efficiency

An important point in our comparison is the efficiency of the schemes, i.e. the computation time to obtain a given accuracy. Obviously this is machine as well as programming dependent. The different schemes were implemented in an efficient and consistent manner in order not to bias any of them. The computation times recorded include the time for matrix setups, inversions and boundary condition evaluation. All results were computed on the same machine. The number of operations to solve the tridiagonal systems with the Thomas Algorithm is of order $\mathcal{O}(N)$ (see Section 2.3 for the definition of

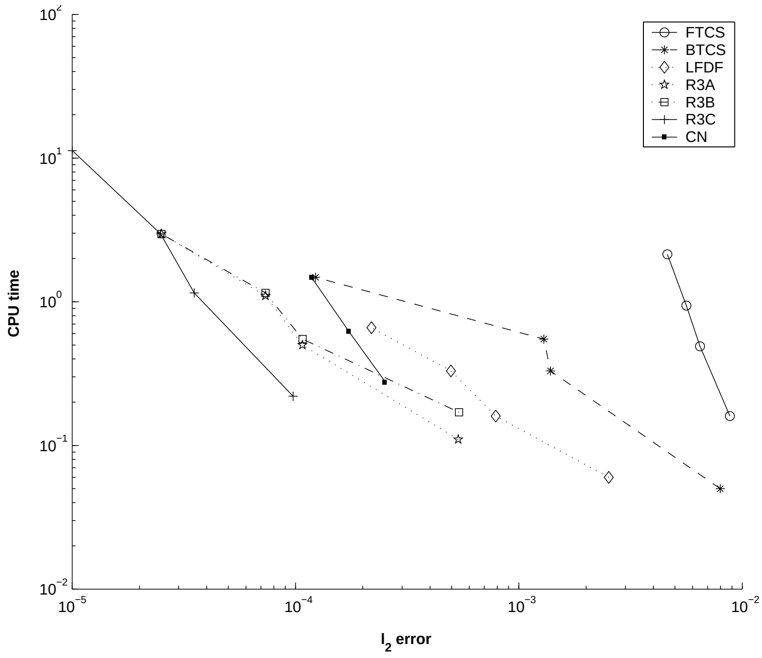

Figure 2.6: Efficiency: CPU–time vs. l_2–error.

N). Hence the dominant factor in the running time is the matrix setup, not the inversion.

We computed solutions on grids with $N = 10, 20, 30, 40$. In Figure 2.6 we plotted the relative l_2–error versus the CPU time for the different grids and schemes. We see that for fixed error the compact schemes take less CPU time than the classic schemes. Due to the strong stability condition of the FTCS scheme, it is generally very time consuming. For fixed time the error of the compact schemes is always significantly smaller. The three compact schemes are the most efficient ones where the R3C scheme seems to be superior to the R3A and R3B schemes.

Using the same values of α and r, the implicit schemes' (BTCS, CN, R3A, R3B, R3C) computation time is not much larger than that of the explicit schemes (FTCS, LFDF), but the accuracy of the compact schemes is significantly better.

2.6 Convergence results

For the convenience of the reader, we briefly recall the notion of viscosity solutions, introduced by Crandall and Lions [**39**]. For a general presentation on viscosity solutions we refer to [**38**]. Following the notation of [**12**], we can write (2.2) as

$$G(x,t,u(x,t),u_t(x,t),u_x(x,t),u_{xx}(x,t)) = 0 \quad \text{in } \bar{\Omega}\times[0,T], \tag{2.44}$$

where G is given by

$$G(x,t,u(x,t),u_t(x,t),u_x(x,t),u_{xx}(x,t))$$
$$=\begin{cases} u_t - (1+\Psi\left[\exp(Kt+x)a^2E(u_{xx}+u_x)\right])(u_{xx}+u_x) - Ku_x & \text{in } Q_t,\\ u(x,0) - \max(1-\exp(-x),0) & \text{in } \Omega,\\ u(-R,t) & \text{in } (0,T),\\ u(R,t) - (1-\exp(-R-Kt)) & \text{in } (0,T), \end{cases}$$

where $Q_t = \Omega\times(0,T)$. Although we have assumed $K=0$ in the previous section, the results of this section hold for any $K\geq 0$ provided that the conclusion of Lemma 2.6 holds. In the following, let z^* and z_* denote the upper semi–continuous and lower semi–continuous envelope of the function $z: C\to\mathbb{R}$, where C is a closed subset of $\mathbb{R}$, defined by

$$z^*(x) = \limsup_{y\to x,\ y\in C} z(y), \qquad z_*(x) = \liminf_{y\to x,\ y\in C} z(y).$$

Definition 2.10. *A locally bounded function $u:\bar{\Omega}\to\mathbb{R}$ is a viscosity subsolution (respectively supersolution) of* (2.44) *if and only if for all $\varphi\in C^2(\bar{\Omega}\times[0,T])$ and for all maximum (respectively minimum) points (x,t) of $u^*-\varphi$ (respectively $u_*-\varphi$), one has*

$$G_*(x,t,u^*(x,t),\varphi_t(x,t),\varphi_x(x,t),\varphi_{xx}(x,t)) \leq 0$$
$$(\textit{respectively } G^*(x,t,u_*(x,t),\varphi_t(x,t),\varphi_x(x,t),\varphi_{xx}(x,t)) \geq 0).$$

A locally bounded function is a viscosity solution of (2.44) *if it is a viscosity subsolution and a viscosity supersolution.*

2.6.1 Analytical convergence result

The solution of (2.2) involves two approximation processes. One is imposing the Dirichlet boundary conditions (2.4) and (2.5). The existence and uniqueness proof in [**11**] uses the boundary conditions (1.26). It is easy to carry over the existence and uniqueness proof with only small changes, so we omit the proof. The convergence of the solution on a bounded domain to the solution on the half–space has been studied in [**7, 103**] for the *linear* case, i.e. $a = 0$.

The other approximation arises when solving (1.24). Since the right–hand side of (1.24) is unbounded for $A \to 0$, it is necessary to solve the ordinary differential equation approximately with a bounded approximation of the right–hand side. This gives rise to an approximate function Ψ_ε which is used in the numerical solution of (2.2). Thus, we are in fact solving an approximate problem,

$$u_t - \left(1 + \Psi_\varepsilon\left[\exp(Kt + x)a^2 E(u_{xx} + u_x)\right]\right)(u_{xx} + u_x) - K u_x = 0, \quad (2.45)$$

with (2.3)–(2.5). In the following we show that the solution of this approximate problem converges to the solution of the original problem.

Proposition 2.11. *Let Ψ_ε be a monotone smooth approximation of Ψ with bounded derivative such that $\Psi_\varepsilon \to \Psi$ locally uniformly as $\varepsilon \to 0$. Then the viscosity solution u_ε of* (2.45) *and* (2.3)–(2.5) *converges to the viscosity solution u of* (2.2) *and* (2.3)–(2.5) *as $\varepsilon \to 0$.*

Proof. We use the "half–relaxed limits" technique which has been introduced by Barles and Perthame [**9, 10**] and Ishii [**82**]. Let u_ε denote a solution to the approximate problem (2.45), (2.3)–(2.5) with $\varepsilon > 0$. We omit an existence proof which is very similar to the one for the original problem. Since $u_1 \equiv 1$ and $u_2 \equiv 0$ are super– and subsolutions, respectively, comparison arguments show that u_ε is bounded independently of ε. Then

$$\overline{u}(x,t) = \limsup_{\varepsilon\to 0}{}^* u_\varepsilon(x,t) = \limsup_{\varepsilon'\to 0}\{u_\varepsilon(x',t') : \varepsilon \le \varepsilon',\ \|(x,t)-(x',t')\| \le \varepsilon'\},$$
$$\underline{u}(x,t) = \liminf_{\varepsilon\to 0}{}_* u_\varepsilon(x,t) = \liminf_{\varepsilon'\to 0}\{u_\varepsilon(x',t') : \varepsilon \le \varepsilon',\ \|(x,t)-(x',t')\| \le \varepsilon'\},$$

are well–defined. By [**38**, Lemma 6.1] both limits are discontinuous viscosity solutions of (2.2), (2.3)–(2.5), since $\Psi_\varepsilon \to \Psi$ locally uniformly as $\varepsilon \to 0$. The strong comparison result for (2.2) in [**11**, App. B, pp. 395] shows that $\overline{u} = \underline{u} = u$. □

2.6.2 Convergence of the compact scheme R3C

The convergence of approximation schemes for fully nonlinear parabolic equations has been studied in an abstract setting in [**12**]. We want to apply Theorem 2.1 in [**12**] to show the convergence of the compact scheme R3C to the viscosity solution of (2.45), (2.3)–(2.5). We start by recalling the assumptions of Theorem 2.1 in [**12**]. The numerical scheme R3C approximating (2.44) can be written as

$$S(k, h, n, i, U_i^{n+1}, U) = 0, \tag{2.46}$$

where U_i^{n+1} is the desired approximate solution that is computed using elements of U. Roughly speaking, Theorem 2.1 in [**12**] states that any stable, consistent and monotone scheme converges to the solution of (2.44), provided (2.44) satisfies a "strong uniqueness" condition. Therefore the scheme S is expected to have the following properties, at least for some sequence (k, h) converging to zero.

(S1) For all (k, h), there exists a solution U of (2.46) that is bounded independently of (k, h).

(S2) For any smooth function ϕ and for any (x, t) in $\bar{\Omega} \times [0, T]$, it holds

$$\liminf_{(k,h)\to 0,\ (x_i,t_n)\to(x,t),\ \xi\to 0} \frac{S(k, h, n, i, \phi_i^{n+1} + \xi, \phi + \xi)}{\rho(k, h)} \geq G_*(x, t, \phi(x, t), \phi_t(x, t), \phi_x(x, t), \phi_{xx}(x, t)),$$

$$\limsup_{(k,h)\to 0,\ (x_i,t_n)\to(x,t),\ \xi\to 0} \frac{S(k, h, n, i, \phi_i^{n+1} + \xi, \phi + \xi)}{\rho(k, h)} \leq G^*(x, t, \phi(x, t), \phi_t(x, t), \phi_x(x, t), \phi_{xx}(x, t))$$

for some function $\rho(k, h) > 0$ such that $\rho(k, h) \to 0$ as $(k, h) \to 0$.

(S3) If $U \geq V$ (the inequality holds for all components) and $U_i^{n+1} = V_i^{n+1}$, then

$$S(k, h, n, i, U_i^{n+1}, U) \leq S(k, h, n, i, V_i^{n+1}, V)$$

for any $k, h > 0$, $1 \leq n \leq M$, $1 \leq i \leq 2N + 1$ and for all $U, V \in \mathbb{R}^{M(2N+1)}$.

(S4) If the locally bounded upper semi–continuous (lower semi–continuous) function u (v) is a viscosity subsolution (supersolution) of (2.44) then

$$u \leq v \quad \text{in } \bar{\Omega}.$$

Our main result on the convergence of the compact scheme R3C is the following theorem.

Theorem 2.12. *Assume that Ψ' is bounded, the constant transaction cost parameter a is sufficiently small (see below) and the assumptions of Lemma 2.6 are fulfilled. Then the solution U converges to the unique viscosity solution of* (2.45), (2.3)–(2.5) *as $(k, h) \to 0$, uniformly on each compact subset of $\bar{\Omega}$.*

Proof. In order to be able to apply Theorem 2.1 in [**12**], we have to check the assumptions **(S1)**–**(S4)**. The proof of **(S4)** is given in [**11**, App. B, pp. 395].

We show that $\|U^n\|_\infty$ is bounded for arbitrary $n \in \mathbb{N}$ if $\|U^0\|_\infty$ is bounded. For arbitrary $n \in \mathbb{N}$ let $i_0 \in \{-N, \ldots, N\}$ be such that $\|U^{n+1}\|_\infty = |U_{i_0}^{n+1}|$. Employing Lemma 2.6 and using $a_{-1} + a_0 + a_1 = b_{-1} + b_0 + b_1 = 1$, we can estimate

$$\begin{aligned}
\|U^{n+1}\|_\infty = |U_{i_0}^{n+1}| &= a_{-1}|U_{i_0}^{n+1}| + a_0|U_{i_0}^{n+1}| + a_1|U_{i_0}^{n+1}| \\
&\leq a_{-1}|U_{i_0-1}^{n+1}| + a_0|U_{i_0}^{n+1}| + a_1|U_{i_0+1}^{n+1}| \\
&\leq |a_{-1}U_{i_0-1}^{n+1} + a_0U_{i_0}^{n+1} + a_1U_{i_0+1}^{n+1}| \\
&= |b_{-1}U_{i_0-1}^{n} + b_0U_{i_0}^{n} + b_1U_{i_0+1}^{n}| \\
&\leq ||B^nU^n||_\infty \leq ||B^n||_\infty||U^n||_\infty = ||U^n||_\infty,
\end{aligned}$$

where $||B^n||_\infty$ is the row-sum norm of B^n. Thus, $||U^n||_\infty \leq \|U^0\|_\infty$ for $n \in \mathbb{N}$, yielding **(S1)**.

The consistency assumption **(S2)** follows from Theorem 2.5. It remains to show that **(S3)** holds. For simplicity, we will only consider the case $K = 0$. This relates to the case of zero interest rate in the financial model. The case $K > 0$ can be proved analogously. Define $F : \mathbb{R}^4 \to \mathbb{R}$ by

$$\begin{aligned}
&F(\Delta_2U_i^{n+1}, \Delta_0U_i^{n+1}, \Delta_2U_i^n, \Delta_0U_i^n) \\
=&\beta\big[(\tfrac{1}{2} + A_1)\Delta_2U_i^n + (\tfrac{1}{2} + A_2)\Delta_2U_i^{n+1} + (\tfrac{1}{2} + B_1)\Delta_0U_i^n + (\tfrac{1}{2} + B_2)\Delta_0U_i^{n+1}\big].
\end{aligned}$$

Note that $\beta, A_1, A_2, B_1, B_2$ depend on U as well. Using this definition we can write

$$S(k,h,n,i,U_i^{n+1},U) = U_i^{n+1} - U_i^n - kF(\Delta_2 U_i^{n+1}, \Delta_0 U_i^{n+1}, \Delta_2 U_i^n, \Delta_0 U_i^n).$$

Let $U, V \in \mathbb{R}^{M(2N+1)}$ with $U \geq V$ and $U_i^{n+1} = V_i^{n+1}$. We need to show that

$$S(k,h,n,i,V_i^{n+1},V) - S(k,h,n,i,U_i^{n+1},U) \geq 0.$$

With the abbreviations $W_i^{n+1} = U_i^{n+1} - V_i^{n+1}$, $W_i^n = U_i^n - V_i^n$ we infer

$$\begin{aligned} &S(k,h,n,i,V_i^{n+1},V) - S(k,h,n,i,U_i^{n+1},U) \\ &= (U_i^n - V_i^n) + k\big[F(\Delta_2 U_i^{n+1}, \Delta_0 U_i^{n+1}, \Delta_2 U_i^n, \Delta_0 U_i^n) \\ &\quad - F(\Delta_2 V_i^{n+1}, \Delta_0 V_i^{n+1}, \Delta_2 V_i^n, \Delta_0 V_i^n)\big] \\ &= W_i^n + k\nabla F(z)(\Delta_2 W_i^{n+1}, \Delta_0 W_i^{n+1}, \Delta_2 W_i^n, \Delta_0 W_i^n), \end{aligned}$$

using the mean value theorem

$$\begin{aligned} &F(\Delta_2 U_i^{n+1}, \Delta_0 U_i^{n+1}, \Delta_2 U_i^n, \Delta_0 U_i^n) - F(\Delta_2 V_i^{n+1}, \Delta_0 V_i^{n+1}, \Delta_2 V_i^n, \Delta_0 V_i^n) \\ &= \nabla F(z)[(\Delta_2 U_i^{n+1}, \Delta_0 U_i^{n+1}, \Delta_2 U_i^n, \Delta_0 U_i^n) \\ &\quad - (\Delta_2 V_i^{n+1}, \Delta_0 V_i^{n+1}, \Delta_2 V_i^n, \Delta_0 V_i^n)] \end{aligned}$$

for some

$$z \in \big[(\Delta_2 U_i^{n+1}, \Delta_0 U_i^{n+1}, \Delta_2 U_i^n, \Delta_0 U_i^n), (\Delta_2 V_i^{n+1}, \Delta_0 V_i^{n+1}, \Delta_2 V_i^n, \Delta_0 V_i^n)\big],$$

where $z = (z_1, z_2, z_3, z_4)$, and $[p,q]$ denotes the line between $p, q \in \mathbb{R}^4$. We compute

$$\nabla F(z) = \begin{pmatrix} \beta(\frac{1}{2} + A_2) \\ \beta(\frac{1}{2} + B_2) \\ \beta[(\frac{1}{2} + A_1) + c(a^2)A_{1,\beta}\Psi' z_3] + c(a^2)\Psi'(\frac{1}{2} + A_1)z_3 \\ \beta[(\frac{1}{2} + B_1) + c(a^2)B_{1,\beta}\Psi' z_4] + c(a^2)\Psi'(\frac{1}{2} + B_1)z_4 \end{pmatrix},$$

where $\beta, A_1, A_2, B_1, B_2, \Psi'$ and $A_{1,\beta}, B_{1,\beta}$, the derivatives with respect to β, are evaluated in z and $c(a^2)$ is a positive constant depending on a^2 with $c(a^2) \to 0$ as $a \to 0$. We obtain

$$S(k,h,n,i,V_i^{n+1},V) - S(k,h,n,i,U_i^{n+1},U) = W_i^n + k\beta T_1 + c(a^2)kT_2, \tag{2.47}$$

where

$$
\begin{aligned}
T_1 =& \left(\tfrac{1}{2} + A_2\right) \frac{W_{i+1}^{n+1} - 2W_i^{n+1} + W_{i-1}^{n+1}}{h^2} + \left(\tfrac{1}{2} + B_2\right) \frac{W_{i+1}^{n+1} - W_{i-1}^{n+1}}{2h} \\
&+ \left(\tfrac{1}{2} + A_1\right) \frac{W_{i+1}^{n} - 2W_i^{n} + W_{i-1}^{n}}{h^2} + \left(\tfrac{1}{2} + B_1\right) \frac{W_{i+1}^{n} - W_{i-1}^{n}}{2h}
\end{aligned}
$$

and

$$
\begin{aligned}
T_2 =& \left[\beta A_{1,\beta} \Psi' + \Psi' \left(\tfrac{1}{2} + A_1\right)\right] z_3 \frac{W_{i+1}^{n} - 2W_i^{n} + W_{i-1}^{n}}{h^2} \\
&+ \left[\beta B_{1,\beta} \Psi' + \Psi' \left(\tfrac{1}{2} + B_1\right)\right] z_4 \frac{W_{i+1}^{n} - W_{i-1}^{n}}{2h}.
\end{aligned}
$$

The term T_1 collects the terms known from the linear scheme (2.29), where β is simply a positive constant. The term T_2 involves additional expressions for the nonlinear case. Note also that the nonlinear terms only involve the coefficients at time level n, since the nonlinearity is discretized explicitly. We will make use of this observation in the following by employing Lemma 2.6 to obtain the positivity of T_1 and use this to control the term T_2 for suitably small values of a.

We collect the terms in (2.47) according to the grid points, make use of $W_i^{n+1} = 0$ and obtain

$$
\begin{aligned}
& S(k,h,n,i,V_i^{n+1},V) - S(k,h,n,i,U_i^{n+1},U) \\
=& [1 - r\beta(1+2A_1) + c(a^2) 2r(\beta A_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + A_1)) z_3] W_i^n \\
&+ [r\beta(\tfrac{1}{2} + A_2) + \tfrac{\mu}{2}(\tfrac{1}{2} + B_2)] W_{i+1}^{n+1} \\
&+ [r\beta(\tfrac{1}{2} + A_2) - \tfrac{\mu}{2}(\tfrac{1}{2} + B_2)] W_{i-1}^{n+1} \\
&+ [r\beta(\tfrac{1}{2} + A_1) + \tfrac{\mu}{2}(\tfrac{1}{2} + B_1)] W_{i+1}^{n} \\
&+ c(a^2)[r(\beta A_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + A_1)) z_3 + \tfrac{k}{2h}(\beta B_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + B_1)) z_4] W_{i+1}^n \\
&+ [r\beta(\tfrac{1}{2} + A_1) - \tfrac{\mu}{2}(\tfrac{1}{2} + B_1)] W_{i-1}^{n} \\
&+ c(a^2)[r(\beta B_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + A_1)) z_3 - \tfrac{k}{2h}(\beta B_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + B_1)) z_4] W_{i-1}^n \\
=& \big[b_0 + c(a^2) 2r(\beta A_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + A_1)) z_3\big] W_i^n - a_{-1} W_{i+1}^{n+1} - a_1 W_{i-1}^{n+1} \\
&+ \big[b_{-1} + c(a^2)[r(\beta A_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + A_1)) z_3 \\
&+ \tfrac{k}{2h}(\beta B_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + B_1)) z_4]\big] W_{i+1}^n \\
&+ \big[b_1 + c(a^2)[r(\beta A_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + A_1)) z_3 \\
&- \tfrac{k}{2h}(\beta B_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + B_1)) z_4]\big] W_{i-1}^n
\end{aligned}
$$

$$\begin{aligned}
&\geq \big[b_0 + c(a^2)2r(\beta A_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + A_1))z_3\big]W_i^n \\
&\quad + \big[b_{-1} + c(a^2)[r(\beta A_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + A_1))z_3 \\
&\quad + \tfrac{k}{2h}(\beta B_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + B_1))z_4]\big]W_{i+1}^n \\
&\quad + \big[b_1 + c(a^2)[r(\beta A_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + A_1))z_3 \\
&\quad - \tfrac{k}{2h}(\beta B_{1,\beta}\Psi' + \Psi'(\tfrac{1}{2} + B_1))z_4]\big]W_{i-1}^n.
\end{aligned}$$

where $a_{-1}, a_1, b_{-1}, b_0, b_1$ are the coefficients of the linear scheme (2.29) and where we have employed Lemma 2.6 in the last inequality. Making use of the assumption that Ψ' is bounded, we can control the nonlinear terms by the positive coefficients b_{-1}, b_0, and b_1 if a is sufficiently small. We conclude

$$S(k, h, n, i, V_i^{n+1}, V) - S(k, h, n, i, U_i^{n+1}, U) \geq 0,$$

which completes the proof. □

2.7 Financial example

The Black–Scholes analysis requires continuous trading of the hedged portfolio and this may be expensive in a market with proportional transaction costs. To show the influence of the transaction costs on the price of the European Call option, we compute the price given by the numerical solution of (2.1) and the standard Black–Scholes value for the following choice of parameters

$$\sigma_0 = 0.2, \quad \rho = 0.1, \quad E = 100, \quad T = 0.04.$$

The solutions at time $t = 0.02 = 1$ year are plotted in Figure 2.7 for different values of the transaction cost parameter a. Figure 2.8 shows the difference between the Black–Scholes price and the price given by the solution of (2.1). Since the nonlinear volatility depends on the Gamma (V_{SS}), the difference is small in regions with small Gamma. The difference is not symmetric. The position of the maximal difference is moving in negative direction in time, relating to the negative sign of the convective term in (2.2). At one year the maximal difference is at $S = 95$. The linear Black–Scholes price is about 9.93 whereas the nonlinear price ($a = 0.02$) is about 12.28. The nonlinear price is 23.6 % higher than the linear Black–Scholes price.

In financial context the option price sensitivities are known as 'Greeks'. Mathematically, they are the derivatives of the option price with respect

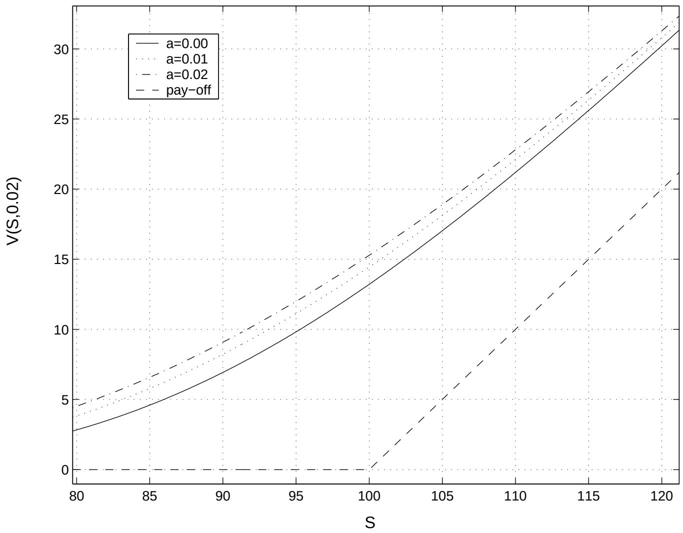

Figure 2.7: Solution of partial differential equation (2.1).

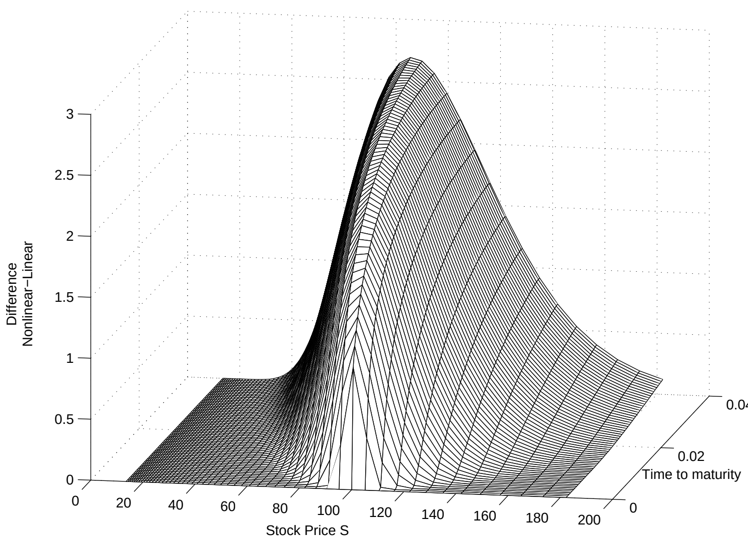

Figure 2.8: Influence of transaction costs.

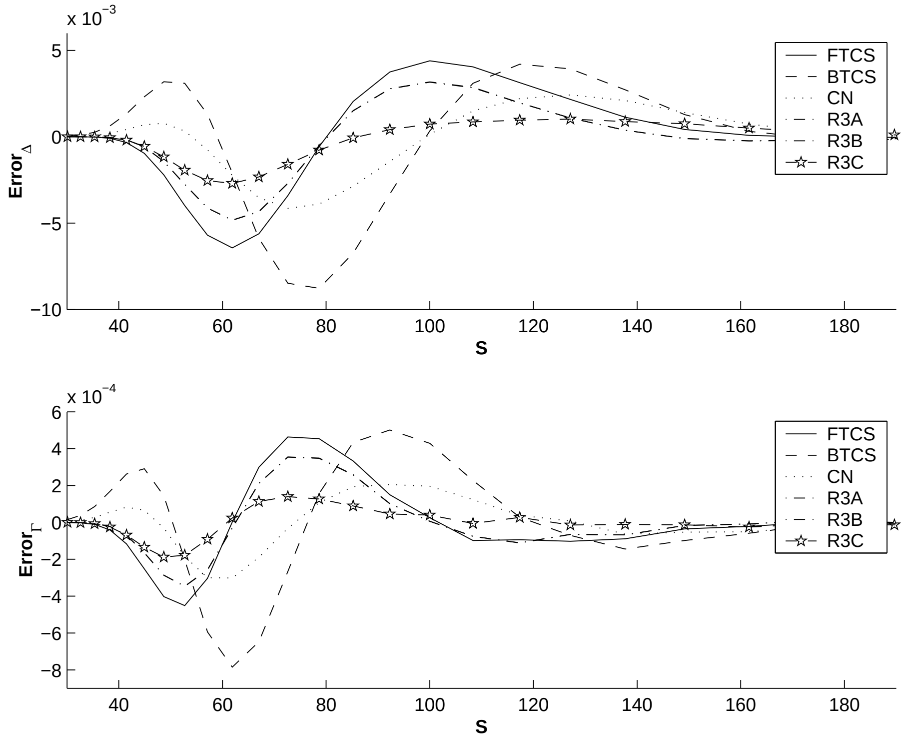

Figure 2.9: Error in the Greeks Delta (upper figure) and Gamma (lower figure).

to the variables or parameters. The most important ones are the first and second derivatives with respect to the price of the underlying stock, called 'Delta' and 'Gamma', respectively. Since price sensitivities are a distinctive measure of risk, growing emphasis on risk management issues has suggested a greater need for their efficient computation.

Figure 2.9 shows the error of the Greeks of the numerical solution computed using 50 grid points. The following parameters were used in the computation

$$a = 0.02, \quad \sigma_0 = 0.2, \quad \rho = 0.1, \quad E = 100, \quad T = 0.02.$$

The Greeks were computed using the standard fourth order central difference approximation of the numerical solutions of (2.1). We observe that the compact scheme R3C gives the best approximation. The Crank–Nicolson scheme and the compact schemes R3A and R3B also produce acceptable results. The errors of the classical schemes (FTCS, BTCS, LFDF) are up to three times larger than those of the compact schemes. The Leap–Frog Du Fort–Frankel scheme even produces spurious oscillations in the derivatives (not shown).

Chapter 3

Parameter Estimation in Option Pricing

3.1 Introduction

The Black–Scholes equation has been derived under several assumptions, in particular the asset price $S(t)$ is supposed to follow a stochastic process

$$dS(t) = \mu S(t)\,dt + \sigma S(t)\,dW(t),$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ are the constant drift and constant volatility of the underlying asset, respectively, and $W(t)$ denotes a Brownian motion. The drift and the volatility are not directly observable. The drift is removed from the model by a hedging argument and does not enter explicitly in the Black–Scholes equation. Obtaining values for σ is often done by computing the so–called *implied volatility* out of observed option prices by inverting the Black–Scholes formula. A widely observed phenomenon is that these computed volatilities are not constant.

The pattern of implied volatilities for different exercise prices sometimes forms a *smile* shape, i.e. implied volatilities of in-the-money and out-of-the-money options are generally higher than that of at-the-money options. This is observed, for example in coffee options markets [**79**]. In equity option markets, typically, one observes a so–called volatility *skew*, i.e. the implied volatility for in-the-money calls is significantly higher than the implied volatility of at-the-money calls and out-of-the-money calls. Additionally, often variation with respect to time to maturity is present as well. This is usually referred to as *volatility term structure.*

These observations lead to a natural generalization of the Black–Scholes model replacing the constant volatility σ in the model by a (deterministic) *local volatility function* $\sigma = \sigma(T, E)$, where T denotes the time to maturity and E the exercise price. It arises the question of how to determine this volatility function from option prices observed in markets, such that the generalized Black–Scholes model replicates the market prices. This problem is often referred to as the *calibration problem*.

As first observed by Dupire [**47**], the option price $V = V(T, E)$ as a function of the exercise time T and the exercise price E satisfies the (forward) differential equation

$$V_T(T,E) - \frac{1}{2}\sigma^2(T,E)E^2V_{EE}(T,E) + rEV_E(T,E) = 0, \quad T > 0, E > 0 \tag{3.1a}$$

with the initial condition

$$V(0,E) = V_0(E) = \max(S_0 - E, 0), \quad E > 0, \tag{3.1b}$$

and boundary conditions

$$V(T,0) = S_0, \quad \lim_{E\to\infty} V(T,E) = 0, \quad T > 0. \tag{3.1c}$$

It is derived from a Fokker–Planck equation integrated twice with respect to the space variable E and using the (formal) identity $(S_0 - E)^+_{EE} = \delta_{S_0}(E)$, where δ_{S_0} denotes the Dirac mass at S_0. Solving (3.1a) for the volatility leads to *Dupire's formula*

$$\sigma(E,T) = \left(\frac{2[V_T(T,E) + rEV_E(T,E)]}{E^2V_{EE}(T,E)}\right)^{1/2}. \tag{3.2}$$

Note that typical option prices are convex in E which implies positivity of the denominator.

Dupire's local volatility function model has received great attention as well as some criticism [**46**]. It was extended in [**48, 90**] who define the local variance as the expectation of the future instantaneous variance conditional on a given asset price. Therein, the (stochastic) instantaneous variance can be quite general, such that this approach is consistent with (univariate diffusion) stochastic volatility models, for example [**80**]. However, if one stays within the completely deterministic setting, (3.1) is the most elaborate model up to our knowledge.

The problem of determining the volatility in (3.1) from observed option prices is an ill–posed optimization problem in the sense of the lack of continuous dependence of the minimizers with respect to perturbations of the problem. In the mathematical literature, there are two main approaches to address the *calibration problem.* The first is to apply equation (3.2) on interpolated data sets of option prices observed in the market [**31, 47, 70**]. It depends largely on the interpolation method but is computationally cheap.

The second one is to use different regularization techniques. Either the problem is reformulated as a stochastic optimal control problem and a so–called entropic regularization [**4**] is performed or a Tikhonov regularized cost functional is used in a (deterministic) inverse problem [**101**]. The last approach has been adopted in many works [**1, 85**]. For a complete review of the literature we refer to [**41**], for a survey on Tikhonov regularization see [**53**].

Most of the references mentioned above focus on the numerical results obtained by standard methods without analyzing in–depth the algorithms used. A theoretical foundation is given in [**40, 41**]. In [**41**] a trinomial tree method using Tikhonov regularization and a probabilistic interpretation of the cost function's gradient is analyzed and numerical results are shown. Convergence rate for Tikhonov regularization under interpretable conditions have been derived in [**51**].

Our goal is to identify from given option prices $V(T, E)$ the volatility function σ in (3.1). We follow the optimal control approach using a Lagrangian framework. The proposed algorithm is based on a Sequential Quadratic Programming Method (SQP) and on a primal–dual active set strategy that guarantees L^∞ constraints for the volatility, in particular assuring its positivity. The algorithm proposed is founded on a thorough analysis of first– and second–order optimality conditions. Furthermore, we prove the existence of a Lagrange multiplier associated with the inequality constraints.

SQP methods have been widely applied to optimization problems of the form

$$\text{minimize } J(x) \text{ subject to } e(x) = 0,$$

where the cost functional $J : X \to \mathbb{R}$ and the constraint $e : X \to Y$ are sufficiently smooth functions and X, Y are real Hilbert spaces. Such problems occur frequently in optimal control of systems described by partial differential equations [**6**]. SQP methods for constrained optimal control of partial differential equations have been studied widely, see for example, [**72**] for projected SQP methods, [**106**] on inexact SQP interior point methods, and [**64**] about

SQP–multigrid techniques. For parameter identification problems they have been applied for example in [**99, 140**]. For a general survey on SQP methods we refer to [**28**] and the references therein.

The basic idea of SQP methods is to minimize at each iteration a quadratic approximation of the Lagrangian associated with the cost functional over an affine subspace of solutions of the linearized constraint. In each level of the SQP method a linear–quadratic subproblem has to be solved. In the presence of bilateral coefficient constraints, this subproblem involves linear inequality constraints. For the solution of the subproblems we use a primal–dual active set method based on a generalized Moreau–Yosida approximation of the indicator function of the admissible control set that was developed in [**18, 76**].

This chapter is organized in the following manner: In Section 3.2 we formulate the parameter estimation as an optimal control problem and prove existence of local optimal solutions. Moreover, any optimal solution is characterized by an optimality system involving an adjoint equation for the Lagrange multiplier. The optimization method is proposed in Section 3.3. We apply a globalized SQP method with a modified Hessian matrix to ensure that every SQP step is a descent direction and implement a line search strategy. In each level of the SQP method a linear–quadratic optimal control problem with box constraints is solved by a primal–dual active set strategy. In Section 3.4 numerical examples are presented and discussed.

3.2 The optimal control problem

In this section the parameter identification problem is introduced as an optimal control problem. We prove existence of at least one optimal solution and present first–order necessary optimality conditions. Furthermore, we investigate sufficient second–order optimality conditions. To streamline the presentation, in the analytical parts of this chapter we restrict ourselves to the case of zero interest rate, i.e. $r = 0$.

3.2.1 Formulation of the optimal control problem

We start by introducing some notation. For $R > E > M > 0$ and $T > 0$ let $\Omega = (M, R)$ be the one–dimensional spatial domain and $Q = (0, T) \times \Omega$ the time–spatial domain. Concerning the error inflicted by introducing artificial boundary conditions, see our related comments in Section 2.2 and [**7, 89**].

We define

$$V = \big\{\varphi \in H^1(\Omega) : \varphi(R) = 0\big\},$$

which is a Hilbert space endowed with the inner product

$$\langle \varphi, \psi \rangle_V = \int_0^R \varphi_x \psi_x \,\mathrm{d}x \quad \text{for all } \varphi, \psi \in V.$$

By $L^2(0,T;V)$ we denote the space of (equivalence classes) of measurable functions $\varphi : [0,T] \to V$, which are square integrable, i.e.,

$$\int_0^T \|\varphi(t)\|_V^2 \,\mathrm{d}t < \infty.$$

Analogously, the spaces $L^2(0,T;H^1(\Omega))$ and $L^2(0,T;L^\infty(\Omega))$ are defined. In particular, $L^2(0,T;L^2(\Omega))$ can be identified with $L^2(Q)$. Moreover we make use of the space

$$W(0,T) = \{\varphi \in L^2(0,T;V) : \varphi_t \in L^2(0,T;V')\},$$

which is a Hilbert space endowed with the common inner product; see [**42**, p. 473]. Let us recall the Hilbert space

$$\begin{aligned} H^{2,1}(Q) &= H^1(0,T;L^2(\Omega)) \cap L^2(0,T;H^2(\Omega)) \\ &= \big\{\varphi : Q \to \mathbb{R} \;\big|\; \varphi, \varphi_t, \varphi_x, \varphi_{xx} \in L^2(Q)\big\}, \end{aligned}$$

supplied with the inner product

$$\langle \varphi, \psi \rangle_{H^{2,1}(Q)} = \int_0^T \int_\Omega \varphi_t \psi_t + \varphi_{xx}\psi_{xx} + \varphi_x \psi_x + \varphi\psi \,\mathrm{d}x\mathrm{d}t \quad \text{for } \varphi, \psi \in H^{2,1}(Q)$$

and the induced norm $\|\cdot\|_{H^{2,1}(Q)} = \langle\cdot\,,\cdot\rangle_{H^{2,1}(Q)}^{1/2}$. Recall that from $\Omega \subset \mathbb{R}$ it follows that $H^{2,1}(Q)$ is continuously embedded into $L^\infty(Q)$; see, e.g. [**110**, p. 24]. We introduce the subspace

$$H_0^{2,1}(Q) = \{\varphi \in H^{2,1}(Q) \mid \varphi(\cdot\,, R) = 0 \text{ in } (0,T)\}.$$

of $H^{2,1}(Q)$. The space $H_0^{2,1}(Q)$ is a Hilbert space endowed with the topology of $H^{2,1}(Q)$.

When t is fixed, the expression $\varphi(t)$ stands for the function $\varphi(t,\cdot)$ considered as a function in Ω only.

Next we specify the set of admissible coefficient functions. Suppose that q_{min} and q_{max} are given functions in $H^{2,1}(Q)\cap L^\infty(0,T;H^2(\Omega))$ satisfying $q_{\mathsf{min}}<q_{\mathsf{max}}$ in Q almost everywhere (a.e.), it exists $C_{\mathsf{ad}}>0$ such that

$$\max\{||q_{\mathsf{min}}||_{L^\infty(0,T;H^2(\Omega))},||q_{\mathsf{max}}||_{L^\infty(0,T;H^2(\Omega))}\}\leq C_{\mathsf{ad}},$$

and $\overline{q}_{\mathsf{min}}=\operatorname{essinf}_{(t,x)\in Q}q_{\mathsf{min}}(t,x)>0$. We introduce the set for the admissible coefficient functions by

$$\mathcal{Q}_{\mathsf{ad}}=\big\{q\in H^{2,1}(Q)\,\big|\,||q||_{L^\infty(0,T;H^2(\Omega))}\leq C_{\mathsf{ad}}\text{ and }q_{\mathsf{min}}\leq q\leq q_{\mathsf{max}}\text{ in }Q\text{ a.e.}\big\},\tag{3.3}$$

which is a closed, bounded and convex set in $H^{2,1}(Q)$. Note, that the bound $C_{\mathsf{ad}}>0$ is purely technical, and can be chosen arbitrarily large.

The goal of the parameter identification is to determine the volatility in (3.1), for streamlining the presentation we restrict ourselves to the case $r=0$ of zero interest rate in the analytical part of this chapter. Therefore, we need to determine the coefficient function $q=q(t,x)=\frac{1}{2}E^2\sigma^2(T,E)$ in the parabolic problem

$$
\begin{aligned}
u_t(t,x)-q(t,x)u_{xx}(t,x)&=0 &&\text{for all }(t,x)\in Q, &&(3.4a)\\
u(t,M)&=u_D(t) &&\text{for all }t\in(0,T), &&(3.4b)\\
u(t,R)&=0 &&\text{for all }t\in(0,T), &&(3.4c)\\
u(0,x)&=u_0(x) &&\text{for all }x\in\Omega &&(3.4d)
\end{aligned}
$$

from given, observed option data $u_T\in L^2(\Omega)$ for the solution u of (3.4) at the final time T. In (3.4c) and (3.4d) we suppose that the boundary and initial data satisfy $u_D\in H^1(0,T)$ and $u_0\in L^2(\Omega)$.

Definition 3.1. *For given $q\in\mathcal{Q}_{\mathsf{ad}}$, $u_D\in H^1(0,T)$ and $u_0\in L^2(\Omega)$ a function u is called a* weak solution to (3.4) *if $u\in W(0,T)$, $u(\cdot,M)=u_D$ in $L^2(0,T)$, $u(0)=u_0$ in $L^2(\Omega)$ and*

$$\int_0^T\langle u_t,\varphi\rangle_{H^{-1},H_0^1}+\Big(\int_\Omega qu_x\varphi_x+q_xu_x\varphi\,\mathrm{d}x\Big)\mathrm{d}t=0,\tag{3.5}$$

for all $\varphi\in L^2(0,T;H_0^1(\Omega))$. In (3.5) $\langle\cdot\,,\cdot\rangle_{H^{-1},H_0^1}$ denotes the duality pairing between $H_0^1(\Omega)$ and its dual space $H^{-1}(\Omega)$.

Remark 3.2. Since $H^{2,1}(Q) \hookrightarrow L^\infty(Q)$ and $q_x \in W(0,T) \hookrightarrow C([0,T], L^2(\Omega))$, the integral in (3.5) is well–defined for every $\varphi \in L^2(0,T;H_0^1(\Omega))$.

The following theorem ensures existence of a weak solution to (3.4) for positive coefficient functions and follows from standard arguments [**100**].

Theorem 3.3. *Suppose that $u_0 \in L^2(\Omega)$ and $u_D \in H^1(0,T)$. Then, for every $q \in \mathcal{Q}_{\mathsf{ad}}$ there exists a unique weak solution u to* (3.4) *and a constant $C > 0$ such that*

$$\|u\|_{W(0,T)} \le C \left(\|u_0\|_{L^2(\Omega)} + \|u_D\|_{H^1(0,T)} \right). \tag{3.6}$$

If the initial condition u_0 is more regular, we have the following result, we omit its proof, because it is standard.

Corollary 3.4. *If $u_0 \in V$ holds with the compatibility condition $u_0(M) = u_D(0)$, it follows that $u \in H_0^{2,1}(Q)$ and there exists a constant $C > 0$ such that*

$$\|u\|_{H^{2,1}(Q)} \le C \left(\|u_0\|_V + \|u_D\|_{H^1(0,T)} \right). \tag{3.7}$$

To write the state equations (3.4) in an abstract form we define the two Hilbert spaces

$$X = H^{2,1}(Q) \times W(0,T) \quad \text{and} \quad Y = L^2(0,T;H_0^1(\Omega)) \times L^2(0,T) \times L^2(\Omega)$$

endowed with their product topologies. Moreover, let

$$K_{\mathsf{ad}} = \mathcal{Q}_{\mathsf{ad}} \times W(0,T)$$

which is closed and convex. In the sequel we identify the dual Y' of Y with the product space $L^2(0,T;H^{-1}(\Omega)) \times L^2(0,T) \times L^2(\Omega)$.

Next we introduce the bilinear operator $e = (e_1, e_2, e_3) : X \to Y'$ by

$$e_1(\omega) = u_t - q u_{xx}, \tag{3.8a}$$
$$e_2(\omega) = u(\cdot, M) - u_D, \tag{3.8b}$$
$$e_3(\omega) = u(0) - u_0, \tag{3.8c}$$

where $\omega = (q,u)$ holds and the identity $e_1(\omega) = u_t - q u_{xx}$ in $L^2(0,T;H^{-1}(\Omega))$ stands for

$$\begin{aligned}&\langle e_1(\omega), \varphi \rangle_{L^2(0,T;H^{-1}(\Omega)), L^2(0,T;H_0^1(\Omega))} \\ &= \int_0^T \langle u_t, \varphi \rangle_{H^{-1}, H_0^1} \,\mathrm{d}t + \int_0^T \int_\Omega q u_x \varphi_x + q_x u_x \varphi \,\mathrm{d}x\mathrm{d}t \quad \text{for } \varphi \in L^2(0,T;H_0^1(\Omega)).\end{aligned}$$

The Fréchet-derivatives with respect to ω are denoted by primes, where subscripts denote as usual the associated partial Fréchet-derivative.

Remark 3.5. From $q \in H^{2,1}(Q)$ we infer that $q_x \in C([0,T];L^2(\Omega))$. Thus, for $\varphi \in L^2(0,T;H_0^1(\Omega))$

$$\Big|\int_0^T\int_\Omega qu_x\varphi_x + q_xu_x\varphi \,\mathrm{d}x\mathrm{d}t\Big| \le \|q\|_{L^\infty(Q)}\|u_x\|_{L^2(Q)}\|\varphi_x\|_{L^2(Q)} + \|q_x\|_{C([0,T];L^2(\Omega))}\|u_x\|_{L^2(Q)}\|\varphi\|_{L^2(0,T;L^\infty(\Omega))}.$$

It follows that the bilinear operator e_1 is well–defined for every $\omega \in X$.

Now we address properties of the operator e. In particular, we prove that e is Fréchet–differentiable and its linearization $e'(\omega)$ is surjective at any point $\omega \in K_{\mathsf{ad}}$ which will later be crucial for the existence of a (unique) Lagrange multiplier λ^* such that the first–order optimality condition is satisfied (see Theorem 3.10).

Proposition 3.6. *The bilinear operator $e : X \to Y'$ is twice continuously Fréchet–differentiable and its second Fréchet–derivative is Lipschitz–continuous on X. Moreover, its linearization $e'(\omega)$ at any point $\omega = (q,u) \in K_{\mathsf{ad}}$ is surjective. Furthermore,we have*

$$\|\delta u\|_{W(0,T)} \le C_1\,\|\delta q\|_{H^{2,1}(Q)} \quad \textit{for all } \delta\omega = (\delta q, \delta u) \in N(e'(\omega)), \tag{3.9}$$

where $N(e'(\omega)) \subset X$ denotes the null space of the linear operator $e'(\omega)$.

Proof. First we prove that e is twice continuously Fréchet–differentiable. For arbitrary directions $(\delta q, \delta u), (\widetilde{\delta q}, \widetilde{\delta u}) \in X$ we compute the directional derivatives as

$$e'(q,u)(\delta q,\delta u) = \begin{pmatrix} \delta u_t - q\delta u_{xx} - \delta q u_{xx} \\ \delta u(\cdot, M) \\ \delta u(0) \end{pmatrix} \tag{3.10}$$

and

$$e''(q,u)\big((\delta q,\delta u),(\widetilde{\delta q},\widetilde{\delta u})\big) = \begin{pmatrix} -\widetilde{\delta q}\delta u_{xx} - \delta q\widetilde{\delta u}_{xx} \\ 0 \\ 0 \end{pmatrix}. \tag{3.11}$$

Using Young's inequality we obtain

$$
\begin{aligned}
&\int_0^T\int_\Omega \delta q\delta u_x\varphi_x + \delta q_x\delta u_x\varphi \,\mathrm{d}x\mathrm{d}t\\
\le&\big(\|\delta q\|_{L^\infty(Q)}\|\delta u_x\|_{L^2(Q)} + \|\delta q_x\|_{C([0,T];L^2(\Omega))}\|\delta u_x\|_{L^2(\Omega)}\big)\|\varphi\|_{L^2(0,T;H_0^1(\Omega))}\\
\le&C\big(\|\delta q\|^2_{H^{2,1}(Q)} + \|\delta u\|^2_{W(0,T)}\big)\|\varphi\|_{L^2(0,T;H_0^1(\Omega))}\\
\le&C\|\delta\omega\|^2_X\|\varphi\|_{L^2(0,T;H_0^1(\Omega))}
\end{aligned}
$$

for all $\varphi \in L^2(0,T;H_0^1(\Omega))$ and $\delta\omega = (\delta q, \delta u) \in X$. Hence,

$$
\begin{aligned}
&\|e_1(q+\delta q, u+\delta u) - e_1(q,u) - \nabla e_1(q,u)(\delta q,\delta u)\|_{L^2(0,T;H^{-1}(\Omega))}\\
=&\sup_{\|\varphi\|_{L^2(0,T;H_0^1(\Omega))}=1}\int_0^T\int_\Omega \delta q\delta u_x\varphi_x + \delta q_x\delta u_x\varphi\,\mathrm{d}x\mathrm{d}t\\
\le&C\,\|\delta\omega\|^2_X
\end{aligned}
$$

and thus

$$
\lim_{\|\delta\omega\|_X\to 0}\frac{\|e_1(\omega+\delta\omega) - e_1(\omega) - e_1'(\omega)\delta\omega\|_{L^2(0,T;H^{-1}(\Omega))}}{\|\delta\omega\|_X} = 0. \tag{3.12}
$$

Notice that — due to the linearity of the operators e_2 and e_3 — we have

$$
\|e_2(\omega+\delta\omega) - e_2(\omega) - e_2'(\omega)\delta\omega\|_{L^2(0,T)} = 0 \tag{3.13}
$$

and

$$
\|e_3(\omega+\delta\omega) - e_3(\omega) - e_3'(\omega)\delta\omega\|_{L^2(\Omega)} = 0. \tag{3.14}
$$

Consequently, we infer from (3.12), (3.13) and (3.14) that the operator e is Fréchet differentiable with Fréchet derivative (3.10). Now we turn to the second derivative. In view of (3.11)

$$
\|e_1'(\omega+\widetilde{\delta\omega})\delta\omega - e_1'(\omega)\delta\omega - e_1''(\omega)(\delta\omega,\widetilde{\delta\omega})\|_{L^2(0,T;H^{-1}(\Omega))} = 0,
$$

and $e_2''(\omega) = e_3''(\omega) = 0$ holds, we infer that e is twice Fréchet–differentiable and the directional derivative, given in (3.11), is the second Fréchet derivative of e. Since $e''(\omega)$ does not depend on $\omega \in X$, the Lipschitz–continuity on X is obvious.

It remains to prove that $e'(\omega)$ is surjective and that the estimate (3.9) is satisfied for all $\delta\omega \in N(e'(\omega))$. Suppose that $r = (r_1, r_2, r_3) \in Y'$ is arbitrary. Then the operator $e'(\omega)$ is surjective, if there exists a pair $\delta\omega = (\delta q, \delta u) \in X$ such that $e'(\omega)(\delta\omega) = r$, which is equivalent to

$$\delta u_t - q\delta u_{xx} = r_1 + \delta q u_{xx} \qquad \text{in } L^2(0,T;H^{-1}(\Omega)), \tag{3.15a}$$

$$\delta u(\cdot, M) = r_2 \qquad \text{in } L^2(0,T), \tag{3.15b}$$

$$\delta u(0) = r_3 \qquad \text{in } L^2(\Omega). \tag{3.15c}$$

Choosing $\delta q = 0$ there exists a unique $\delta u \in W(0,T)$, which solves (3.15). Hence $e'(\omega)$ is surjective.

Let $\delta\omega = (\delta q, \delta u) \in N(e'(\omega))$. Estimate (3.9) follows from standard arguments. For that reason we only estimate the additional right–hand side in (3.15a), namely the term $\delta q u_{xx}$. Since $H^1(\Omega)$ and V are continuously embedded into $L^\infty(\Omega)$, the space $H_0^{2,1}(Q)$ is continuously embedded into $L^\infty(Q)$ as well as into $C([0,T];V)$, we infer from Hölder's and Young's inequalities

$$\begin{aligned}\int_0^t \int_\Omega \big(\delta q \delta u\big)_x u_x \,\mathrm{d}x\mathrm{d}s &\le \int_0^t \|\delta q\|_{L^\infty(\Omega)} \|\delta u_x\|_{L^2(\Omega)} \|u_x\|_{L^2(\Omega)} \,\mathrm{d}s \\ &\quad + \int_0^t \|\delta q_x\|_{L^2(\Omega)} \|\delta u\|_{L^\infty(\Omega)} \|u_x\|_{L^2(\Omega)} \,\mathrm{d}s \\ &\le C(\varepsilon)\, \|\delta q\|^2_{H^{2,1}(Q)} + \varepsilon\, \|\delta u\|^2_{W(0,T)}\end{aligned}$$

for almost all $t \in [0,T]$ and for every $\varepsilon > 0$, where the constant $C(\varepsilon) > 0$ depends on $\|u\|_{L^2(0,T;V)}$ and ε. Choosing ε appropriately and using standard arguments the estimate follows. □

Remark 3.7. It follows from the proof of Proposition 3.6 that at any point $\omega \in K_{\mathsf{ad}}$ the operator $e_u : W(0,T) \to Y'$ is even bijective.

Next we introduce the cost functional $J : X \to [0,\infty)$ by

$$J(\omega) = \frac{1}{2} \int_\Omega |u(T) - u_T|^2 \,\mathrm{d}x + \frac{\beta}{2} \|q\|^2_{H^{2,1}(Q)} \quad \text{for } \omega = (q,u) \in X, \tag{3.16}$$

where u_T is a given observed option price at the end–time T, and $\beta > 0$ is a regularization parameter.

Lemma 3.8. *The cost functional $J : X \to [0,\infty)$ is Fréchet–differentiable and its Fréchet derivatives are given by*

$$J'(\omega)\delta\omega = \int_\Omega \big(u(T)-u_T\big)\delta u(T)\,\mathrm{d}x + \beta \int_0^T \int_\Omega q_t\delta q_t + q\delta q + q_x\delta q_x + q_{xx}\delta q_{xx}\,\mathrm{d}x\mathrm{d}t \tag{3.17}$$

and

$$J''(\omega)(\delta\omega,\widetilde{\delta\omega}) = \int_\Omega \delta u(T)\widetilde{\delta u}(T)\,\mathrm{d}x + \beta \int_0^T \int_\Omega \delta q_t\widetilde{\delta q}_t + \delta q\widetilde{\delta q} + \delta q_x\widetilde{\delta q}_x + \delta q_{xx}\widetilde{\delta q}_{xx}\,\mathrm{d}x\mathrm{d}t \tag{3.18}$$

for arbitrary directions $\delta\omega = (\delta q, \delta u), \widetilde{\delta\omega} = (\widetilde{\delta q}, \widetilde{\delta u}) \in X$. In particular, the second Fréchet derivative is Lipschitz–continuous on X.

Proof. For all $\delta u \in W(0,T)$ we have $\delta u(T) \in L^2(\Omega)$ (see, e.g. [**42**, p. 480]) so that the integrals are well–defined. It follows by standard arguments that the first and second Fréchet derivative are given by (3.17) and (3.18), respectively. Since $\omega \mapsto J''(\omega)$ does not depend on ω, the second Fréchet derivative is clearly Lipschitz–continuous on X. □

Using the operator e we can express the parameter identification problem as a constrained optimal control problem in the following form

$$\min J(\omega) \quad \text{s.t.} \quad \omega \in K_{\mathsf{ad}} \text{ and } e(\omega) = 0. \tag{$\mathbf{P}$}$$

Note that in our formulation, both the state variable u and the coefficient q are considered as independent variables while the realization of (3.4) is an explicit constraint. Alternatively, one could use the equality constraint to treat $u = u(q)$ as a variable dependent on the unknown coefficient q and solve the nonlinear least–squares problem by the Gauss–Newton method.

In this chapter, we choose the SQP approach with independent variables. SQP methods can be viewed as a natural extension of Newton and quasi–Newton methods to the constrained optimization setting [**77**], and are hence expected to inherit its fast local convergence property. Indeed, the iterates of the SQP method are identical to those generated by Newton's method when applied to the system composed of the first–order necessary conditions for the Lagrangian associated with ($\mathbf{P}$) and the constraint equation. Note that SQP methods are not feasible–point methods, i.e. its iterates need not be points satisfying the constraints. This is a great advantage, in particular if the constraints are nonlinear.

3.2.2 Existence of optimal solutions

The next theorem guarantees that $(\mathbf{P})$ has an optimal solution.

Theorem 3.9. *Problem* $(\mathbf{P})$ *possesses at least one (global) solution* $\omega^* = (q^*, u^*) \in K_{\mathsf{ad}}$.

Proof. In view of Theorem 3.3 the admissible set

$$E = \{\omega = (q, u) \in X : e(\omega) = 0 \text{ in } Y' \text{ and } \omega \in K_{\mathsf{ad}}\} \tag{3.19}$$

is non–empty (from $q_{\mathsf{min}} \in \mathcal{Q}_{\mathsf{ad}}$ follows $(q_{\mathsf{min}}, u(q_{min})) \in E$). Moreover, $J(\omega) \geq 0$ holds for all $\omega \in E$. Thus there exists a $\zeta \geq 0$ such that

$$\zeta = \inf\{J(\omega) : \omega \in E\}. \tag{3.20}$$

We infer that there exists a minimizing sequence $(\omega^n)_{n\in\mathbb{N}} \subset E$, $\omega^n = (q^n, u^n)$, with

$$\lim_{n\to\infty} J(\omega^n) = \zeta.$$

Due to (3.6) and

$$J(\omega^n) \geq \frac{\beta}{2} \|q^n\|^2_{H^{2,1}(Q)} \quad \text{for all } n,$$

we infer that the sequence $(\omega^n)_{n\in\mathbb{N}}$ is bounded in X. Thus, there exist subsequences, again denoted by $(\omega^n)_{n\in\mathbb{N}}$, and a pair $\omega^* = (q^*, u^*) \in X$ satisfying

$$\begin{aligned} q^n &\rightharpoonup q^* && \text{in } H^{2,1}(Q) && \text{as } n \to \infty, \\ u^n &\rightharpoonup u^* && \text{in } W(0,T) && \text{as } n \to \infty. \end{aligned} \tag{3.21}$$

Furthermore, since $q^n \in \mathcal{Q}_{\mathsf{ad}}$ and it holds $L^\infty(0,T;H^2(\Omega)) \cap H^1(0,T;L^2(\Omega)) \hookrightarrow C([0,T];H^1(\Omega))$ compactly due to Aubin's lemma [**132**], we obtain

$$q^n \to q^* \quad \text{in } L^\infty(0,T;H^1(\Omega)) \quad \text{as } n \to \infty. \tag{3.22}$$

In view of (3.22), (3.21) it holds

$$\int_0^T \int_\Omega q^n u^n_x \varphi_x + q^n_x u^n_x \varphi \,\mathrm{d}x\mathrm{d}t \to \int_0^T \int_\Omega q^* u^*_x \varphi_x + q^*_x u^*_x \varphi \,\mathrm{d}x\mathrm{d}t$$

as $n \to \infty$ for every $\varphi \in L^2(0,T;H^1_0(\Omega))$. Therefore,

$$\lim_{n\to\infty} e_1(\omega^n) = e_1(\omega^*) \quad \text{in } L^2(0,T;H^{-1}(\Omega)).$$

From $e_1(\omega^n) = 0$ for all $n \in \mathbb{N}$ we conclude that $e_1(\omega^*) = 0$. Since the operators e_2 and e_3 are linear, we find $e(\omega^*) = 0$. Since J is convex and continuous, and therefore weakly lower semi–continuous, we obtain $J(\omega^*) \leq \lim_{n\to\infty} J(\omega^n) = \zeta$. Finally, since $\mathcal{Q}_{\mathsf{ad}}$ is convex and closed in $H^{2,1}(Q)$, and therefore weakly closed, we have $q^* \in \mathcal{Q}_{\mathsf{ad}}$, and the claim follows. □

3.2.3 First–order necessary optimality conditions

Problem (**P**) is a non–convex programming problem so that different local minima might occur. A numerical method will produce a local minimum close to its starting value. Hence, we do not restrict our investigations to global solutions of (**P**). We will assume that a fixed reference solution $\omega^* = (q^*, u^*) \in K_{\mathsf{ad}}$ is given satisfying certain first– and second–order optimality conditions (ensuring local optimality of the solution).

In this section we introduce the Lagrange functional associated with (**P**) and derive first–order necessary optimality conditions. Furthermore, we show that there exists a unique Lagrange multiplier associated with the inequality constraints for the optimal coefficient q^*.

To formulate the optimality conditions we introduce the Lagrange functional $L : X \times Y \to \mathbb{R}$ associated with (**P**) by

$$\begin{aligned} L(\omega, p) &= J(\omega) + \langle e(\omega), (\lambda, \mu, \nu)\rangle_{Y',Y} \\ &= \frac{1}{2}\left\|u(T) - u_T\right\|^2_{L^2(\Omega)} + \frac{\beta}{2}\left\|q\right\|^2_{H^{2,1}(Q)} + \int_\Omega \big(u(0) - u_0\big)\nu \,\mathrm{d}x \\ &\quad + \int_0^T \langle u_t, \lambda\rangle_{H^{-1},H^1_0} \,\mathrm{d}t + \int_0^T\int_\Omega (q\lambda)_x u_x \,\mathrm{d}x\mathrm{d}t + \int_0^T \big(u(\cdot, M) - u_D\big)\mu \,\mathrm{d}t, \end{aligned}$$

with $\omega = (q, u) \in X$ and $p = (\lambda, \mu, \nu) \in Y$. Due to Proposition 3.6 and Lemma 3.8 the Lagrangian is twice continuously Fréchet–differentiable with respect to $\omega \in X$ for each fixed $p \in Y$ and its second Fréchet derivative is Lipschitz–continuous.

We infer from Proposition 3.6 and Remark 3.7 that an optimal solution to (**P**) can be characterized by first–order necessary optimality conditions. This is formulated in the next theorem. Recall that the set E has been introduced in (3.19). Moreover, let

$$B_\rho(\omega) = \big\{\tilde{\omega} \in X : \left\|\tilde{\omega} - \omega\right\|_X < \rho\big\}$$

be the open ball in X with radius $\rho > 0$ and mid point $\omega \in X$.

Theorem 3.10. *Suppose that $\omega^* = (q^*, u^*) \in K_{\mathsf{ad}}$ is a local solution to* $(\mathbf{P})$, *i.e., $\omega^* \in E$ and there exists a constant $\rho > 0$ such that*

$$J(\omega^*) \leq J(\omega) \quad \textit{for all } \omega \in E \cap B_\rho(\omega^*).$$

Then there are unique Lagrange multipliers $p^* = (\lambda^*, \mu^*, \nu^*) \in Y$ satisfying the adjoint equations

$$-\lambda_t^* - (q^*\lambda^*)_{xx} = 0 \qquad \textit{in } Q, \tag{3.23a}$$
$$\lambda^*(\cdot\,, M) = \lambda^*(\cdot\,, R) = 0 \qquad \textit{in } (0,T), \tag{3.23b}$$
$$\lambda^*(T) = -(u^*(T) - u_T) \qquad \textit{in } \Omega \tag{3.23c}$$

in the weak sense and the identities

$$\mu^* = (q^*\lambda^*)_x(\cdot, M) \qquad \textit{in } L^2(0,T), \tag{3.24}$$
$$\nu^* = \lambda^*(0) \qquad \textit{in } L^2(\Omega) \tag{3.25}$$

hold. Moreover, the variational inequality

$$\langle \beta q^* - \mathcal{R}(\lambda^* u_{xx}^*), q - q^* \rangle_{H^{2,1}(Q)} \geq 0 \quad \textit{for all } q \in \mathcal{Q}_{\mathsf{ad}} \tag{3.26}$$

holds, where $\mathcal{R} : (H^{2,1}(Q))' \to H^{2,1}(Q)$ denotes the Riesz isomorphism, i.e., $q = \mathcal{R}(f) \in H^{2,1}(Q)$ solves

$$\int_0^T \int_\Omega q_t\varphi_t + q_{xx}\varphi_{xx} + q_x\varphi_x + q\varphi \,\mathrm{d}x\mathrm{d}t = \langle f, \varphi \rangle_{(H^{2,1}(Q))', H^{2,1}(Q)} \quad \textit{for all } \varphi \in H^{2,1}(Q)$$

with $f \in (H^{2,1}(Q))'$. Here, $\langle \cdot\,, \cdot \rangle_{(H^{2,1}(Q))', H^{2,1}(Q)}$ denotes the duality pairing between $H^{2,1}(Q)$ and its dual.

Proof. We infer from Proposition 3.6 and Remark 3.7 that a standard constraint qualification holds at (q^*, u^*) [**111**]. Therefore, there exists a unique Lagrange multiplier $p^* = (\lambda^*, \mu^*, \nu^*) \in Y$ such that

$$L_q(\omega^*, p^*)(q - q^*) \geq 0 \qquad \text{for all } q \in \mathcal{Q}_{\mathsf{ad}}, \tag{3.27}$$
$$L_u(\omega^*, p^*)u = 0 \qquad \text{for all } u \in W(0,T), \tag{3.28}$$
$$L_p(\omega^*, p^*)p = 0 \qquad \text{for all } p \in Y. \tag{3.29}$$

Equation (3.29) is equivalent to the equality constraint $e(\omega^*) = 0$ and is fulfilled since ω^* solves $(\mathbf{P})$. Next we turn to (3.28), which is equivalent to

$$\begin{aligned} 0 \;=\; & \int_\Omega (u^*(T) - u_T)u(T)\,\mathrm{d}x + \int_0^T \langle u_t, \lambda^*\rangle_{H^{-1},H_0^1(\Omega)}\,\mathrm{d}t \\ & + \int_0^T\int_\Omega (q^*\lambda^*)_x u_x\,\mathrm{d}x\mathrm{d}t + \int_0^T u(\cdot, M)\mu^*\,\mathrm{d}t + \int_\Omega u(0)\nu^*\,\mathrm{d}x \end{aligned} \tag{3.30}$$

for all $u \in W(0,T)$. In particular, (3.30) holds for all $u(t,x) = \chi(t)\psi(x)$ with $\chi \in C_0^1(0,T)$ and $\psi \in H_0^1(\Omega) \subset V$. Consequently,

$$\int_0^T\int_\Omega \chi_t\psi\lambda^* + (q^*\lambda^*)_x\chi\psi'\,\mathrm{d}x\mathrm{d}t = 0 \tag{3.31}$$

for all $\chi \in C_0^1(0,T)$ and $\psi \in H_0^1(\Omega)$. Notice that

$$\int_0^T\int_\Omega \chi_t\psi\lambda^*\,\mathrm{d}x\mathrm{d}t = \int_\Omega \Big(\int_0^T \chi_t\lambda^*\,\mathrm{d}t\Big)\psi\,\mathrm{d}x = \Big\langle -\int_0^T \lambda_t^*\chi\,\mathrm{d}t, \psi\Big\rangle_{H^{-1},H_0^1}, \tag{3.32}$$

where λ_t^* denotes the distributional derivative of λ^* with respect to t. The remaining term in (3.31) leads to

$$\int_0^T\int_\Omega (q^*\lambda^*)_x\psi'\chi\,\mathrm{d}x\mathrm{d}t = \Big\langle -\int_0^T (q^*\lambda^*)_{xx}\chi\,\mathrm{d}t, \psi\Big\rangle_{H^{-1},H_0^1}. \tag{3.33}$$

Inserting (3.32) and (3.33) into (3.31) we get

$$\Big\langle \int_0^T \big(-\lambda_t^* - (q^*\lambda^*)_{xx}\big)\chi\,\mathrm{d}t, \psi\Big\rangle_{H^{-1},H_0^1} = 0, \tag{3.34}$$

for all $\chi \in C_0^1(0,T)$ and $\psi \in H_0^1(\Omega)$. Notice that $q^* \in \mathcal{Q}_{\mathsf{ad}}$ implies $q^* \in L^\infty(Q)$ as well as $q_x^* \in L^\infty(0,T;L^2(\Omega))$. Therefore, it follows $(q^*\lambda^*)_x \in L^2(Q)$ and, consequently, $(q^*\lambda^*)_{xx} \in L^2(0,T;H^{-1}(\Omega))$. The set

$$\{\varphi \in L^2(0,T;H_0^1(\Omega)) : \varphi(t,x) = \chi(t)\psi(x) \text{ with } \chi \in C_0^1(0,T) \text{ and } \psi \in H_0^1(\Omega)\}$$

is dense in $L^2(0,T;H_0^1(\Omega))$ so that $\lambda_t^* \in L^2(0,T;H^{-1}(\Omega))$ and (3.23a) hold. Moreover,

$$\int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \langle \lambda^*, u\rangle_{L^2(\Omega)} \,\mathrm{d}t \;=\; \langle \lambda_t^*, u\rangle_{L^2(0,T;H^{-1}(\Omega)),L^2(0,T;H_0^1(\Omega))} + \langle u_t, \lambda^*\rangle_{L^2(0,T;H^{-1}(\Omega)),L^2(0,T;H_0^1(\Omega))} \tag{3.35}$$

for $u \in W(0,T)$. Hence, we may apply (3.30), (3.34), and (3.35) to obtain

$$\begin{aligned}
0 &= \int_\Omega (u^*(T) - u_T)u(T)\,\mathrm{d}x + \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \langle \lambda^*, u\rangle_{L^2(\Omega)} \,\mathrm{d}t \\
&\quad + \int_0^T \langle -\lambda_t^* - (q^*\lambda^*)_{xx}, u\rangle_{H^{-1},H_0^1(\Omega)} + \int_0^T (q^*\lambda^*)_x u\Big|_{x=M}^{x=R} \,\mathrm{d}t \\
&\quad + \int_0^T \mu^* u(\cdot, M)\,\mathrm{d}t + \int_\Omega \nu^* u(0)\,\mathrm{d}x \\
&= \langle (u^*(T) - u_T + \lambda^*(T), u(T)\rangle_{L^2(\Omega)} + \langle \nu^* - \lambda^*(0), u(0)\rangle_{L^2(\Omega)} \\
&\quad + \langle \mu^* - (q^*(\cdot,M)\lambda^*(\cdot,M))_x, u(\cdot,M)\rangle_{L^2(0,T)}.
\end{aligned}$$

Choosing appropriate test functions in $W(0,T)$, we find (3.23c), (3.24), and (3.25). Finally, we consider (3.27). We compute

$$L_q(q^*,u^*,p^*)q = \int_0^T \int_\Omega \beta\big(q_t^* q_t + q^* q + q_x^* q_x + q_{xx}^* q_{xx}\big) + (q\lambda^*)_x u_x^* \,\mathrm{d}x\mathrm{d}t \tag{3.36}$$

for all $q \in \mathcal{Q}_{\mathsf{ad}}$. For $\lambda^* \in L^2(0,T;H_0^1(\Omega))$ and $u_x^* \in L^2(Q)$ the integral

$$\int_0^T \int_\Omega (q\lambda^*)_x u_x^* \,\mathrm{d}x\mathrm{d}t$$

is bounded for all $q \in \mathcal{Q}_{\mathsf{ad}}$. Moreover, $(\lambda^* u_x^*)(\cdot, M) = (\lambda^* u_x^*)(\cdot, R) = 0$ holds. Thus, the function $g = -\lambda^* u_{xx}^*$ can be identified with an element in $(H^{2,1}(Q))'$ and we derive from (3.36)

$$L_q(q^*,u^*,p^*)q = \beta \langle q^*, q\rangle_{H^{2,1}(Q)} + \langle g, q\rangle_{(H^{2,1}(Q))',H^{2,1}(Q)} \tag{3.37}$$

for all $q \in \mathcal{Q}_{\mathsf{ad}}$. Employing the Riesz isomorphism $\mathcal{R}$, inserting (3.37) into (3.27) we find

$$\langle \beta q^* - \mathcal{R}(\lambda^* u^*_{xx}), q - q^* \rangle_{H^{2,1}(Q)} \geq 0 \quad \text{for all } q \in \mathcal{Q}_{\mathsf{ad}},$$

which is the variational inequality (3.26). □

Remark 3.11. The usage of the Riesz operator $\mathcal{R} : (H^{2,1}(Q))' \to H^{2,1}(Q)$ in (3.26) requires to solve a problem of the form

$$-u_{tt} + u_{xxxx} - u_{xx} + u = f \quad \text{in } Q,$$

including initial and boundary conditions. Hence, in our numeric realization we will employ the 'weaker' norm in $L^2(0, T; H^1(\Omega))$, see Section 3.3. Then $\mathcal{R}$ can be replaced by the Riesz operator $\widetilde{\mathcal{R}} : H^{-1}(\Omega) \to H^1(\Omega)$, that requires only the solution of the Neumann problem

$$-u(t)_{xx} + u(t) = f(t) \quad \text{in } \Omega, \quad u(t)_x|_{\delta\Omega} = 0,$$

for a.e. $t \in (0, T)$.

Utilizing variational techniques we can prove the following error estimate for the adjoint variable λ^*.

Corollary 3.12. *Let all hypotheses of Theorem* 3.10 *hold. Then there exists a constant $C_2 > 0$ depending on $\|q^*\|_{L^\infty(0,T;L^4(\Omega))}$ and $\overline{q}_{\mathsf{min}}$ such that*

$$\|\lambda^*\|_{L^2(0,T;H^1_0(\Omega))} \leq C_2 \, \|u^*(T) - u_T\|_{L^2(\Omega)}$$

Hence, if the residual $\|u^*(T) - u_T\|_{L^2(\Omega)}$ becomes small the norm of the Lagrange multiplier λ^* is small. We will make use of this estimate in the next section.

From Theorem 3.10 we infer the existence of a Lagrange multiplier associated with the constraint $q^* \in \mathcal{Q}_{\mathsf{ad}}$. To formulate the result we introduce the following sets.

Definition 3.13. *Let K be a convex subset of a (real) Banach space Z and $z^* \in K$. The* cone of feasible directions R_K at the point z^**, the* tangent cone T_K at the point z^* *and the* normal cone N_K at the point z^* *are defined by*

$$\begin{aligned} R_K(z^*) &= \{z \in Z : \exists \sigma > 0 : z^* + \sigma z \in K\}, \\ T_K(z^*) &= \{z \in Z : \exists z^*(\sigma) = z^* + \sigma z + o(\sigma) \in K, \sigma \geq 0\}, \\ N_K(z^*) &= \{z \in Z' : \langle z, \tilde{z} - z^* \rangle_{Z',Z} \leq 0 \textit{ for all } \tilde{z} \in K\}. \end{aligned}$$

In case of $z^* \notin K$ the normal cone $N_K(z^*)$ is set equal to the empty set.

In the following we choose $Z = H^{2,1}(Q)$, $K = \mathcal{Q}_{\mathsf{ad}}$ and $z^* = q^*$.

Corollary 3.14. *Let all hypotheses of Theorem* 3.10 *be satisfied. Then there exists a Lagrange multiplier $\tilde{\xi}^* \in N_{\mathcal{Q}_{\mathsf{ad}}}(q^*)$ associated with the inequality constraints such that*

$$L_q(\omega^*, p^*) + \tilde{\xi}^* = 0 \quad \textit{in } (H^{2,1}(Q))'. \tag{3.38}$$

Proof. Defining $\tilde{\xi}^* = -\beta q^* + \lambda^* u^*_{xx} \in (H^{2,1}(Q))'$ and using (3.26) we obtain $\tilde{\xi}^* \in N_{\mathcal{Q}_{\mathsf{ad}}}(q^*)$. In particular, (3.38) follows. □

Remark 3.15. Using the Riesz isomorphism $\mathcal{R}$ introduced in Theorem 3.10 we can identify $\tilde{\xi}^* \in (H^{2,1}(Q))'$ with an element in the Hilbert space $H^{2,1}(Q)$ by setting $\xi^* = -\beta q^* + \mathcal{R}(\lambda^* u^*_{xx})$.

Let $\omega^* = (q^*, u^*) \in K_{\mathsf{ad}}$ denote a local solution to $(\mathbf{P})$. If the solution $q^* \in \mathcal{Q}_{\mathsf{ad}}$ is inactive with respect to the norm constraint, i.e., $\|q^*\|_{L^\infty(0,T;H^2(\Omega))} < C_{\mathsf{ad}}$, then $(\mathbf{P})$ is locally equivalent to

$$\min J(\omega) \quad \text{s.t.} \quad \omega \in \hat{K}_{\mathsf{ad}} \text{ and } e(\omega) = 0, \tag{$\hat{\mathbf{P}}$}$$

where $\hat{K}_{\mathsf{ad}} = \hat{\mathcal{Q}}_{\mathsf{ad}} \times W(0,T)$ and

$$\hat{\mathcal{Q}}_{\mathsf{ad}} = \big\{ q \in H^{2,1}(Q) \;\big|\; q_{\mathsf{min}} \le q \le q_{\mathsf{max}} \text{ in } Q \text{ a.e.} \big\},$$

which is a closed, convex and bounded subset in $L^2(Q)$. We define by

$$\hat{E} = \{ \omega \in \hat{K}_{\mathsf{ad}} : e(\omega) = 0 \}$$

the admissible set of $(\hat{\mathbf{P}})$. Monitoring the sequence $\|q^n\|_{L^\infty(0,T;H^2(\Omega))}$ we solve $(\hat{\mathbf{P}})$ in our numerical experiments, see Section 3.4 below. For that reason we focus on $(\hat{\mathbf{P}})$ in the remainder of this section.

The Lagrange multiplier ξ^* associated with the inequality constraints for the optimal coefficient q^* is characterized by the following corollary.

Corollary 3.16. *Let all hypotheses of Theorem* 3.10 *be satisfied. Suppose that $\|q^*\|_{L^\infty(0,T;H^2(\Omega))} < C_{\mathsf{ad}}$. Then ξ^* satisfies*

$$\xi^*\big|_{A^*_-} \le 0, \quad \xi^*\big|_{A^*_+} \ge 0, \quad \xi^*\big|_{I^*} = 0, \tag{3.39}$$

where

$$\begin{aligned}A_-^* &= \{(t,x) \in \overline{Q} : q^*(t,x) = q_{\mathsf{min}}(t,x)\},\\ A_+^* &= \{(t,x) \in \overline{Q} : q^*(t,x) = q_{\mathsf{max}}(t,x)\},\\ I^* &= \{(t,x) \in \overline{Q} : q_{\mathsf{min}}(t,x) < q^*(t,x) < q_{\mathsf{max}}(t,x)\}\end{aligned}$$

are the active and inactive sets for the optimal coefficient q^*.

Proof. The proof uses similar arguments as the proof of Theorem 2.3 in [**76**]. Therefore, we give only the proof of $\xi^*\big|_{A_-^*} \le 0$. Define

$$\begin{aligned}A_-^{\geq} =&\{(t,x) \in \overline{Q} : (q^*(t,x) = q_{\mathsf{min}}(t,x)) \wedge (\xi^* > 0)\},\\ A_-^{\geq,l} =&\{(t,x) \in A_-^{\geq} : \xi^* > \frac{1}{l}\},\\ C_-^l =&\{(t,x) \in A_-^{\geq,l} : q_{\mathsf{max}}(t,x) - q_{\mathsf{min}}(t,x) > \frac{1}{l}\}.\end{aligned}$$

Assume that $A_-^{\geq}$ has positive measure $\mu(A_-^{\geq}) > \varepsilon > 0$. Since

$$\mu\{(t,x) \in \overline{Q} : q_{\mathsf{max}}(t,x) = q_{\mathsf{min}}(t,x)\} = 0$$

and $A_-^{\geq,l} \uparrow A_-^{\geq}$, it follows $\mu(C_-^l) > 0$ for l sufficiently large and $C_-^l \uparrow A_-^{\geq}$. Hence there exists $l > 0$ such that $\mu(C_-^l) > \varepsilon$ because of the lower continuity of μ. Define $\delta \in (H^{2,1}(Q))'$ by $\varphi \mapsto \int_0^T \int_\Omega (q_{\mathsf{max}} - q_{\mathsf{min}})\chi_{C_-^l}\varphi \,\mathrm{d}x\mathrm{d}t$ and its Riesz representative by $\mathcal{R}(\delta) \in H^{2,1}(Q)$. Recall that $\xi^* = -\beta q^* + \mathcal{R}(\lambda^* u_{xx}^*)$ by Remark 3.15 and consider the directional derivative (see (3.37))

$$\begin{aligned}L_q(q^*,u^*,p^*)\mathcal{R}(\delta) &= \langle \mathcal{R}(\delta), \beta q^* - \mathcal{R}(\lambda^* u_{xx}^*)\rangle_{H^{2,1}}\\ &= \langle \delta, -\xi^*\rangle_{(H^{2,1})',H^{2,1}}\\ &= -\int_0^T \int_\Omega (q_{\mathsf{max}} - q_{\mathsf{min}})\chi_{C_-^l}\xi^* \,\mathrm{d}x\mathrm{d}t\\ &< -\frac{\varepsilon}{l^2} < 0.\end{aligned}$$

This contradicts the optimality of q^*. Hence, $\mu(A_-^{\geq}) = 0$. □

The primal–dual active set algorithm used below makes use of the following result from convex analysis [**75, 84**]. Using the generalized Moreau–Yosida regularization of the indicator function $\chi_{\hat{\mathcal{Q}}_{\mathsf{ad}}}$ of the convex set $\hat{\mathcal{Q}}_{\mathsf{ad}}$ of admissible controls, i.e.,

$$\chi_{\hat{\mathcal{Q}}_{\mathsf{ad}}}(q^*) = \inf_{q\in H^{2,1}(Q)} \left\{\chi_{\hat{\mathcal{Q}}_{\mathsf{ad}}}(q^* - q) + \langle \xi^*, q\rangle_{H^{2,1}(Q)} + \frac{c}{2}\|q\|^2_{H^{2,1}(Q)}\right\}$$

with $c > 0$, one can replace $q^* \in \hat{\mathcal{Q}}_{\mathsf{ad}}$ and condition (3.39) by

$$q^* = \mathcal{P}_{\mathsf{ad}}\Big(q^* + \frac{\xi^*}{c}\Big) \text{ for every } c > 0, \tag{3.40}$$

where

$$\mathcal{P}_{\mathsf{ad}} : L^2(Q) \to \{q \in L^2(Q) \mid q_{\mathsf{min}} \le q \le q_{\mathsf{max}} \text{ in } Q \text{ a.e.}\}$$

by

$$\mathcal{P}_{\mathsf{ad}}(q)(t,x) = \begin{cases} q_{\mathsf{min}}(t,x) & \text{if } q(t,x) < q_{\mathsf{min}}(t,x), \\ q(t,x) & \text{if } q_{\mathsf{min}}(t,x) \le q(t,x) \le q_{\mathsf{max}}(t,x), \\ q_{\mathsf{max}}(t,x) & \text{if } q(t,x) > q_{\mathsf{max}}(t,x) \end{cases}$$

for almost all $(t,x) \in Q$. The primal–dual active set method uses identification (3.40) as a prediction strategy, i.e. given a current primal–dual iteration pair (q_k, ξ_k) and arbitrarily fixed $c > 0$ the next active and inactive sets are given by

$$\mathcal{A}_-^k = \big\{(t,x) \in Q \,|\, q_k + \tfrac{\xi_k}{c} < q_{\mathsf{min}}\big\}, \quad \mathcal{A}_+^k = \big\{(t,x) \in Q \,|\, q_k + \tfrac{\xi_k}{c} > q_{\mathsf{max}}\big\},$$
$$\mathcal{I}^k = Q \setminus (\mathcal{A}_-^k \cup \mathcal{A}_+^k).$$

Note that (3.40) is equivalent to the differential inclusion [**6**] $\xi^* \in \partial\chi_{\hat{\mathcal{Q}}_{\mathsf{ad}}}(q^*)$ where $\partial\chi_{\hat{\mathcal{Q}}_{\mathsf{ad}}}$ denotes the subdifferential of the indicator function $\chi_{\hat{\mathcal{Q}}_{\mathsf{ad}}}$.

3.2.4 Second–order analysis

In Section 3.2.3 we have investigated first–order necessary optimality conditions for ($\hat{\mathbf{P}}$). To ensure that a solution (ω^*, p^*) satisfying $\omega^* = (q^*, u^*) \in \hat{E}$, $q^* \in \hat{\mathcal{Q}}_{\mathsf{ad}}$, (3.23) and (3.39) indeed solves ($\hat{\mathbf{P}}$), we have to guarantee second–order sufficient optimality. This is the focus of this section. We review different second–order optimality conditions and set them into relation. Then, we prove that the second–order sufficient optimality condition holds, provided the residual $\|u^*(T) - u_T\|_{L^2(\Omega)}$ is sufficiently small.

For any directions $\delta\omega = (\delta q, \delta u), \widetilde{\delta\omega} = (\widetilde{\delta q}, \widetilde{\delta u}) \in X$ the second Fréchet–derivative of the Lagrangian is given by

$$L''(\omega,p)(\delta\omega,\widetilde{\delta\omega}) = \beta \int_0^T \int_\Omega \delta q_t \widetilde{\delta q}_t + \delta q \widetilde{\delta q} + \delta q_x \widetilde{\delta q}_x + \delta q_{xx} \widetilde{\delta q}_{xx} \,\mathrm{d}x\mathrm{d}t$$
$$+ \int_\Omega \delta u(T) \widetilde{\delta u}(T) \,\mathrm{d}x + \int_0^T \int_\Omega (\delta q \lambda)_x \widetilde{\delta u}_x + (\widetilde{\delta q}\lambda)_x \delta u_x \,\mathrm{d}x\mathrm{d}t$$

with $\omega = (q, u) \in X$ and $p = (\lambda, \mu, \nu) \in Y$. In particular, we set

$$\begin{aligned}\mathsf{Q}(\delta\omega) &= L''(\omega, p)(\delta\omega, \delta\omega) \\ &= \|\delta u(T)\|^2_{L^2(\Omega)} + \beta\,\|\delta q\|^2_{H^{2,1}(Q)} + 2\int_0^T\!\!\int_\Omega (\delta q\lambda)_x \delta u_x \,\mathrm{d}x\mathrm{d}t\end{aligned}$$

for $\delta\omega \in X$. From the boundedness of the second derivative of the Lagrangian we infer that Q is continuous.

Lemma 3.17. *The quadratic form Q is weakly lower semi–continuous. Moreover, let $(\delta\omega^n)_{n\in\mathbb{N}}$ be a sequence in $N(e'(\omega))$, $\omega = (q, u) \in X$, with $\delta\omega^n \rightharpoonup 0$ in X and $\mathsf{Q}(\delta\omega^n) \to 0$ as $n \to \infty$. Then it follows that $\delta\omega^n \to 0$ strongly in X.*

Proof. Note that for $\delta\omega = (\delta q, \delta u) \in X$ holds

$$\mathsf{Q}(\delta\omega) = J''(\omega)(\delta\omega, \delta\omega) + 2\int_0^T\!\!\int_\Omega (\delta q\lambda)_x \delta u_x \,\mathrm{d}x\mathrm{d}t.$$

Note that $\delta\omega \mapsto J''(\omega)(\delta\omega, \delta\omega)$ is weakly lower semi–continuous. Since the integral is even weakly continuous (see the proof of Theorem 3.9), it follows that Q is weakly lower semi–continuous on X. Now assume that $(\delta\omega^n)_{n\in\mathbb{N}} = (\delta q^n, \delta u^n)_{n\in\mathbb{N}}$ is a sequence in $N(e'(\omega))$ with $\delta\omega^n \rightharpoonup 0$ in X and $\mathsf{Q}(\delta\omega^n) \to 0$ as $n \to \infty$. Analogously as in the proof of Theorem 3.9 we derive that

$$\lim_{n\to\infty} \int_0^T\!\!\int_\Omega (\delta q^n\lambda)_x \delta u^n_x \,\mathrm{d}x\mathrm{d}t = 0.$$

Since $\mathsf{Q}(\delta\omega^n)$ converges to zero, it follows that for every $\varepsilon > 0$ there exists an $n_\varepsilon \in \mathbb{N}$ such that

$$0 \le J''(\omega)(\delta\omega^n, \delta\omega^n) < \varepsilon \quad \text{for all } n \ge n_\varepsilon.$$

In particular, this implies that

$$\beta\,\|\delta q^n\|^2_{H^{2,1}(Q)} < \varepsilon \quad \text{for all } n \ge n_\varepsilon,$$

which gives $\delta q^n \to 0$ in $H^{2,1}(Q)$ as $n \to \infty$. Since $\delta\omega^n \in N(e'(\omega))$ holds, we infer from Proposition 3.6 that $\delta u^n \to 0$ in $W(0, T)$ as $n \to \infty$. ☐

Let us recall the following definition, see [**30**].

Definition 3.18. *Let* $\omega^* = (q^*, u^*) \in \hat{E}$.

a) *The point* ω^* *is a local solution to* $(\hat{\mathbf{P}})$ *satisfying the* quadratic growth condition *if there exists a* $\rho > 0$ *satisfying*

$$J(\omega) \geq J(\omega^*) + \rho \, \|\omega - \omega^*\|_X^2 + o(\|\omega - \omega^*\|_X^2) \quad \text{for all } \omega \in \hat{E}. \tag{3.41}$$

b) *Suppose that* ω^* *satisfies the first–order necessary optimality conditions with associated unique Lagrange multipliers* $p^* \in Y$ *and* $\xi^* \in N_{\hat{\mathcal{Q}}_{\mathsf{ad}}}(q^*)$. *At* (ω^*, p^*) *the* second–order sufficient optimality condition *holds if there exists a* $\kappa > 0$ *such that*

$$L''(\omega^*, p^*)(\delta\omega, \delta\omega) \geq \kappa \, \|\delta\omega\|_X^2 \quad \text{for all } \delta\omega \in C(\omega^*), \tag{3.42}$$

where

$$C(\omega^*) = \left\{\delta\omega \in \left(T_{\hat{\mathcal{Q}}_{\mathsf{ad}}}(q^*) \cap (\xi^*)^\perp\right) \times W(0,T) : \delta\omega \in N(e'(\omega^*))\right\}$$

denotes the critical cone at ω^*, $^\perp$ *denotes the orthogonal complement in* $H^{2,1}(Q)$ *and* $T_{\hat{\mathcal{Q}}_{\mathsf{ad}}}(q^*)$ *the tangential cone at* q^* *(introduced in Def. 3.13).*

The critical cone $C(\omega^*)$ is the set of directions that are tangent to the feasible set. It turns out that (3.41) and (3.42) are related to the weaker condition

$$L''(\omega^*, p^*)(\delta\omega, \delta\omega) > 0 \quad \text{for all } \delta\omega \in C(\omega^*) \setminus \{0\}, \tag{3.43}$$

which is very close to the necessary optimality condition. In particular, the following theorem holds.

Theorem 3.19. *The quadratic growth condition* (3.41)*, the second–order sufficient optimality condition* (3.42)*, and* (3.43) *are equivalent.*

Proof. The proof is similar to the proof of Theorem 2.3 in [**30**]. We show (3.41) $\Longrightarrow$ (3.42) $\Longrightarrow$ (3.43) $\Longrightarrow$ (3.41). Assume that (3.41) holds, ω^* satisfies the first–oder necessary optimality conditions and let $\delta\omega \in R_{\hat{\mathcal{Q}}_{\mathsf{ad}}}(q^*) \cap (\xi^*)^\perp \cap N(e'(\omega^*))$. Due to (3.41),

$$J(\omega^* + t\delta\omega) - J(\omega^*) \geq \rho t^2 \|\delta\omega\|_X^2 + o(t^2),$$

which implies $L''(\omega^*, p^*)(\delta\omega, \delta\omega) \geq 2\rho\|\delta\omega\|_X^2$. Then (3.42) follows from the polyhedricity of $\hat{\mathcal{Q}}_{\mathsf{ad}}$, i.e., $\overline{R_{\hat{\mathcal{Q}}_{\mathsf{ad}}}(q^*) \cap (\xi^*)^\perp} = T_{\hat{\mathcal{Q}}_{\mathsf{ad}}}(q^*) \cap (\xi^*)^\perp$ (cf.[**30**, Prop.2.2]).

The implication (3.42) $\Longrightarrow$ (3.43) is trivial. Assume now, that (3.43) is satisfied, while (3.41) is not. Then there exists $(\omega_n)_{n\in\mathbb{N}} : \omega_n \to \omega^*$ in $\hat{E}$ such that

$$J(\omega_n) < J(\omega^*) + \frac{1}{n}\|\omega_n - \omega^*\|_X^2. \tag{3.44}$$

Extracting if necessary a subsequence we can write $\omega_n - \omega^* = t_n v_n$, $t_n \in \mathbb{R}^+$, $t_n \downarrow 0$, $\|v_n\|_X = 1$, and $v_n \rightharpoonup v$ in X. Using second order expansion,

$$L(\omega_n, p^*) = L(\omega^*, p^*) + t_n L'(\omega^*, p^*)v_n + \frac{t_n^2}{2} L''(\omega^*, p^*)(v_n, v_n) + o(t_n^2).$$

Noting that $e(\omega^n) = e(\omega^*) = 0$ and using (3.44), it can be seen from

$$\begin{aligned} J(\omega_n) &= J(\omega^*) + t_n L'(\omega^*, p^*)v_n + \frac{t_n^2}{2} L''(\omega^*, p^*)(v_n, v_n) + o(t_n^2) \\ &< J(\omega^*) + \frac{t_n^2}{n}\|v_n\|_X^2, \end{aligned}$$

that $L'(\omega^*, p^*)v = 0$. Since $L'(\omega^*, p^*)v_n \geq 0$ due to (3.27), (3.28), it follows

$$\limsup_n L''(\omega^*, p^*)(v_n, v_n) = \limsup_n \mathsf{Q}(v_n) \leq 0.$$

Since $v_n \in T_{\hat{\mathcal{Q}}_{\mathsf{ad}}}(q^*) \cap (\xi^*)^\perp$ and $T_{\hat{\mathcal{Q}}_{\mathsf{ad}}}(q^*) \cap (\xi^*)^\perp$ is weakly closed, v is a critical direction. Due to Lemma 3.17, Q is weakly lower semi–continuous and therefore, $\mathsf{Q}(v) \leq 0$. Since the integral is weakly continuous, we obtain with $v = (q, u)$ and $v_n = (q_n, u_n)$

$$\begin{aligned} 2\int_0^T \int_\Omega (q\lambda)_x u_x \,\mathrm{d}x\mathrm{d}t &= 2\lim_n \int_0^T \int_\Omega (q_n\lambda)_x u_{n,x} \,\mathrm{d}x\mathrm{d}t \\ &\leq -\liminf_n (\|u_n(T)\|_{L^2(\Omega)}^2 + \beta\,\|q_n\|_{H^{2,1}(Q)}^2) \\ &\leq -C \end{aligned}$$

for some $C > 0$ since $\|v_n\|_X = 1$. Hence $v \neq 0$, which contradicts (3.43). $\square$

In the next theorem we present a sufficient condition for the second–order sufficient optimality condition (3.42).

Theorem 3.20. *Let all hypotheses of Theorem* 3.10 *be satisfied. Then* (3.42) *holds provided* $\|u^*(T) - u_T\|_{L^2(\Omega)}$ *is sufficiently small.*

Proof. Applying estimate (3.9) and Hölder's inequality, we estimate for arbitrary $\delta\omega = (\delta q, \delta u) \in N(e'(\omega^*))$

$$
\begin{aligned}
& L''(\omega^*, p^*)(\delta\omega, \delta\omega) \\
\geq & \beta \, \|\delta q\|^2_{H^{2,1}(Q)} + 2 \int_0^T \int_\Omega (\delta q \lambda^*)_x \delta u_x \, \mathrm{d}x\mathrm{d}t \\
\geq & \frac{\beta}{2} \, \|\delta q\|^2_{H^{2,1}(Q)} + \frac{\beta}{2C_1} \, \|\delta u\|^2_{W(0,T)} - 2C\|\delta q\|_{H^{2,1}(Q)}\|\lambda^*\|_{L^2(0,T;H^1_0(\Omega))}\|\delta u\|_{W(0,T)} \\
\geq & \Big(\frac{\beta}{2}\min\{1, 1/C_1\} - CC_2\|u^*(T) - u_T\|_{L^2(\Omega)}\Big) \Big(\|\delta q\|^2_{H^{2,1}(Q)} + \|\delta u\|^2_{W(0,T)}\Big),
\end{aligned}
$$

where we used Corollary 3.12 and Young's inequality in the last inequality. Suppose that

$$
\|u^*(T) - u_T\|_{L^2(\Omega)} < \frac{\beta \min\{1, 1/C_1\}}{2CC_2},
$$

we conclude (3.42). □

3.3 The optimization method

In this section we turn to the optimization algorithm used to solve the parameter identification problem (**P**). We suppose that $\|q^*\|_{L^\infty(0,T;H^2(\Omega))} < C_{\mathsf{ad}}$. Since empirical results suggest that the volatility is quite regular, this is no severe restriction for our application. Hence we solve ($\hat{\mathbf{P}}$) instead of (**P**). To solve ($\hat{\mathbf{P}}$) we apply a globalized SQP method. The globalization is realized by a modification of the Hessian matrix to ensure that every SQP step is a descent direction and by a line search strategy. Since in each level of the SQP method a linear–quadratic optimal control problem with box constraints has to be solved we utilize a primal–dual active set strategy.

The SQP method is addressed in Section 3.3.1. The solution of the equality constrained optimal control problem is done by an SQP method. Section 3.3.2 is concerned with the primal–dual active set method. The line search strategy is discussed in Section 3.3.3.

3.3.1 The SQP method

In the numerical realization we initialize our SQP method by taking a function $u^0 \in W(0,T)$, which satisfies $u(\cdot, M) = u_D$ in $(0,T)$ and $u(0) = u_0$ in Ω. Hence, the next iterate $u^{n+1} = u^n + \delta u^n$, $n \geq 0$, can be determined by choosing δu^n in the linear space

$$\mathfrak{U} = \big\{\delta u \in W(0,T) \,|\, \delta u(t,\cdot) \in H_0^1(\Omega) \text{ for } t \in (0,T) \text{ and } \delta u(0) = 0 \text{ in } \Omega\big\},$$

which is a Hilbert space endowed with the topology of $W(0,T)$. Thus, the constraints $e_2(\omega^n) = 0$ and $e_3(\omega^n) = 0$ are guaranteed by construction for any $n \geq 0$. Consequently, there is only one constraint that is $e_1(\omega) = 0$ in $L^2(0,T;H^{-1}(\Omega))$. Therefore, the Lagrange variables μ and ν are redundant. It follows from

$$\mu^* = (q^*\lambda^*)_x|_{(\cdot,M)} \quad \text{and} \quad \nu^* = \lambda^*(0)$$

that μ^* and ν^* can be computed after determining the optimal coefficient q^* and the Lagrange multiplier λ^* associated with the constraint $e_1(\omega^*) = 0$. We set $Y_1 = L^2(0,T;H_0^1(\Omega))$ and identify its dual space Y_1' with $L^2(0,T;H^{-1}(\Omega))$.

In numerical tests we tried out different norms for the regularization term. It turned out that by replacing the $H^{2,1}(Q)$–norm in (3.16) by the 'weaker' $L^2(0,T;H^1(\Omega))$–norm, relative to a prior $q_d \in X_1$, we get good results while saving computational effort, since we only have to solve a Neumann–problem for q in every time–step (see Remark 3.11).

The linear–quadratic minimization problems that have to be solved in each step of the SQP method are well–defined provided $L''(\omega^n, \lambda^n)$ is coercive on $N(e'(\omega^n))$ and $e'(\omega^n)$ on $N(e'(\omega^n))$ is surjective for every iterate. Often these requirements hold only locally so we consider in the following a globalization strategy using a modified Hessian. For $\gamma \in [0,1]$ let us define the function ${}^\gamma\mathsf{L} : X \times Y_1 \to \mathbb{R}$ by

$${}^\gamma\mathsf{L}(\omega,\lambda) = J(\omega) + \gamma \, \langle e_1(\omega), \lambda \rangle_{Y_1',Y_1} \quad \text{for } (\omega,\lambda) \in X \times Y_1.$$

Notice that for $\gamma = 1$ the function ${}^\gamma\mathsf{L}$ is the usual Lagrangian associated with the single constraint $e_1(\omega) = 0$ in Y_1'. Therefore, we set $\mathsf{L} = {}^\gamma\mathsf{L}$ for $\gamma = 1$.

Algorithm 3.21 (SQP method).

1) *Choose* $\omega^0 = (q^0, u^0) \in X = X_1 \times X_2$ *that satisfies* $u^0(\cdot, M) = u_D$ *in* $(0,T)$ *and* $u^0(0) = u_0$ *in* Ω, $\lambda^0 \in Y_1$ *with* $\lambda^0(T) = u_T - u^0(T)$ *in* Ω. *Fix relative and absolute stopping tolerances* $1 > \varepsilon_{\mathsf{rel}} \geq \varepsilon_{\mathsf{abs}} > 0$. *Choose maximal number of SQP iterations* $n_{\mathsf{sqp}} \in \mathbb{N}$. *Set* $n := 0$ *and* $\bar\kappa \in (0,1]$.

2) *For* $\omega^n = (q^n, u^n)$ *evaluate* $e_1(\omega^n)$, $e_1'(\omega^n)$, $J(\omega^n)$, $J'(\omega^n)$. *If*

$$\|\nabla \mathsf{L}(\omega^n, \lambda^n)\|^2_{X' \times Y_1'} = \|\mathsf{L}'(\omega^n, \lambda^n)\|^2_{X'} + \|e_1(\omega^n)\|^2_{Y_1'} < \varepsilon_{\mathsf{abs}}$$

or

$$\|\nabla \mathsf{L}(\omega^n, \lambda^n)\|^2_{X' \times Y_1'} < \varepsilon_{\mathsf{rel}} \, \|\nabla \mathsf{L}(\omega^0, \lambda^0)\|^2_{X' \times Y_1'}$$

or $n = n_{\mathsf{sqp}}$, *then STOP. Otherwise continue with step* 3).

3) *Set* $\gamma = 1$.

4) *Solve for* $\delta\omega^n = (\delta q^n, \delta u^n)$ *the following linear–quadratic optimal control problem*

$$\begin{aligned} &\min \mathsf{L}(\omega^n, \lambda^n) + \mathsf{L}'(\omega^n, \lambda^n)\delta\omega + \frac{1}{2}\,{}^{\gamma}\mathsf{L}''(\omega^n, \lambda^n)(\delta\omega^n, \delta\omega^n) \\ &\text{s.t.} \begin{cases} e_1'(\omega^n)\delta\omega^n + e_1(\omega^n) = 0 & \text{in } Y_1', \\ q_{\mathsf{min}} \le q^n + \delta q^n \le q_{\mathsf{max}} & \text{in } X_1. \end{cases} \end{aligned} \tag{QP}$$

5) *If* $\|\delta\omega\|_X > 0$ *evaluate the quotient*

$$\kappa^n = \frac{{}^{\gamma}\mathsf{L}''(\omega^n, \lambda^n)(\delta\omega^n, \delta\omega^n)}{\|\delta q^n\|^2_{X_1} + \|\delta u^n\|^2_{L^2(0,T;L^2(\Omega))}}.$$

where

$${}^{\gamma}\mathsf{L}''(\omega^n, \lambda^n)(\delta\omega^n, \delta\omega^n) = J''(\omega^n)(\delta\omega^n, \delta\omega^n) + \gamma \, \langle e_1''(\omega^n)(\delta\omega^n, \delta\omega^n), \lambda^n \rangle_{Y_1', Y_1}$$

holds. If $\kappa^n \le 0$ *then set* $\gamma = 0$ *and go back to step* 4). *If* $\kappa^n \in (0, \bar{\kappa})$ *then set* $\bar{\kappa} = \kappa^n$.

6) *Determine a step size parameter* $\alpha_n \in (0, 1]$ *by a backtracking line search (see Section 3.3.3).*

7) *Set* $q^{n+1} = q^n + \alpha_n \delta q^n$, $u^{n+1} = u^n + \alpha_n \delta u^n$, $\lambda^{n+1} = \lambda^n + \alpha_n \delta\lambda^n$, *set* $n := n + 1$, *and go back to step 2).*

Remark 3.22. Alternatively, γ can be adjusted by using the following iteration strategy after step 3):

a) *Choose* $\eta \in (0, 1)$ *and set* $i = 0$.

b) *Perform steps 4) and 5).*

c) *If* $\kappa^n \leq 0$ *set*

$$\gamma := \min \left\{ \eta, \frac{\bar{\kappa}\,\|\delta\omega^n\|_X^2 - J''(\omega^n)(\delta\omega^n, \delta\omega^n)}{\langle e_1''(\omega^n)(\delta\omega^n, \delta\omega^n), \lambda^n\rangle_{Y_1', Y_1}} \right\}. \tag{3.45}$$

Set $i := i + 1$ *and go back to* b)*.*

Since $J''(\omega^n)(\delta\omega^n, \delta\omega^n) \geq 0$, it follows that $\langle e_1''(\omega^n)(\delta\omega^n, \delta\omega^n), \lambda^n\rangle_{Y_1', Y_1} < 0$ from $\kappa^n \leq 0$. The procedure using (3.45) is less strict then setting directly $\gamma = 0$ if $\kappa^n \leq 0$. However, it may involve solving **(QP)** several times, i.e., more often than at most two times as needed by the strategy that switches directly from $\gamma = 1$ to $\gamma = 0$.

3.3.2 The primal–dual active set method

To solve the linear–quadratic optimal control problem (**QP**) in step 4) of Algorithm 3.21 we apply the primal–dual active set method [**76**]. It involves both primal variable and dual variables and is therefore different from conventional active set strategies that involve primal variables only, see e.g. [**130**]. In practice, the algorithm behaves like an infeasible one, since its iterates violate the constraints up to the last–but–one iterate. The algorithm stops at a feasible and optimal solution. Based on the identification (3.40) for the inactive and active sets, the easy–to–implement algorithm exhibits a low number of iterations to find the optimal solution and is very robust [**76, 18**]. Note also that the algorithm uses only one Lagrange multiplier to realize both inequality constraints, which reduces the number of variables and hence the amount of memory needed by its implementation.

Suppose that we are at level n of the SQP method. Thus, we have iterates (q^n, u^n, λ^n) and start our primal–dual active set strategy to determine $(\delta q^n, \delta u^n, \delta\lambda^n)$. Let us define $q^n_{\mathsf{min}} = q_{\mathsf{min}} - q^n$ and $q^n_{\mathsf{max}} = q_{\mathsf{max}} - q^n$. Then we are looking for a step δq^n satisfying the inequality constraints $q^n_{\mathsf{min}} \leq \delta q^n \leq q^n_{\mathsf{max}}$ in Q. The method uses identification (3.40) as a prediction strategy, i.e. given a current primal–dual iteration pair $(\delta q_k, \xi_k)$ the next inactive and active sets are given by (3.46)–(3.48) (see below). The method stops at a feasible and optimal solution as soon as two consecutive tuples of active and inactive sets are equal. This solution is then used in Algorithm 3.21.

Instead of solving problem (**QP**) with inequality constraints we solve

$$\min \mathsf{L}(\omega^n,\lambda^n)+\mathsf{L}'(\omega^n,\lambda^n)\delta\omega+\frac{1}{2}{}^{\gamma}\mathsf{L}''(\omega^n,\lambda^n)(\delta\omega^n,\delta\omega^n)$$
$$\text{s.t.}\begin{cases} e_1'(\omega^n)\delta\omega^n+e_1(\omega^n)=0 & \text{in } Y_1', \\ \delta q_k = q^n_{\mathsf{min}} & \text{in } \mathcal{A}^k_-, \\ \delta q_k = q^n_{\mathsf{max}} & \text{in } \mathcal{A}^k_+, \\ q^n_{\mathsf{min}} \le q^n+\delta q_k \le q^n_{\mathsf{max}} & \text{in } X_1. \end{cases} \qquad (\mathbf{QP}_{\mathcal{I}_k})$$

Hence, it is solved for δq_k only on the inactive set $\mathcal{I}_k$, but δq_k is fixed on the active set $\mathcal{A}^k=\mathcal{A}^k_-\cup\mathcal{A}^k_+$. We have the following algorithm.

Algorithm 3.23 (Primal–dual active set method).

(1) Choose initial values $(\delta q_0,\xi_0)\in X_1\times X_1$, the parameter $c>0$ and set $k:=0$.

(2) Determine the active and inactive sets

$$\mathcal{A}^k_- = \big\{(t,x)\in Q \,|\, \delta q_k+\frac{\xi_k}{c} < q^n_{\mathsf{min}}\big\}, \tag{3.46}$$

$$\mathcal{A}^k_+ = \big\{(t,x)\in Q \,|\, \delta q_k+\frac{\xi_k}{c} > q^n_{\mathsf{max}}\big\}, \tag{3.47}$$

$$\mathcal{I}^k = Q\setminus\mathcal{A}^k. \tag{3.48}$$

with $\mathcal{A}^k=\mathcal{A}^k_-\cup\mathcal{A}^k_+$, and set

$$\widetilde{\delta q_k}=\begin{cases} q^n_{\mathsf{min}} & \text{in } \mathcal{A}^k_-, \\ q^n_{\mathsf{max}} & \text{in } \mathcal{A}^k_+, \\ \delta q_k & \text{in } \mathcal{I}^k. \end{cases}$$

(3) If $k>0$ holds then STOP provided $\mathcal{A}^k=\mathcal{A}^{k-1}$ holds.

(4) Solve for $(\delta q_k,\delta u_k,\delta\lambda_k)\in X_1\times\mathfrak{U}\times Y_1$ the linearized state equations

$$(\delta u_k)_t - q^n(\delta u_k)_{xx} - \widetilde{\delta q_k}u^n_{xx} = -(u^n_t-q^nu^n_{xx}) \quad \text{in } Q, \tag{3.49a}$$
$$\delta u_k(\cdot\,,M)=\delta u_k(\cdot\,,R)=0 \quad \text{in } (0,T), \tag{3.49b}$$
$$\delta u_k(0)=0 \quad \text{in } \Omega, \tag{3.49c}$$

the linearized adjoint system

$$-(\delta\lambda_k)_t - (q^n\delta\lambda_k)_{xx} - \gamma(\widetilde{\delta q_k}\lambda^n)_{xx} = -(-\lambda^n_t - (q^n\lambda^n)_{xx}) \quad \text{in } Q, \tag{3.49d}$$

$$\delta\lambda_k(\cdot, M) = \delta\lambda_k(\cdot, R) = 0 \quad \text{in } (0,T), \tag{3.49e}$$

$$\delta\lambda_k(T) + \delta u_k(T) = u_T - u^n(T) - \lambda^n(T) \quad \text{in } \Omega, \tag{3.49f}$$

and the linearized optimality condition on the inactive set

$$\begin{aligned}\beta(-(\delta q_k)_{xx} + \delta q_k) - \delta\lambda_k u^n_{xx} - \gamma\lambda^n(\delta u_k)_{xx}& \\ = -\Big(\beta\big(-(q^n - q_d)_{xx} + (q^n - q_d)\big) - \widetilde{\mathcal{R}}(\lambda^n u^n_{xx})\Big) &\quad \text{in } \mathcal{I}^k,\end{aligned} \tag{3.49g}$$

where (q^n, u^n, λ^n) *denotes the current SQP iterates and* $\widetilde{\mathcal{R}} = (-\Delta + \mathrm{Id})^{-1}$ *is the Riesz operator between* $H^{-1}(\Omega)$ *and* $H^1(\Omega)$. *Recall that we have chosen* λ^0 *in such a way that* $\lambda^0(T) = u_T - u^0(T)$ *holds in* Ω. *Thus, by induction we have* $\lambda^n(T) = u_T - u^n(T)$ *in* Ω *and* (3.49f) *reads*

$$\delta\lambda_k(T) + \delta u_k(T) = 0 \quad \text{in } \Omega.$$

(5) Set $q^n_k = q^n + \delta q_k$, $u^n_k = u^n + \delta u_k$, $\lambda^n_k = \lambda^n + \delta\lambda_k$ *and*

$$\xi_k = -\beta(-(q^n_k - q_d)_{xx} + q^n_k - q_d) + \widetilde{\mathcal{R}}(\lambda^n_k(u^n_k)_{xx}) \quad \text{in } Q.$$

From (3.49g) *we infer that* $\xi_k = 0$ *on* $\mathcal{I}_k$. *Set k:=k+1 and go back to step (2).*

Remark 3.24. In [**83**] a global and local convergence analysis is done for the nonlinear primal–dual active set strategy in an abstract setting. It turns out that it converges globally as $k \to \infty$ provided a certain merit function satisfies a sufficient decrease condition. This merit function depends on integrals of the positive part functions $\max(0, \delta q_k - q^n_{\mathsf{max}})$ and $\max(0, q^n_{\mathsf{min}} - \delta q_k)$ over the set Q as well as of the positive part functions $\max(0, -\xi_k)$ and $\max(0, \xi_k)$ over the active sets. If, in addition, a smoothness and Lipschitz condition hold, the method converges locally superlinearly.

3.3.3 The line search strategy

In this section we address the line search strategy employed in step 6) of Algorithm 3.21. The goal is to find a compromise between the descent of the cost functional J and the reduction of the violation of the constraints $e(\omega) = 0$. This is realized by a line search method utilizing a suitable chosen merit function. Of course, there are many proposals for merit functions in the literature; see, for instance, in [**19, 130**]. Here, we use the exact penalty functional

$$\varphi^n(\alpha^n) = J(\omega^n) + \mu \left\| e_1(\omega^n + \alpha^n \delta\omega^n) \right\|_{Y_1'}, \tag{3.50}$$

where $\alpha^n \in (0, 1]$ is the step size parameter, which has to be determined, and $\mu > 0$ is a parameter penalizing violations of the constraints $e_1(\omega) = 0$. The reason for our choice (3.50) comes from the fact that — apart from evaluating the norm — our merit function does not introduce an additional nonlinearity into our problem.

In a first–order variation we use the approximation

$$\begin{aligned}\bar{\varphi}^n(\alpha^n) &= J(\omega^n) + \alpha^n J'(\omega^n)\delta\omega^n + \mu \left\| e_1(\omega^n) + \alpha^n e_1'(\omega^n)\delta\omega^n \right\|_{Y_1'} \\ &= J(\omega^n) + \alpha^n J'(\omega^n)\delta\omega^n + \mu(1 - \alpha^n) \left\| e_1(\omega^n) \right\|_{Y_1'},\end{aligned}$$

where we have used that $e_1(\omega^n) + e_1'(\omega^n)\delta\omega^n = 0$ holds for the SQP step $\delta\omega^n$. Notice that $\bar{\varphi}^n(1)$ does not depend on μ. For appropriately chosen penalty parameter $\mu > 0$, our line search strategy is based on the well–known Armijo rule (see e.g. [**29, 68**])

$$\varphi^n(\alpha^n) - \varphi^n(0) \le c\alpha^n \left(\bar{\varphi}^n(1) - \bar{\varphi}^n(0) \right) \quad \text{for } c \in \left(0, \frac{1}{2} \right). \tag{3.51}$$

In [**76**] sufficient conditions are given that there exists a sufficiently large penalty parameter $\bar{\mu} > 0$ such that

$$\bar{\varphi}^n(1) - \bar{\varphi}^n(0) < 0 \quad \text{for all } \mu \ge \bar{\mu}. \tag{3.52}$$

To find an appropriate value for μ we check whether (3.52) holds and increase μ if not and iterate. If (3.52) is fulfilled we determine α^n from (3.51) by a backtracking strategy starting with $\alpha^n = 1$. If (3.51) is violated we decrease α^n by setting $\alpha^n := \zeta\alpha^n$ with $\zeta \in (0, 1)$ and iterate until (3.51) holds or α^n falls below a minimal step size parameter.

3.4 Numerical experiments

In this section we report the results of our numerical experiments. The algorithm proposed in Section 3.3 is discretized for the numerical realization. We make use of a finite element method with piecewise linear finite elements. To solve (3.49) we apply a preconditioned GMRES method.

A typical amount of data noise in option prices, that can be caused for example by bid–ask spreads, is $\delta = 0.1\%$. For the choice of the regularization parameter β we follow here the strategy proposed in [**51**], and define a decreasing sequence of admissible regularization parameters by

$$(\beta_1, \ldots, \beta_6) = (10^2, 10^1, \ldots, 10^{-3})\delta = (10^{-1}, 10^{-2}, \ldots, 10^{-6}).$$

We start minimizing the functional with the highest value $\beta = \beta_1$ and subsequently decrease the regularization parameter, starting the method at the minimizers obtained in the previous step. For the SQP method we choose stopping tolerances $\varepsilon_{\mathsf{abs}} = 10^{-6}$ and $\varepsilon_{\mathsf{rel}} = 10^{-3}$ and a maximal number of iterations $n_{\mathsf{sqp}} = 20$. We use a non–uniform grid with 140 nodes with local refinement around $x = S_0$ for the spatial discretization. In time, we employ a fixed, non–equidistant grid consisting of 35 points with small time steps close to $t = 0$.

As first example we apply our method to an artificial data set of Black–Scholes prices, i.e. prices computed with the Black–Scholes formula, with $S_0 = 100$, $r = 0$, one month to maturity and constant volatility $\sigma = 0.15$. We consider four different cases with *a priori* guess $q_d = \frac{1}{2}\sigma_d^2 x^2$. In our first simulation we use the 'good' *a priori* guess $\sigma_d = 0.16$, in a second we add 0.1% uniformly distributed noise. We compare these results to the ones from a third and fourth run using a 'bad' *a priori* guess $\sigma_d = 0.1$, where in the fourth run again we added 0.1% uniformly distributed noise.

The resulting option prices are given in Table 3.1 and the corresponding volatilities are shown in Table 3.2. The error–free true values for the option prices were computed using the Black–Scholes formula. Table 3.3 displays the residuals $\|u(T)-u_T\|$ remaining after reconstruction. In all four runs, the identified option prices correspond very well to the true values, the difference is neglectably small. The corresponding volatilities are identified well, with small differences remaining due to discretization errors, which can be reduced by using a finer grid. This can be seen from the results of a fifth run which was executed on a grid with halved mesh size in space and time.

Strike E	95	97.5	100	102.5	105
True value	5.24433	3.23921	1.72734	0.77577	0.28866
Good guess	5.24429	3.23921	1.72734	0.77576	0.28861
Good guess & noise	5.24430	3.23921	1.72734	0.77576	0.28861
Bad guess	5.24435	3.23922	1.72733	0.77578	0.28866
Bad guess & noise	5.24435	3.23922	1.72733	0.77578	0.28866
Good guess, fine grid	5.24433	3.23921	1.72734	0.77577	0.28866

Table 3.1: True option price and reconstructed option prices computed for different strikes E and different a priori guesses q_d. The true values were computed using the Black–Scholes formula with constant volatility $\sigma = 0.15$.

Strike E	95	97.5	100	102.5	105
True value	0.1500	0.1500	0.1500	0.1500	0.1500
Good guess	0.1454	0.1500	0.1517	0.1506	0.1470
Good guess & noise	0.1457	0.1500	0.1517	0.1506	0.1470
Bad guess	0.1458	0.1500	0.1517	0.1506	0.1472
Bad guess & noise	0.1460	0.1500	0.1517	0.1506	0.1472
Good guess, fine grid	0.1488	0.1500	0.1508	0.1502	0.1494

Table 3.2: Reconstructed volatilities for different strikes E and different *a priori* guesses q_d.

	$\|\|u(T) - u_T\|\|$
Good guess	6.76×10^{-3}
Good guess & noise	3.56×10^{-2}
Bad guess	6.50×10^{-3}
Bad guess & noise	3.47×10^{-2}
Good guess, fine grid	1.51×10^{-3}

Table 3.3: Residuals remaining after reconstruction.

Strike E	Price C
5825	469.5
6175	223.5
6225	195.5
6275	169
6325	144.5
6575	56.5
6875	10.5
7225	0.5

Table 3.4: Call option prices for maturity 0.09589.

Overall, the method shows only very small dependence on the chosen *a priori* guess and is robust regarding to additional data noise. In all runs, Algorithm 3.21 needs very few iterations, typically one to three, to meet the prescribed stopping tolerances. We find that the identification process is very stable and the option prices and associated volatilities can be recovered well.

In our second example we use market data from [**41**]. These data involve FTSE index call option prices from February 11, 2000. The option prices are given in Table 3.4. The spot price is $S_0 = 6219$, the constant interest rate is $r = 0.061451$, and the maturity is 0.095890.

Note that meaningful empirical data, i.e. prices of options that are actually traded, are usually only available for strike prices E in a small region around the spot price S_0, typically $S_0 \pm 10\%$ or even only $S_0 \pm 5\%$ [**86**]. Restricting the computational domain to this small region is not advisable, hence one needs to extrapolate the data out of this region. Here, we compute the implied volatilities of the options with the largest and the smallest strike price available by inverting the Black–Scholes formula, and use prices computed by the Black–Scholes formula with these volatilities in the regions, where no market data are available. The data are interpolated using a cubic spline.

The resulting local volatility function is shown in Figure 3.1. It is skewed, as is typical for equity index options. The volatility function is higher for options in–the–money, i.e. for options with $E < S_0$ than for options at–the–money and out–of–the–money. Furthermore, it shows a term structure, with volatility decreasing as time approaches maturity. These characteristics are

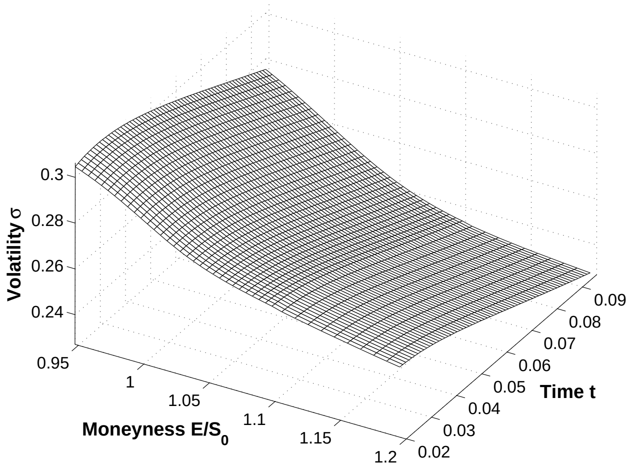

Figure 3.1: Local volatility function computed from market data.

consistent with empirically observed patterns in equity index options [**87**].

Chapter 4

Optimal Portfolio in Incomplete Markets

4.1 Introduction

The main aim of this chapter is to prove the existence and uniqueness of generalized Sobolev solutions to the initial–boundary–value problem (1.30) and to the Cauchy problem (1.30a), (1.30c) in $\hat{\Omega} = \mathbb{R}^d \times \mathbb{R}^{d'}$. To our knowledge, the well–posedness of problem (1.30) has not been considered in the mathematical literature up to now.

The main mathematical difficulty is the treatment of the terms with the quadratic gradients. In order to deal with this problem, several approaches can be found in the mathematical literature. The first one uses a sign condition of the form $f(u, \nabla u)u \geq 0$ (where $f(u, \cdot)$ is a function with quadratic growth) when deriving a priori H^1 bounds [**15, 21, 22, 126**]. This idea cannot be used here. The second idea is to study the connections between backward stochastic differential equations and partial differential equations with quadratic gradient terms similar to (1.30) and to consider viscosity solutions [**95**]. However, the results presented here are not covered by those in [**95**], as we consider nonlinear covariance matrices. A third idea consists in first analyzing an approximate problem (with, for instance, linearly growing gradient terms) and in establishing L^∞ bounds independent of the approximation parameter, which leads to a priori H^1 estimates [**23, 24, 25, 26, 55, 56, 67, 112**]. We use this technique since the equation (1.30a) allows for a maximum principle in the sense of Stampacchia. In particular we consider *generalized* (weak) solutions. Then uniform H^1 bounds are derived using nonlinear test functions of the type $\sinh(\lambda u)$ for sufficiently large $\lambda > 0$. The uniform H^1 bounds only imply *weak* convergence in H^1 of the sequence of approximating solutions. However, the quasilinear structure of the problem requires that the sequence converges *strongly* in H^1. This is achieved by employing the monotonicity method of Frehse [**59**], originally used for *elliptic* problems, which we extend to *parabolic* equations. With these techniques we prove our first main result, the existence of solutions to the initial–boundary–value problem (1.30) (Section 4.2) and to the Cauchy problem (1.30a), (1.30c), which is the original formulation in [**107**] (Section 4.3).

Notice that the sign of the quadratic gradient terms depends on whether $p < 1$ or $p > 1$ but our proof does not depend on the sign of $p - 1$ and holds for arbitrary values of p (except $p = 0$ and $p = 1$). Furthermore, we mention that our results can be extended to equations fulfilling more general structure conditions, but the emphasis here lies on studying the particular problem (1.30).

Our second main result is a proof of the uniqueness of generalized solutions to (1.30). The uniqueness proof has to overcome the difficulties arising from both the quadratic gradient terms and the quasilinearity in the elliptic operator. In order to deal with the quadratic gradients, the uniqueness of solutions is often shown in the space of functions whose gradient lies in a smaller space than L^2 (for instance in L^∞) [**35, 147**]. Quasilinear terms can be handled, for instance, using duality methods [**3**]. However, there are much less uniqueness results (and techniques) for problems with *both* difficulties in the context of Sobolev solutions. We are only aware of the paper of Barles and Murat [**8**], where the uniqueness of weak solutions to general elliptic problems is proved under a structure condition on the nonlinearities. We adapt their method in order to show the uniqueness of generalized solutions to (1.30) either if the covariance matrices C and C' do *not* depend on S and S', respectively, or if $p > 1$ and some (smallness) conditions on the derivatives of C and C' with respect to u are satisfied (Section 4.4). Notice that we do *not* need additional regularity assumptions on the solution.

Finally, we give a few details on the background of the model and present some numerical results by solving problem (1.30) with a finite element method for two risky assets and one state variable (Section 4.5). The experiments are showing that the optimal value function varies only slowly with respect to the state variable.

4.2 Existence of solutions

In this section we prove the existence of (generalized) solutions to (1.30). Let $Q_T = \hat{\Omega} \times (0,T)$. We call u a (generalized) solution of (1.30) if $u - u_D \in L^2(0,T;H_0^1(\hat{\Omega}))$, $u \in H^1(0,T;H^{-1}(\hat{\Omega}))$, u fulfills the initial condition (1.30c) in the sense of $L^2(\hat{\Omega})$ and

$$\begin{aligned}
&\int_0^T \langle u_t, \phi\rangle \, dt + \frac{1}{2}\int_{Q_T} (\nabla\phi)^\top C(u)\nabla u \, dx\, dt + \frac{1}{2}\int_{Q_T} (\nabla'\phi)^\top C'(u)\nabla' u \, dx\, dt \\
= &\int_{Q_T} (\mu\cdot\nabla u + \mu'\cdot\nabla' u - q(\mu - rS)\cdot\nabla u - \frac{q}{2}\beta(u)^2 + pr)\phi \, dx\, dt \qquad (4.1)\\
&- \frac{1}{2(p-1)}\int_{Q_T} (\nabla u)^\top C(u)\nabla u\phi \, dx\, dt + \frac{1}{2}\int_{Q_T} (\nabla' u)^\top C'(u)\nabla' u\phi \, dx\, dt \\
&- \frac{1}{2}\int_{Q_T} ((\operatorname{div} C)(u)\cdot\nabla u + (\operatorname{div}' C')(u)\cdot\nabla' u)\phi \, dx\, dt
\end{aligned}$$

holds for any $\phi \in L^\infty(Q_T) \cap L^2(0,T;H_0^1(\hat{\Omega}))$. Here, $u_t = \partial_t u$, $(\operatorname{div} C)(u)$ denotes the vector with components $((\operatorname{div} C)(u))_j = \sum_{i=1}^d \partial c_{ij}(u)/\partial S_i$ (analogously for $\operatorname{div}' C'(u)$), $\langle\cdot,\cdot\rangle$ is the dual product between $H^{-1}(\hat{\Omega})$ and $H_0^1(\hat{\Omega})$ and dx is an abbreviation for $dS\, dS'$. The notion of solution for the whole–space problem is analogous.

The basic hypotheses for the initial–boundary–value problem are as follows:

(H1) Domain: $\hat{\Omega} = \Omega \times \Omega' \subset \mathbb{R}^d \times \mathbb{R}^{d'}$ is a bounded domain with boundary $\partial\hat{\Omega} \in C^1$, $d \geq 1$, $d' \geq 0$.

(H2) Coercivity: $\exists \alpha, \alpha' > 0 : \forall \xi \in \mathbb{R}^d \backslash \{0\}, \xi' \in \mathbb{R}^{d'} \backslash \{0\} : \forall\, S, S', t, u :$

$$\xi^\top C(S,S',t,u)\xi \geq \alpha|\xi|^2 \quad \text{and} \quad \xi'^\top C'(S,S',t,u)\xi' \geq \alpha'|\xi'|^2.$$

(H3) Symmetry: $c_{ij} = c_{ji}$ for all $i,j \in \{1,\ldots,d\}$ and $c'_{ij} = c'_{ji}$ for all $i,j \in \{1,\ldots,d'\}$.

(H4) Data:

(i) $C(\cdot,\cdot,\cdot,u)$, $C(\cdot,\cdot,\cdot,u) \in L^\infty(0,T;W^{1,\infty}(\hat\Omega))$ for all $u \in \mathbb{R}$ and $C(S,S',t,\cdot)$, $C'(S,S',t,\cdot) \in C^1(\mathbb{R}) \cap W^{1,\infty}(\mathbb{R})$ for all S,S',t; $p \in \mathbb{R}\backslash\{0,1\}$.

(ii) $\mu \in L^\infty(0,T;L^\infty(\hat\Omega))$, $\mu' \in L^\infty(0,T;L^\infty(\hat\Omega))$, $r \in L^\infty(0,T;L^\infty(\hat\Omega))$.

(iii) $u_D \in L^2(0,T;H^2(\hat\Omega)) \cap L^\infty(0,T;L^\infty(\hat\Omega)) \cap H^1(0,T;L^1(\hat\Omega))$, $u_0 \in L^\infty(\hat\Omega) \cap H^1(\hat\Omega)$.

First we prove that there exists a solution of a truncated approximate problem. Define $s_K = \max(-K_2, \min(s, K_1))$ for $s \in \mathbb{R}$, where

$$K_1 = K_1(t) = (t+1)\overline{M}, \quad K_2 = K_2(t) = (t+1)\underline{M}$$

and

$$\begin{aligned}
\overline{M} &= \max\{\sup_{\hat\Omega} u_0, \sup_{\partial\hat\Omega\times(0,T)} u_D, M_2(r,\beta,p)\},\\
\underline{M} &= \min\{\inf_{\hat\Omega} u_0, \inf_{\partial\hat\Omega\times(0,T)} u_D, M_1(r,\beta,p)\},
\end{aligned}$$

with

$$\begin{aligned}
M_1(r,\beta,p) &= -\sup_{S,S',t,u}\Big(\frac{q}{2}\beta(S,S',t,u)^2 - pr(S,S',t)\Big),\\
M_2(r,\beta,p) &= -\inf_{S,S',t,u}\Big(\frac{q}{2}\beta(S,S',t,u)^2 - pr(S,S',t)\Big).
\end{aligned}$$

Consider the approximate problem

$$\begin{aligned}
&\int_0^T \langle u_t^\varepsilon, \phi\rangle\, dt + \frac{1}{2}\int_{Q_T} (\nabla\phi)^\top C(u^\varepsilon)\nabla u^\varepsilon\, dx\, dt + \frac{1}{2}\int_{Q_T} (\nabla'\phi)^\top C'(u^\varepsilon)\nabla' u^\varepsilon\, dx\, dt\\
&= \int_{Q_T} (\mu\cdot\nabla u^\varepsilon + \mu'\cdot\nabla' u^\varepsilon - q(\mu - rS)\cdot\nabla u^\varepsilon - \frac{q}{2}\beta(u^\varepsilon)^2 + pr)\phi\, dx\, dt\\
&\quad - \frac{1}{2(p-1)}\int_{Q_T} \frac{(\nabla u^\varepsilon)^\top C(u^\varepsilon)\nabla u_K^\varepsilon}{1+\varepsilon(\nabla u^\varepsilon)^\top C(u^\varepsilon)\nabla u^\varepsilon}\phi\, dx\, dt\\
&\quad + \frac{1}{2}\int_{Q_T} \frac{(\nabla' u^\varepsilon)^\top C'(u^\varepsilon)\nabla' u_K^\varepsilon}{1+\varepsilon(\nabla' u^\varepsilon)^\top C'(u^\varepsilon)\nabla' u^\varepsilon}\phi\, dx\, dt
\end{aligned}$$

$$- \frac{1}{2} \int\limits_{Q_T} ((\operatorname{div} C)(u^\varepsilon) \cdot \nabla u^\varepsilon + (\operatorname{div}' C')(u^\varepsilon) \cdot \nabla' u^\varepsilon) \phi \, dx \, dt \tag{4.2}$$

for any $\phi \in L^2(0,T; H_0^1(\hat{\Omega}))$ and $\varepsilon > 0$ subject to boundary and initial conditions (1.30b), (1.30c).

Lemma 4.1. *There exists a solution u^ε of* (4.2), (1.30b), (1.30c) *such that $u^\varepsilon - u_D \in L^2(0,T; H_0^1(\hat{\Omega}))$ and $u^\varepsilon \in L^2(0,T; H^2(\hat{\Omega})) \cap H^1(0,T; L^2(\hat{\Omega}))$.*

Proof. We use a fixed point argument. For given $w \in L^2(0,T; H^1(\hat{\Omega}))$ we consider the linear equation

$$\begin{aligned}
&\int\limits_0^T \langle u_t^\varepsilon, \phi \rangle \, dt + \frac{1}{2} \int\limits_{Q_T} (\nabla \phi)^\top C(w) \nabla u^\varepsilon \, dx \, dt + \frac{1}{2} \int\limits_{Q_T} (\nabla' \phi)^\top C'(w) \nabla' u^\varepsilon \, dx \, dt \\
&= \int\limits_{Q_T} (\mu \cdot \nabla u^\varepsilon + \mu' \cdot \nabla' u^\varepsilon - q(\mu - rS) \cdot \nabla u^\varepsilon - \frac{q}{2} \beta(w)^2 + pr) \phi \, dx \, dt \\
&\quad - \frac{1}{2(p-1)} \int\limits_{Q_T} \frac{(\nabla w)^\top C(w) \nabla w_K}{1 + \varepsilon (\nabla w)^\top C(w) \nabla w} \phi \, dx \, dt \\
&\quad + \frac{1}{2} \int\limits_{Q_T} \frac{(\nabla' w)^\top C'(w) \nabla' w_K}{1 + \varepsilon (\nabla' w)^\top C'(w) \nabla' w} \phi \, dx \, dt \\
&\quad - \frac{1}{2} \int\limits_{Q_T} ((\operatorname{div} C)(w) \cdot \nabla u^\varepsilon + (\operatorname{div}' C')(w) \cdot \nabla' u^\varepsilon) \phi \, dx \, dt
\end{aligned} \tag{4.3}$$

for any $\phi \in L^2(0,T; H_0^1(\hat{\Omega}))$ subject to the boundary and initial conditions (1.30b), (1.30c).

Since

$$0 \le \frac{(\nabla w)^\top C(w) \nabla w_K}{1 + \varepsilon (\nabla w)^\top C(w) \nabla w} \le \frac{1}{\varepsilon}, \quad 0 \le \frac{(\nabla' w)^\top C'(w) \nabla' w_K}{1 + \varepsilon (\nabla' w)^\top C'(w) \nabla' w} \le \frac{1}{\varepsilon}, \tag{4.4}$$

(4.3) is a linear parabolic equation with bounded coefficients and bounded inhomogeneity. By standard results [**54**], (4.3) admits a unique solution u^ε such that $u^\varepsilon - u_D \in L^2(0,T; H_0^1(\hat{\Omega}))$, $u^\varepsilon \in L^2(0,T; H^2(\hat{\Omega})) \cap H^1(0,T; L^2(\hat{\Omega}))$. Thus the fixed point operator

$$S : L^2(0,T; H^1(\hat{\Omega})) \to L^2(0,T; H^1(\hat{\Omega})), \quad w \mapsto u^\varepsilon,$$

is well defined and $S(L^2(0,T;H^1(\hat\Omega))) \subset L^2(0,T;H^2(\hat\Omega)) \cap H^1(0,T;L^2(\hat\Omega))$. The following estimate holds [**54**]

$$\|u^\varepsilon\|_{L^2(0,T;H^2(\hat\Omega))} + \|u^\varepsilon\|_{L^\infty(0,T;H^1(\hat\Omega))} + \|u^\varepsilon_t\|_{L^2(0,T;L^2(\hat\Omega))} \leq c,$$

where in general $c > 0$ is a generic constant depending on ε, the data and on the inhomogeneity. Here, in fact, the inhomogeneity is bounded independently of w. Thus c only depends on ε and the data, but not on w. Standard arguments show that S is continuous. In view of the compact embedding $L^2(0,T;H^2(\hat\Omega)) \cap H^1(0,T;L^2(\hat\Omega))) \subset L^2(0,T;H^1(\hat\Omega))$ [**132**], S is compact in $L^2(0,T;H^1(\hat\Omega))$. The hypotheses for Schauder's fixed point theorem are fulfilled and (4.2), (1.30b), (1.30c) admits at least one solution u^ε. □

The existence proof for the original problem is based on the following uniform a priori estimates.

Lemma 4.2. *Let u^ε be a generalized solution to* (4.2), (1.30b), (1.30c) *in* $(0,T)$. *Then there exist constants* $\overline{K}, \underline{K} > 0$ *(independent of ε) such that*

$$\underline{K} \leq u^\varepsilon \leq \overline{K},$$

where $\underline{K} = \min_{0\leq t\leq T} K_2(t)$, $\overline{K} = \max_{0\leq t\leq T} K_1(t)$.

Remark 4.3. The sign of one of the quadratic terms depends on whether $p < 1$ or $p > 1$. Without truncation in the quadratic terms it is easy to obtain upper or lower L^∞ estimates for $p < 1$ or $p > 1$, respectively, using standard test functions, but it is not possible to obtain the missing lower (upper) estimate in this way. Our proof does not rely on the sign of $p - 1$, since by truncating the solution and choosing appropriate test functions, these terms vanish completely.

Proof. Let $\varphi(u^\varepsilon) = u^\varepsilon - K_1(t)$. Using $\varphi(u^\varepsilon)^+ = \max(0, \varphi(u^\varepsilon)) \in L^2(0,T;H^1_0(\hat\Omega))$ as test function in (4.2) yields, in view of $\nabla u^\varepsilon_K \varphi(u^\varepsilon)^+ \equiv 0$,

$$\frac{1}{2}\int_{\hat\Omega} (\varphi(u^\varepsilon)^+(t)^2 - \underbrace{\varphi(u^\varepsilon_0)^{+2}}_{=0})\,dx + \frac{1}{2}\int_{Q_T} (\nabla\varphi(u^\varepsilon)^+)^\top C(u^\varepsilon)\nabla u^\varepsilon\,dx\,dt$$
$$+\frac{1}{2}\int_{Q_T} (\nabla'\varphi(u^\varepsilon)^+)^\top C'(u^\varepsilon)\nabla' u^\varepsilon\,dx\,dt$$
$$= \int_{Q_T} (\mu\cdot\nabla u^\varepsilon + \mu'\cdot\nabla' u^\varepsilon - q(\mu - rS)\cdot\nabla u^\varepsilon - \underbrace{(\frac{q}{2}\beta(u^\varepsilon)^2 - pr + \overline{M}))}_{\geq 0}\,\varphi(u^\varepsilon)^+\,dx\,dt$$

$$
\begin{aligned}
&\quad - \frac{1}{2} \int\limits_{Q_T} ((\operatorname{div} C)(u^\varepsilon) \cdot \nabla u^\varepsilon + (\operatorname{div}' C')(u^\varepsilon) \cdot \nabla' u^\varepsilon)\, \varphi(u^\varepsilon)^+ \, dx\, dt \\
&\le \int\limits_{Q_t} (\mu \cdot \nabla \varphi(u^\varepsilon)^+ + \mu' \cdot \nabla' \varphi(u^\varepsilon)^+ - q(\mu - rS) \cdot \nabla \varphi(u^\varepsilon)^+)\, \varphi(u^\varepsilon)^+ \, dx\, dt \\
&\quad - \frac{1}{2} \int\limits_{Q_T} ((\operatorname{div} C)(u^\varepsilon) \cdot \nabla \varphi(u^\varepsilon)^+ + (\operatorname{div}' C')(u^\varepsilon) \cdot \nabla' \varphi(u^\varepsilon)^+)\, \varphi(u^\varepsilon)^+ \, dx\, dt \\
&=: I.
\end{aligned} \tag{4.5}
$$

We use Young's inequality and (H4) to estimate the right hand side:

$$
I \le \int\limits_{Q_T} (\delta |\nabla \varphi(u^\varepsilon)^+|^2 + \delta |\nabla' \varphi(u^\varepsilon)^+|^2 + \frac{c}{\delta} (\varphi(u^\varepsilon)^+)^2)\, dx\, dt,
$$

where $\delta > 0$, and $c > 0$ is a constant independent of ε and varying in the following from occurrence to occurrence.

We employ the coercivity (H2) of C and C' to estimate the left hand side of (4.5) from below. Then the gradient terms on the right hand side can be controlled, for sufficiently small $\delta > 0$, by the left hand side. More precisely, we obtain

$$
\begin{aligned}
&\frac{1}{2} \int\limits_{\hat\Omega} (\varphi(u^\varepsilon)^+(t))^2 \, dx + \frac{1}{2} \int\limits_{Q_T} \underbrace{(\alpha - 2\delta)}_{\ge 0} |\nabla \varphi(u^\varepsilon)^+|^2 \, dx\, dt \\
&\qquad + \frac{1}{2} \int\limits_{Q_T} \underbrace{(\alpha' - 2\delta))}_{\ge 0} |\nabla' \varphi(u^\varepsilon)^+|^2 \, dx\, dt \le \frac{2c}{\delta} \int\limits_{Q_T} (\varphi(u^\varepsilon)^+)^2 \, dx,
\end{aligned}
$$

and applying Gronwall's lemma yields $u^\varepsilon \le K_1 \le \overline{K}$ a.e. in $\hat\Omega \times (0, T)$.

In order to derive the lower bound, set $\varphi(u^\varepsilon) = u^\varepsilon - K_2$. Using $\varphi(u^\varepsilon)^- = \min(0, \varphi(u^\varepsilon)) \in L^2(0, T; H_0^1(\hat\Omega))$ as test function in (4.2) yields

$$
\begin{aligned}
&\frac{1}{2} \int\limits_{\hat\Omega} ((\varphi(u^\varepsilon)^-(t))^2 - \underbrace{\varphi(u_0^\varepsilon)^{-2}}_{=0})\, dx + \frac{1}{2} \int\limits_{Q_T} (\nabla \varphi(u^\varepsilon)^-)^\top C(u^\varepsilon) \nabla u^\varepsilon \, dx\, dt \\
&+ \frac{1}{2} \int\limits_{Q_T} (\nabla' \varphi(u^\varepsilon)^-)^\top C'(u^\varepsilon) \nabla' u^\varepsilon \, dx\, dt
\end{aligned}
$$

$$= \int_{Q_T} \Big(\mu \cdot \nabla u^\varepsilon + \mu' \cdot \nabla' u^\varepsilon - q(\mu - rS) \cdot \nabla u^\varepsilon - \underbrace{(\frac{q}{2}\beta(u^\varepsilon)^2 - pr + \underline{M})}_{\leq 0}\Big) \varphi(u^\varepsilon)^- \, dx\, dt$$

$$- \frac{1}{2} \int_{Q_T} \big((\operatorname{div} C)(u^\varepsilon) \cdot \nabla u^\varepsilon + (\operatorname{div}' C')(u^\varepsilon) \cdot \nabla' u^\varepsilon\big)\, \varphi(u^\varepsilon)^- \, dx\, dt$$

$$\leq \int_{Q_T} \big(\mu \cdot \nabla \varphi(u^\varepsilon)^- + \mu' \cdot \nabla' \varphi(u^\varepsilon)^- - q(\mu - rS) \cdot \nabla \varphi(u^\varepsilon)^-\big)\, \varphi(u^\varepsilon)^- \, dx\, dt$$

$$- \frac{1}{2} \int_{Q_T} \big((\operatorname{div} C)(u^\varepsilon) \cdot \nabla \varphi(u^\varepsilon)^- + (\operatorname{div}' C')(u^\varepsilon) \cdot \nabla' \varphi(u^\varepsilon)^-\big)\, \varphi(u^\varepsilon)^- \, dx\, dt.$$

We can estimate similarly as above and applying Gronwall's lemma yields $u^\varepsilon \geq K_2 \geq \underline{K}$ a.e. in $\hat{\Omega} \times (0, T)$. □

Lemma 4.4. *Let u^ε be a weak solution to* (4.2), (1.30b), (1.30c). *Then there exists a constant $k > 0$ (independent of ε) such that*

$$\|u^\varepsilon\|_{L^2(0,T;H^1(\hat{\Omega}))} \leq k.$$

Proof. Inspired by [**59**], we use $\sinh(\lambda u^\varepsilon) - \sinh(\lambda u_D)$, $\lambda > 0$, as test function in (4.2) to obtain

$$\int_0^T \langle u_t^\varepsilon, \sinh(\lambda u^\varepsilon) - \sinh(\lambda u_D)\rangle \, dt + \frac{1}{2} \int_{Q_T} \lambda \cosh(\lambda u^\varepsilon)(\nabla u^\varepsilon)^\top C(u^\varepsilon) \nabla u^\varepsilon \, dx\, dt$$

$$+ \frac{1}{2} \int_{Q_T} \lambda \cosh(\lambda u^\varepsilon)(\nabla' u^\varepsilon)^\top C'(u^\varepsilon) \nabla' u^\varepsilon \, dx\, dt$$

$$= \int_{Q_T} \Big(\mu \cdot \nabla u^\varepsilon + \mu' \cdot \nabla' u^\varepsilon - q(\mu - rS) \cdot \nabla u^\varepsilon - \frac{q}{2}\beta(u^\varepsilon)^2 + pr\Big)$$

$$\times \big(\sinh(\lambda u^\varepsilon) - \sinh(\lambda u_D)\big) \, dx\, dt$$

$$- \frac{1}{2(p-1)} \int_{Q_T} \frac{(\nabla u^\varepsilon)^\top C(u^\varepsilon) \nabla u^\varepsilon}{1 + \varepsilon(\nabla u^\varepsilon)^\top C(u^\varepsilon) \nabla u^\varepsilon} \big(\sinh(\lambda u^\varepsilon) - \sinh(\lambda u_D)\big) \, dx\, dt$$

$$+ \frac{1}{2} \int_{Q_T} \frac{(\nabla' u^\varepsilon)^\top C'(u^\varepsilon) \nabla' u^\varepsilon}{1 + \varepsilon(\nabla' u^\varepsilon)^\top C'(u^\varepsilon) \nabla' u^\varepsilon} \big(\sinh(\lambda u^\varepsilon) - \sinh(\lambda u_D)\big) \, dx\, dt$$

$$
\begin{aligned}
&- \frac{1}{2} \int\limits_{Q_T} ((\operatorname{div} C)(u^\varepsilon) \cdot \nabla u^\varepsilon + (\operatorname{div}' C')(u^\varepsilon) \cdot \nabla' u^\varepsilon) \\
&\qquad\qquad\qquad\qquad \times (\sinh(\lambda u^\varepsilon) - \sinh(\lambda u_D)) \, dx \, dt \\
&+ \frac{1}{2} \int\limits_{Q_T} \lambda \cosh(\lambda u_D) \big[(\nabla u_D)^\top C(u^\varepsilon) \nabla u^\varepsilon + (\nabla' u_D)^\top C'(u^\varepsilon) \nabla' u^\varepsilon \big] \, dx \, dt.
\end{aligned}
$$

Since u^ε is uniformly bounded in $L^\infty(Q_T)$ and $|\sinh(x)| \leq \cosh(x)$, $x \in \mathbb{R}$, we obtain

$$
\begin{aligned}
&\frac{1}{2} \int\limits_{Q_T} \lambda \cosh(\lambda u^\varepsilon) \Big[(\nabla u^\varepsilon)^\top C(u^\varepsilon) \nabla u^\varepsilon + (\nabla' u^\varepsilon)^\top C'(u^\varepsilon) \nabla' u^\varepsilon \Big] \, dx \, dt \\
&\leq \int\limits_{Q_T} |(\mu \cdot \nabla u^\varepsilon + \mu' \cdot \nabla' u^\varepsilon - q(\mu - rS) \cdot \nabla u^\varepsilon)(\sinh(\lambda u^\varepsilon) - \sinh(\lambda u_D))| \, dx \, dt \\
&\quad \underbrace{+ \int\limits_{Q_T} |\frac{q}{2} \beta(u^\varepsilon)^2 - pr| (\cosh(\lambda u^\varepsilon) + \cosh(\lambda u_D)) \, dx \, dt}_{\leq L_1} \\
&\quad + \frac{1}{2|p-1|} \int\limits_{Q_T} (\nabla u^\varepsilon)^\top C(u^\varepsilon) \nabla u^\varepsilon \, (\cosh(\lambda u^\varepsilon) + \cosh(\lambda u_D)) \, dx \, dt \\
&\quad + \frac{1}{2} \int\limits_{Q_T} (\nabla' u^\varepsilon)^\top C'(u^\varepsilon) \nabla' u^\varepsilon \, (\cosh(\lambda u^\varepsilon) + \cosh(\lambda u_D)) \, dx \, dt \\
&\quad + \frac{1}{2} \int\limits_{Q_T} |(\operatorname{div} C)(u^\varepsilon) \cdot \nabla u^\varepsilon + (\operatorname{div}' C')(u^\varepsilon) \cdot \nabla' u^\varepsilon| \\
&\qquad\qquad\qquad\qquad \times |\sinh(\lambda u^\varepsilon) - \sinh(\lambda u_D)| \, dx \, dt \\
&\quad + \frac{1}{2} \int\limits_{Q_T} \lambda \cosh(\lambda u_D) \Big[|(\nabla u_D)^\top C(u^\varepsilon) \nabla u^\varepsilon| + |(\nabla' u_D)^\top C'(u^\varepsilon) \nabla' u^\varepsilon| \Big] \, dx \, dt \\
&\quad \underbrace{+ \frac{1}{\lambda} \int\limits_{\hat{\Omega}} |\cosh(\lambda u^\varepsilon)(t) - \cosh(\lambda u_0)| \, dx}_{\leq L_2} + \underbrace{\int\limits_{Q_T} |u^\varepsilon \cosh(\lambda u_D) u_{D,t}| \, dx \, dt}_{\leq L_3}.
\end{aligned}
$$

Here we use the assumption that $u_D \in H^1(0, T; L^1(\hat{\Omega}))$. Choosing λ sufficiently large and using Young's inequality for some $\delta > 0$, we can further estimate

$$
\begin{aligned}
&\frac{1}{2}\int\limits_{Q_T} \underbrace{(\lambda\cosh(\lambda u^\varepsilon) - \frac{1}{|p-1|}(\cosh(\lambda u^\varepsilon) + \cosh(\lambda u_D)))}_{=:\kappa>0}(\nabla u^\varepsilon)^\top C(u^\varepsilon)\nabla u^\varepsilon \, dx\, dt \\
&\quad + \frac{1}{2}\int\limits_{Q_T} \underbrace{(\lambda\cosh(\lambda u^\varepsilon) - (\cosh(\lambda u^\varepsilon) + \cosh(\lambda u_D)))}_{=:\kappa'>0}(\nabla' u^\varepsilon)^\top C'(u^\varepsilon)\nabla' u^\varepsilon \, dx\, dt \\
\leq & L_1 + L_2 + L_3 + \int\limits_{Q_T} |\mu\cdot\nabla u^\varepsilon + \mu'\cdot\nabla' u^\varepsilon - q(\mu - rS)\cdot\nabla u^\varepsilon| \\
&\qquad\qquad\qquad\qquad \times |\sinh(\lambda u^\varepsilon) - \sinh(\lambda u_D)| \, dx\, dt \\
&\quad + \frac{1}{2}\int\limits_{Q_T} |(\operatorname{div} C)(u^\varepsilon)\cdot\nabla u^\varepsilon + (\operatorname{div}' C')(u^\varepsilon)\cdot\nabla' u^\varepsilon|(\cosh(\lambda u^\varepsilon) + \cosh(u_D)) \, dx\, dt \\
&\quad + \frac{1}{2}\int\limits_{Q_T} \lambda\cosh(\lambda u_D)\Big[\underbrace{|(\nabla u_D)^\top C(u^\varepsilon)\nabla u^\varepsilon|}_{\leq ||C||_2|\nabla u_D||\nabla u^\varepsilon|} + \underbrace{|(\nabla' u_D)^\top C'(u^\varepsilon)\nabla' u^\varepsilon|}_{\leq ||C'||_2|\nabla' u_D||\nabla' u^\varepsilon|}\Big] \, dx\, dt \\
\leq & L_1 + L_2 + L_3 + \int\limits_{Q_T} \big(\delta|\nabla u^\varepsilon|^2 + \frac{c}{\delta}(\cosh(\lambda u^\varepsilon) + \cosh(\lambda u_D))^2\big) \, dx\, dt \\
&\quad + \int\limits_{Q_T} \big(\delta|\nabla' u^\varepsilon|^2 + \frac{c}{\delta}(\cosh(\lambda u^\varepsilon) + \cosh(\lambda u_D))^2\big) \, dx\, dt \\
&\quad + \int\limits_{Q_T} \big(\delta(|\nabla u^\varepsilon|^2 + |\nabla' u^\varepsilon|^2) + \frac{1}{\delta}(|\operatorname{div} C(u^\varepsilon)|^2 + |\operatorname{div}' C'(u^\varepsilon)|^2)\big) \\
&\qquad\qquad\qquad\qquad \times (\cosh(\lambda u^\varepsilon) + \cosh(\lambda u_D)) \, dx\, dt \\
&\quad + \frac{1}{2}\int\limits_{Q_T} \lambda\cosh(\lambda u_D)\Big[||C||_2(\frac{1}{\delta}|\nabla u_D|^2 + \delta|\nabla u^\varepsilon|^2) \\
&\qquad\qquad\qquad\qquad + ||C'||_2(\frac{1}{\delta}|\nabla' u_D|^2\delta|\nabla' u^\varepsilon|^2)\Big] \, dx\, dt,
\end{aligned}
$$

where $\|\cdot\|_2$ denotes the matrix norm defined by $||C||_2 = \sup_{|x|=1}|Cx|$ and $|\cdot|$ is the euclidian norm. For sufficiently small $\delta > 0$ the gradient terms on the right hand side can now be estimated by the left hand side using the coercivity (H2) of C and C':

$$\frac{1}{2}\int\limits_{Q_T}\Big\{\alpha\kappa - \delta\big[2c + 2\cosh(\lambda u^\varepsilon) + 2\cosh(\lambda u_D)$$

$$+ \lambda||C||_2\cosh(\lambda u_D)\big]\Big\}|\nabla u^\varepsilon|^2\,dx\,dt$$

$$+\frac{1}{2}\int\limits_{Q_T}\Big\{\alpha'\kappa' - \delta\big[2c + 2\cosh(\lambda u^\varepsilon) + 2\cosh(\lambda u_D)$$

$$+ \lambda||C'||_2\cosh(\lambda u_D)\big]\Big\}|\nabla' u^\varepsilon|^2\,dx\,dt$$

$$\leq L_1 + L_2 + L_3 + \int\limits_{Q_T}\frac{2}{\delta}(\cosh(\lambda u^\varepsilon) + \cosh(\lambda u_D))^2\,dx\,dt$$

$$+\int\limits_{Q_T}\frac{1}{\delta}(|(\operatorname{div} C)(u^\varepsilon)|^2 + |(\operatorname{div}' C')(u^\varepsilon)|^2)(\cosh(\lambda u^\varepsilon) + \cosh(\lambda u_D))\,dx\,dt$$

$$+\frac{1}{2}\int\limits_{Q_T}\lambda\cosh(\lambda u_D)\left[\frac{1}{\delta}||C||_2\frac{1}{\delta}|\nabla u_D|^2 + \frac{1}{\delta}||C'||_2|\nabla' u_D|^2\right]\,dx\,dt,$$

By Lemma 4.2, the right hand side is bounded and we conclude

$$\int\limits_{Q_T}(|\nabla u^\varepsilon|^2 + |\nabla' u^\varepsilon|^2)\,dx\,dt \leq k.$$

Due to the Poincaré inequality we obtain the desired H^1–bound. □

The main result of this section is the following theorem.

Theorem 4.5. *Assume hypotheses (H1)–(H4) hold. Then there exists a solution u of* (1.30) *such that $u - u_D \in L^\infty(0,T;L^\infty(\hat\Omega)) \cap L^2(0,T;H^1_0(\hat\Omega))$ and $u \in H^1(0,T;H^{-1}(\hat\Omega))$.*

Proof. Let u^ε be a solution of (4.2), (1.30b), (1.30c). In view of Lemma 4.4, $\|u^\varepsilon\|_{L^2(0,T;H^1(\hat\Omega))}$ is uniformly bounded and we can extract a subsequence u^ε (not relabeled) such that, as $\varepsilon \to 0$,

$$u^\varepsilon \rightharpoonup u \quad \text{in } L^2(0,T;H^1(\hat\Omega)), \tag{4.6}$$

using, e.g., [**148**, Theorem 21.D]. Since also $\|u_t^\varepsilon\|_{L^2(0,T;H^{-1}(\hat\Omega))}$ is uniformly bounded, again for a subsequence which is not relabeled,

$$u_t^\varepsilon \rightharpoonup u_t \quad \text{in } L^2(0,T;H^{-1}(\hat\Omega)). \tag{4.7}$$

By Aubin's lemma [**132**] we obtain

$$u^\varepsilon \to u \quad \text{in } L^2(0,T;L^2(\hat\Omega)). \tag{4.8}$$

In order to pass to the limit as $\varepsilon \to 0$ in the quadratic gradient terms of the truncated approximate equation (4.2), we need the strong convergence of $u^\varepsilon \to u$ in $L^2(0,T;H^1(\hat\Omega))$. The proof of this result is the main step of the proof.

To establish the strong convergence of $u^\varepsilon \to u$ we use the so–called monotonicity method of Frehse [**59**], extended here to parabolic problems. Let $\bar u^\varepsilon = u^\varepsilon - u$ and choose $\sinh(\lambda \bar u^\varepsilon)$, $\lambda > 0$, as test function in the approximate problem (4.2):

$$\begin{aligned}
&\int_0^T \langle u_t^\varepsilon, \sinh(\lambda\bar u^\varepsilon)\rangle \, dt + \frac{1}{2}\int_{Q_T} \lambda\cosh(\lambda\bar u^\varepsilon)(\nabla\bar u^\varepsilon)^\top C(u^\varepsilon)\nabla u^\varepsilon \, dx\, dt \\
&\quad + \frac{1}{2}\int_{Q_T} \lambda\cosh(\lambda\bar u^\varepsilon)(\nabla'\bar u^\varepsilon)^\top C'(u^\varepsilon)\nabla' u^\varepsilon \, dx\, dt \\
&= \int_{Q_T} \Big(\mu\cdot\nabla u^\varepsilon + \mu'\cdot\nabla' u^\varepsilon - q(\mu - rS)\cdot\nabla u^\varepsilon - \frac{q}{2}\beta(u^\varepsilon)^2 + pr\Big)\sinh(\lambda\bar u^\varepsilon)\, dx\, dt \\
&\quad - \frac{1}{2(p-1)}\int_{Q_T} \frac{(\nabla u^\varepsilon)^\top C(u^\varepsilon)\nabla u^\varepsilon}{1+\varepsilon(\nabla u^\varepsilon)^\top C(u^\varepsilon)\nabla u^\varepsilon}\sinh(\lambda\bar u^\varepsilon)\, dx\, dt \\
&\quad + \frac{1}{2}\int_{Q_T} \frac{(\nabla' u^\varepsilon)^\top C'(u^\varepsilon)\nabla' u^\varepsilon}{1+\varepsilon(\nabla' u^\varepsilon)^\top C'(u^\varepsilon)\nabla' u^\varepsilon}\sinh(\lambda\bar u^\varepsilon)\, dx\, dt \\
&\quad - \frac{1}{2}\int_{Q_T} \big((\operatorname{div} C)(u^\varepsilon)\cdot\nabla u^\varepsilon + (\operatorname{div}' C')(u^\varepsilon)\cdot\nabla' u^\varepsilon\big)\sinh(\lambda\bar u^\varepsilon)\, dx\, dt.
\end{aligned} \tag{4.9}$$

The left hand side of this equation can be written as follows:

$$
\begin{aligned}
&\int_0^T \langle \bar{u}^\varepsilon_t, \sinh(\lambda \bar{u}^\varepsilon) \rangle \; dt + \int_0^T \langle u_t, \sinh(\lambda \bar{u}^\varepsilon) \rangle \; dt \\
&\quad + \frac{1}{2} \int_{Q_T} \lambda \cosh(\lambda \bar{u}^\varepsilon) (\nabla \bar{u}^\varepsilon)^\top C(u^\varepsilon) \nabla \bar{u}^\varepsilon \; dx \, dt \\
&\quad + \frac{1}{2} \int_{Q_T} \lambda \cosh(\lambda \bar{u}^\varepsilon) (\nabla' \bar{u}^\varepsilon)^\top C'(u^\varepsilon) \nabla' \bar{u}^\varepsilon \; dx \, dt \\
&\quad + \frac{1}{2} \int_{Q_T} \lambda \cosh(\lambda \bar{u}^\varepsilon) \Big[(\nabla \bar{u}^\varepsilon)^\top C(u^\varepsilon) \nabla u + (\nabla' \bar{u}^\varepsilon)^\top C'(u^\varepsilon) \nabla' u \Big] \; dx \, dt.
\end{aligned}
\tag{4.10}
$$

We claim that the first term is non–negative. Indeed, let $u^\delta \in C^1([0,T]; H^1(\hat{\Omega}))$ be a sequence such that $u^\delta \to u$ in $L^2(0,T; H^1(\hat{\Omega})) \cap H^1(0,T; H^{-1}(\hat{\Omega}))$ as $\delta \to 0$ and $u^\delta(0) = u_0$. Then

$$
\begin{aligned}
&\int_{Q_T} (u^\varepsilon - u^\delta)_t \sinh(\lambda (u^\varepsilon - u^\delta)) \, dt \\
=& \frac{1}{\lambda} \int_{\hat{\Omega}} [\cosh(\lambda (u^\varepsilon - u^\delta)(T)) - \cosh(\lambda (u^\varepsilon - u^\delta)(0))] \, dx \\
=& \frac{1}{\lambda} \int_{\hat{\Omega}} (\cosh(\lambda (u^\varepsilon - u^\delta)(T)) - 1) \; dx \geq 0,
\end{aligned}
$$

and letting $\delta \to 0$ shows that

$$
\int_0^T \langle \bar{u}^\varepsilon_t, \sinh(\lambda \bar{u}^\varepsilon) \rangle \geq 0.
$$

The quadratic gradient terms on the right hand side of (4.9) can be estimated as

$$
\begin{aligned}
\frac{(\nabla u^\varepsilon)^\top C(u^\varepsilon) \nabla u^\varepsilon}{1 + \varepsilon (\nabla u^\varepsilon)^\top C(u^\varepsilon) \nabla u^\varepsilon} \leq & (\nabla \bar{u}^\varepsilon)^\top C(u^\varepsilon) \nabla \bar{u}^\varepsilon + (\nabla u)^\top C(u^\varepsilon) \nabla u \\
& + (\nabla \bar{u}^\varepsilon)^\top C(u^\varepsilon) \nabla u + (\nabla u)^\top C(u^\varepsilon) \nabla \bar{u}^\varepsilon
\end{aligned}
$$

and likewise for the ∇' terms. Taking the modulus and choosing λ sufficiently large, (4.9) and (4.10) become

$$
\begin{aligned}
&\frac{1}{2}\int\limits_{Q_T} (\lambda - \frac{1}{|p-1|}) \cosh(\lambda \bar{u}^\varepsilon)(\nabla \bar{u}^\varepsilon)^\top C(u^\varepsilon)\nabla \bar{u}^\varepsilon \, dx\, dt \\
&\quad + \frac{1}{2}\int\limits_{Q_T} (\lambda - 1) \cosh(\lambda \bar{u}^\varepsilon)(\nabla' \bar{u}^\varepsilon)^\top C'(u^\varepsilon)\nabla' \bar{u}^\varepsilon \, dx\, dt \\
&\le \int\limits_{Q_T} |(\mu \cdot \nabla \bar{u}^\varepsilon + \mu' \cdot \nabla' \bar{u}^\varepsilon - q(\mu - rS)\cdot \nabla \bar{u}^\varepsilon) \sinh(\lambda \bar{u}^\varepsilon)| \, dx\, dt \\
&\quad + \int\limits_{Q_T} |(\mu \cdot \nabla u + \mu' \cdot \nabla' u - q(\mu - rS)\cdot \nabla u - \frac{q}{2}\beta(u^\varepsilon)^2 + pr) \sinh(\lambda \bar{u}^\varepsilon)| \, dx\, dt \\
&\quad + \frac{1}{2|p-1|}\int\limits_{Q_T} |[(\nabla u)^\top C(u^\varepsilon)\nabla u^\varepsilon + (\nabla u^\varepsilon)^\top C(u^\varepsilon)\nabla u] \sinh(\lambda \bar{u}^\varepsilon)| \, dx\, dt \\
&\quad + \frac{1}{2}\int\limits_{Q_T} |[(\nabla' u)^\top C'(u^\varepsilon)\nabla' u^\varepsilon + (\nabla' u^\varepsilon)^\top C'(u^\varepsilon)\nabla' u] \sinh(\lambda \bar{u}^\varepsilon)| \, dx\, dt \\
&\quad + \frac{1}{2|p-1|}\int\limits_{Q_T} |(\nabla u)^\top C(u^\varepsilon)\nabla u \sinh(\lambda \bar{u}^\varepsilon)| \, dx\, dt \\
&\quad + \frac{1}{2}\int\limits_{Q_T} |(\nabla' u)^\top C'(u^\varepsilon)\nabla' u \sinh(\lambda \bar{u}^\varepsilon)| \, dx\, dt \\
&\quad + \frac{1}{2}\int\limits_{Q_T} |[(\operatorname{div} C)(u^\varepsilon)\cdot \nabla \bar{u}^\varepsilon + (\operatorname{div}' C')(u^\varepsilon)\cdot \nabla' \bar{u}^\varepsilon] \sinh(\lambda \bar{u}^\varepsilon)| \, dx\, dt \\
&\quad + \frac{1}{2}\int\limits_{Q_T} |[(\operatorname{div} C)(u^\varepsilon)\cdot \nabla u + (\operatorname{div}' C')(u^\varepsilon)\cdot \nabla' u] \sinh(\lambda \bar{u}^\varepsilon)| \, dx\, dt \\
&\quad + \frac{1}{2}\int\limits_{Q_T} \lambda \cosh(\lambda \bar{u}^\varepsilon)\Big[|(\nabla \bar{u}^\varepsilon)^\top C(u^\varepsilon)\nabla u| + |(\nabla' \bar{u}^\varepsilon)^\top C'(u^\varepsilon)\nabla' u|\Big] \, dx\, dt \\
&\quad + \int\limits_0^T |\langle u_t, \sinh(\lambda \bar{u}^\varepsilon)\rangle| \, dt \\
&=: I_1 + \cdots + I_{10}, \qquad (4.11)
\end{aligned}
$$

where we have used again $|\sinh(x)| \le \cosh(x),\ x \in \mathbb{R}$.

We need to show that the right hand side of (4.11) converges to zero as $\varepsilon \to 0$. In view of (4.8) and since $\bar{u}^\varepsilon$ is uniformly bounded in $L^\infty(Q_T)$, it holds

$$\sinh(\lambda\bar{u}^\varepsilon) \to 0 \text{ in } L^2(0,T;L^2(\hat{\Omega})) \quad \text{and} \quad \sinh(\lambda\bar{u}^\varepsilon) \rightharpoonup 0 \text{ in } L^2(0,T;H^1(\hat{\Omega})), \tag{4.12}$$

which implies that $I_2, I_5, I_6, I_8, I_{10} \to 0$ as $\varepsilon \to 0$. In view of (4.6) and (4.12), we obtain $I_1 \to 0$.

The treatment of the integrals I_3, I_4, I_7 and I_9 is more delicate. In view of (4.8) and since $\bar{u}^\varepsilon \in L^\infty(Q_T)$ uniformly, $\cosh(\lambda\bar{u}^\varepsilon) \to 1$ in $L^2(0,T;L^2(\hat{\Omega}))$ and a.e. in Q_T. Since $\nabla\bar{u}^\varepsilon$ is uniformly bounded in $L^2(0,T;L^2(\hat{\Omega}))$, it holds for a subsequence (not relabeled),

$$\nabla\cosh(\lambda\bar{u}^\varepsilon) \rightharpoonup \nabla z \quad \text{in } L^2(0,T;L^2(\hat{\Omega}))$$

for some z. From identifying $z = 1$ it follows

$$\nabla\cosh(\lambda\bar{u}^\varepsilon) \rightharpoonup 0 \quad \text{in } L^2(0,T;L^2(\hat{\Omega})).$$

Thus

$$\begin{aligned}
&\int_{Q_T} (\nabla u)^\top C(u^\varepsilon)\nabla u^\varepsilon \sinh(\lambda\bar{u}^\varepsilon)\,dx\,dt \\
&\qquad = \frac{1}{\lambda}\int_{Q_T} (\nabla u)^\top C(u^\varepsilon)\nabla\cosh(\lambda\bar{u}^\varepsilon)\,dx\,dt \\
&\qquad\quad + \int_{Q_T} (\nabla u)^\top C(u^\varepsilon)\nabla u \sinh(\lambda\bar{u}^\varepsilon)\,dx\,dt \to 0 \quad \text{as } \varepsilon \to 0.
\end{aligned}$$

All terms in I_3, I_4, I_7 and I_9 can be treated similarly showing that the right hand side of (4.11) converges to zero as $\varepsilon \to 0$.

Employing the coercivity (H2) of C, C' and choosing $\lambda > 0$ sufficiently large, we obtain

$$\lim_{\varepsilon\to 0}\int_{Q_T} \left(|\nabla\bar{u}^\varepsilon|^2 + |\nabla'\bar{u}^\varepsilon|^2\right)dx\,dt \le 0.$$

Thus we achieve

$$\nabla \bar{u}^\varepsilon \to 0,\ \nabla' \bar{u}^\varepsilon \to 0 \quad \text{in } L^2(0,T;L^2(\hat{\Omega})) \text{ as } \varepsilon \to 0,$$

which implies

$$u^\varepsilon \to u \quad \text{in } L^2(0,T;H^1(\hat{\Omega})) \text{ as } \varepsilon \to 0.$$

We can pass to the limit as $\varepsilon \to 0$ in (4.2) and conclude the existence of a solution u to problem (4.1). □

Remark 4.6. As the solution to (1.30) lies a posteriori in the space $L^\infty(Q_T)$, the regularity assumptions on the covariance matrices with respect to u can be relaxed. Indeed, by using a truncation argument by Stampacchia, it is not difficult to see that the hypothesis $C(S,S',t,\cdot), C'(S,S',t,\cdot) \in C^1(\mathbb{R})$ for all S, S', t is sufficient.

4.3 The Cauchy problem

We consider the Cauchy problem (1.30a), (1.30c) in $R_T = \mathbb{R}^d \times \mathbb{R}^{d'} \times (0,T)$. The L^∞ bound for the solutions to problem (1.30) of Section 4.2 depends on $\mu - rS$ which is not bounded if $S \in \mathbb{R}^d$. Therefore, we need the following assumption.

(H5) $\exists\, M > 0 :\ \sup_{(S,S',t)\in R_T} |\mu(S,S',t) - r(S,S',t)S| \le M.$

This assumption can be interpreted as follows: the relative return μ/S tends to the riskless interest rate r for large asset prices. This is known to be the case if the economic model consists of a representative investor with decreasing relative risk aversion or of multiple heterogeneous investors all of whom have constant relative risk aversion [**13**].

In the proof of Lemma 4.4 we made use of the Poincaré inequality to obtain the H^1 estimates. Since the Poincaré inequality cannot be used here, we lack an L^2 estimate for an H^1 estimate independent of $\hat{\Omega}$. The necessary estimate is provided by the following lemma.

Lemma 4.7. *Let (H1)–(H5) hold and let u be a weak solution to* (1.30) *such that $u_D = 0$. Then there exists a constant $L > 0$ (not depending on u) such that*

$$\|u\|_{L^\infty(0,T;L^p(\hat{\Omega}))} \le L \quad \forall\, p < \infty.$$

Proof. As $u \in L^\infty(Q_T)$ and the L^∞ bound is independent of $\hat{\Omega}$ (because of (H5)) it suffices to prove that

$$\|u\|_{L^\infty(0,T;L^1(\hat{\Omega}))} \leq c \tag{4.13}$$

for some $c > 0$, since then the result follows from interpolation. The idea of the proof of (4.13) is to use a smooth and monotone approximation of the sign function $\operatorname{sign}(u)$ as test function in the weak formulation of (1.30).

Let η be convex and smooth such that

$$\eta(0) = 0, \quad \eta'(0) = 0, \quad \eta(x) = |x| - 0.5 \quad \text{for } |x| \geq 1,$$

and define for $\delta > 0$

$$\eta_\delta(x) = \delta\eta\Big(\frac{x}{\delta}\Big), \quad x \in \mathbb{R}.$$

By construction of η_δ,

$$\eta_\delta(u) \leq |u| \quad \text{and} \quad \eta_\delta(u) \to |u| \quad \text{a.e. in } Q_T.$$

Using dominated convergence this implies

$$\eta_\delta(u) \to |u| \quad \text{in } L^2(0,T;L^1(\hat{\Omega})) \text{ as } \delta \to 0.$$

Use $\eta'_\delta(u)$ as test function in (4.1) to obtain

$$\begin{aligned}
&\int_0^T \langle u_t, \eta'_\delta(u)\rangle \, d\tau + \frac{1}{2}\int_{Q_T} \underbrace{[\eta''_\delta(u)(\nabla u)^\top C(u)\nabla u + \eta''_\delta(u)(\nabla' u)^\top C'(u)\nabla' u]}_{\geq 0} \, dx\, dt \\
&= \int_{Q_T} (\mu\cdot\nabla u + \mu'\cdot\nabla' u - q(\mu - rS)\cdot\nabla u)\, \eta'_\delta(u)\, dx\, dt - \int_{Q_T} (\frac{q}{2}\beta(u)^2 - pr)\, \eta'_\delta(u)\, dx\, dt \\
&\quad - \frac{1}{2(p-1)}\int_{Q_T} \underbrace{(\nabla u)^\top C(u)\nabla u}_{\leq ||C(u)||_2|\nabla u|^2}\, \eta'_\delta(u)\, dx\, dt + \frac{1}{2}\int_{Q_T} \underbrace{(\nabla' u)^\top C'(u)\nabla' u}_{\leq ||C'(u)||_2|\nabla' u|^2}\, \eta'_\delta(u)\, dx\, dt \\
&\quad - \frac{1}{2}\int_{Q_T} ((\operatorname{div} C(u))\cdot\nabla u + (\operatorname{div}' C'(u))\cdot\nabla' u)\, \eta'_\delta(u)\, dx\, dt.
\end{aligned} \tag{4.14}$$

Since $u \in L^2(0,T;H^1(\hat{\Omega})) \cap L^\infty(Q_T)$, $u_t \in L^2(0,T;H^{-1}(\hat{\Omega}))$ and η'_δ is smooth it holds [**148**, Prop. 23.20]

$$\int_0^T \langle u_t, \eta'_\delta(u)\rangle\, d\tau = \int_{\hat{\Omega}} \eta_\delta(u(T))\, dx - \int_{\hat{\Omega}} \eta_\delta(u_0)\, dx.$$

Since $|\eta_\delta'(u)| \le 1$, the right hand side of (4.14) is bounded independently of δ (and $\hat\Omega$) and we obtain, after letting $\delta \to 0$,

$$\int_{\hat\Omega} |u(T)|\,dx - \int_{\hat\Omega} |u_0|\,dx \le c.$$

This yields (4.13) for some constant $c = c(T)$. □

We are now able to prove the following theorem.

Theorem 4.8. *Let (H1)–(H5) hold. Then there exists a solution u of the Cauchy problem* (1.30a), (1.30c) *such that $u \in L^2(0,T;H^1(\mathbb{R}^d \times \mathbb{R}^{d'}))) \cap L^\infty(R_T)$ and $u \in H^1(0,T;H^{-1}(\mathbb{R}^d \times \mathbb{R}^{d'}))$.*

Proof. Let $(\hat\Omega^n)_n$ be a sequence of domains with smooth boundaries $\partial\hat\Omega^n$ satisfying $\hat\Omega^n \subset \hat\Omega^{n+1}$ and tending to $\mathbb{R}^d \times \mathbb{R}^{d'}$ in the set–theoretical sense as $n \to \infty$. By Theorem 4.5, in each of the cylinders $Q_T^n := \hat\Omega^n \times (0,T)$ there exists a solution $u^n \in L^2(0,T;H_0^1(\hat\Omega^n)) \cap L^\infty(Q_T^n)$ satisfying $u^n(0) = u_0|_{\hat\Omega^n}$. Under the additional assumption (H5) the constants c in the proof of Lemma 4.2 are independent of $\hat\Omega^n$, implying that these solutions are uniformly bounded in L^∞, i.e., it holds

$$\|u^n\|_{L^\infty(Q_T^n)} \le K,$$

where $K > 0$ is independent of $n \in \mathbb{N}$. Furthermore, the estimates in the proof of Lemma 4.4 are independent of $\hat\Omega^n$ if (H5) holds. In view of Lemma 4.7 we have for $n \ge m$

$$\|u^n\|_{L^2(0,T;H_0^1(\hat\Omega^m))} \le c \tag{4.15}$$

with c independent of n, m.

We can extract a subsequence $(u^{n,m})$ of (u^n) that converges weakly to some $u^{(m)} \in L^2(0,T;H_0^1(\hat\Omega^m)) \cap L^\infty(Q_T^m)$ as $n \to \infty$. Following the lines of the proof of Theorem 4.5 we can see that in fact $u^{n,m} \to u^{(m)}$ strongly in $L^2(0,T;H_0^1(\hat\Omega^m))$ and therefore also a.e. in Q_T^m. We have the following diagonal scheme

$$\begin{array}{ccccccc}
u^{1,1}, & u^{2,1}, & u^{3,1}, & \dots & \to & u^{(1)} = u|_{Q_T^1} \\
 & u^{2,2}, & u^{3,2}, & \dots & \to & u^{(2)} = u|_{Q_T^2} \\
 & & u^{3,3}, & \dots & \to & u^{(3)} = u|_{Q_T^3} \\
 & & & \ddots & & \vdots
\end{array}$$

More precisely, there exists a subsequence $u^{n,1}$ of u^n that converges strongly to some $u^{(1)}$ in $L^2(0,T;H_0^1(\hat\Omega^1))$ (and a.e. in Q_T^1). Furthermore, from this subsequence, we can select a subsequence $u^{n,2}$ that converges strongly to some $u^{(2)}$ in $L^2(0,T;H_0^1(\hat\Omega^2))$ with $u^{(2)}|_{Q_T^1} = u^{(1)}$, etc. The diagonal sequence $u^{n,n}$ tends to some $u \in L^2(0,T;H_0^1(\mathbb{R}^d \times \mathbb{R}^{d'})) \cap L^\infty(R_T)$ which is a solution to the Cauchy problem. $\square$

4.4 Uniqueness of solutions

In this section let either $\hat\Omega \subset \mathbb{R}^d \times \mathbb{R}^{d'}$ be a bounded domain or $\hat\Omega = \mathbb{R}^d \times \mathbb{R}^{d'}$. We make one of the following additional assumptions:

(H6) The matrices $C = C(S,S',t)$, $C' = C'(S,S',t)$ do not depend on u,

or

(H7) $p > 1$, the matrices $\partial C/\partial u$, $\partial C'/\partial u$ are positive semi–definite, the derivatives $\partial(\operatorname{div} C)/\partial u$, $\partial(\operatorname{div}' C')/\partial u$ are uniformly bounded with respect to S, S', t, and there exist positive constants ν, ν' such that $\|\partial C/\partial u\|_2 < \nu$, $\|\partial C'/\partial u\|_2 < \nu'$ with $\nu = \frac{2\varepsilon\eta}{n} - \frac{L_1}{n}$, $\nu' = \frac{2\varepsilon\eta'}{n} - \frac{L_1}{n}$ where $\varepsilon = \min(\alpha,\alpha')/2$ and η, η', L_1, n are defined in (4.22), (4.23), (4.24), respectively.

The smallness assumptions on $\partial C/\partial u$ and $\partial C'/\partial u$ seem quite restrictive. However, the dependence of the (co)variances on the solution, e.g. in the case of a large investor whose portfolio may influence stock prices, and consequently $\partial C/\partial u$ and $\partial C'/\partial u$ will usually be small. Therefore condition (H7) is less limiting in practice.

Lemma 4.9. *Assume (H1)–(H4) and either (H6) or (H7). If $\hat\Omega = \mathbb{R}^d \times \mathbb{R}^{d'}$ we also assume (H5). Then the problem* (1.30) *has a unique solution in the space of generalized solutions.*

Proof. Let u be a solution of (1.30). We introduce the transformation $u = \varphi(v) = -\ln(e^{-KAv} + 1/K)/A$ for some constants $A, K > 0$, which are chosen later. Using the test function $\phi = \psi/\varphi'(v)$ for arbitrary $\psi \in L^2(0,T;H_0^1(\hat\Omega)) \cap$

$L^\infty(0,T;L^\infty(\hat\Omega))$ in (4.1) yields

$$\begin{aligned}
0 =& \int_0^t \langle v_t, \psi\rangle \, dt + \frac{1}{2}\int_{Q_T} \Big[(\nabla\psi)^\top C(\varphi(v))\nabla v + (\nabla'\psi)^\top C'(\varphi(v))\nabla' v\Big] \, dx\, dt \\
&- \frac{1}{2}\int_{Q_T} \Big[\frac{\varphi''(v)}{\varphi'(v)}(\nabla v)^\top C(\varphi(v))\nabla v + \frac{\varphi''(v)}{\varphi'(v)}(\nabla' v)^\top C'(\varphi(v))\nabla' v\Big]\psi \, dx\, dt \\
&- \int_{Q_T} \frac{1}{\varphi'(v)}\Big[\mu\cdot\nabla\varphi(v) + \mu'\cdot\nabla'\varphi(v) \\
&\qquad\qquad - q(\mu - rS)\cdot\nabla\varphi(v) - \frac{q}{2}\beta(\varphi(v))^2 + pr\Big] \, \psi \, dx\, dt \\
&+ \frac{1}{2(p-1)}\int_{Q_T} \frac{1}{\varphi'(v)}(\nabla\varphi(v))^\top C(\varphi(v))\nabla\varphi(v) \, \psi \, dx\, dt \\
&- \frac{1}{2}\int_{Q_T} \frac{1}{\varphi'(v)}(\nabla'\varphi(v))^\top C'(\varphi(v))\nabla'\varphi(v) \, \psi \, dx\, dt \\
&+ \frac{1}{2}\int_{Q_T} \frac{1}{\varphi'(v)}((\operatorname{div} C)(\varphi(v))\cdot\nabla\varphi(v) + (\operatorname{div}' C')(\varphi(v))\cdot\nabla'\varphi(v)) \, \psi \, dx\, dt.
\end{aligned}$$

The transformed problem is of the form

$$v_t - \operatorname{div}_{\hat\xi}(a((S,S'),t,v,(\nabla v,\nabla' v))) + b((S,S'),t,v,(\nabla v,\nabla' v)) = 0, \tag{4.16}$$

with

$$\begin{aligned}
&a(\hat S,t,v,\hat\xi) = \begin{pmatrix} C(\varphi(v))\xi \\ C'(\varphi(v))\xi' \end{pmatrix}, \\
&b(\hat S,t,v,\hat\xi) \\
=&\Big[-\frac{\varphi''(v)}{2\varphi'(v)}\xi^\top C(\varphi(v)) - \mu + q(\mu - rS) + \frac{\varphi'(v)}{2(p-1)}\xi^\top C(\varphi(v)) + \frac{1}{2}\operatorname{div} C(\varphi(v))\Big]\xi \\
&+ \Big[-\frac{\varphi''(v)}{2\varphi'(v)}\xi'^\top C'(\varphi(v)) - \mu' - \frac{\varphi'(v)}{2}\xi'^\top C'(\varphi(v)) + \frac{1}{2}\operatorname{div}' C'(\varphi(v))\Big]\xi' \\
&+ \frac{1}{\varphi'(v)}\Big(\frac{q}{2}\beta(\varphi(v))^2 - pr\Big),
\end{aligned}$$

where $\hat S = (S,S')^\top$, $\hat\xi = (\xi,\xi')^\top$ and $\operatorname{div}_{\hat\xi} = (\operatorname{div}_\xi, \operatorname{div}_{\xi'})$ is the vectorized divergence operator. Notice that (4.16) and (1.30) (including initial and boundary conditions) are equivalent, since φ is invertible for suitable $A, K > 0$.

We now show that (4.16) has a unique solution. This implies the uniqueness of solutions also for problem (1.30) since the transformation φ is invertible. Let v_1, v_2 be two solutions to (4.16) satisfying the same initial condition and set $v := v_1 - v_2$. Using $(v^+)^n = \big(\max(0,v)\big)^n$, $n \in \mathbb{N}$, as test function in the equations satisfied by v_1 and v_2, respectively, and subtracting the resulting equations we get

$$\begin{aligned} 0 = \int_0^t \langle v_t, (v^+)^n\rangle\, dt + \int_{Q_T} n(v^+)^{n-1} & \qquad (4.17)\\ \times (\nabla v^+, \nabla' v^+) \cdot [a(\hat{S}, t, u_1, (\nabla u_1, \nabla' u_1)) - a(\hat{S}, t, u_2, (\nabla u_2, \nabla' u_2))]\, dx\, dt &\\ + \int_{Q_T} [b(\hat{S}, t, u_1, (\nabla u_1, \nabla' u_1)) - b(\hat{S}, t, u_2, (\nabla u_2, \nabla' u_2))](v^+)^n\, dx\, dt. & \end{aligned}$$

The difference in a can be expressed as

$$\begin{aligned} & a(\hat{S}, t, u_1, (\nabla u_1, \nabla' u_1)) - a(\hat{S}, t, u_2, (\nabla u_2, \nabla' u_2)) \\ &= \int_0^1 \frac{\partial}{\partial \tau} a(\hat{S}, t, \tau u_1 + (1-\tau)u_2, (\nabla(\tau u_1 + (1-\tau)u_2), \nabla'(\tau u_1 + (1-\tau)u_2)))\, d\tau \\ &= \int_0^1 \left[\frac{\partial a}{\partial v}(\hat{S}, t, u_\tau, (\nabla u_\tau, \nabla' u_\tau))v + \frac{\partial a}{\partial \hat{\xi}}(\hat{S}, t, u_\tau, (\nabla u_\tau, \nabla' u_\tau))(\nabla v, \nabla' v)^\top\right] d\tau, \end{aligned}$$

where $u_\tau = \tau u_1 + (1-\tau)u_2$, and similarly for the difference in b. Using the above expressions in (4.17) we obtain

$$\begin{aligned} \int_0^t \langle v_t, (v^+)^n\rangle\, dt + \int_{Q_T}\int_0^1 n(v^+)^{n-1}(\nabla v^+, \nabla' v^+) \cdot \Big[\frac{\partial a}{\partial v} v + \frac{\partial a}{\partial \hat{\xi}} \cdot (\nabla v, \nabla' v)^\top\Big]\, d\tau\, dx\, dt & \\ + \int_{Q_T}\int_0^1 \Big[\frac{\partial b}{\partial v} v + \frac{\partial b}{\partial \hat{\xi}} \cdot (\nabla v, \nabla' v)^\top\Big](v^+)^n\, d\tau\, dx\, dt = 0, & \end{aligned}$$

omitting the arguments, where $\partial b/\partial\hat{\xi}$ is the vector containing the partial derivatives of b with respect to ξ and ξ'. Using (H2) this leads to

$$\begin{aligned}
&\frac{1}{n+1}\int_{\hat{\Omega}}(v^+)^{n+1}(t)\,dx\\
&+\int_{Q_T}\int_0^1 n(v^+)^{n-1}\Big(\alpha|\nabla v^+|^2+\alpha'|\nabla' v^+|^2+(\nabla v^+,\nabla' v^+)\cdot\frac{\partial a}{\partial v}v\Big)\,d\tau\,dx\,dt\\
&\qquad+\int_{Q_T}\int_0^1\Big[\frac{\partial b}{\partial v}v+\frac{\partial b}{\partial\hat{\xi}}\cdot(\nabla v^+,\nabla' v^+)\Big](v^+)^n\,d\tau\,dx\,dt\le 0.
\end{aligned}$$

Employing Young's inequality with $\varepsilon=\min(\alpha,\alpha')/2$ we get

$$\begin{aligned}
&\frac{1}{n+1}\int_{\hat{\Omega}}(v^+)^{n+1}(t)\,dx+\int_{Q_T}\int_0^1\underbrace{n(v^+)^{n-1}\Big[(\alpha-\varepsilon)|\nabla v^+|^2+(\alpha'-\varepsilon)|\nabla' v^+|^2\Big]}_{\ge 0}\\
&\le\int_{Q_T}(v^+)^{n+1}\int_0^1\underbrace{\left[-\frac{\partial b}{\partial v}+\frac{n}{2\varepsilon}\Big|\frac{\partial a}{\partial v}\Big|^2+\frac{1}{2\varepsilon n}\Big|\frac{\partial b}{\partial\xi}\Big|^2+\frac{1}{2\varepsilon n}\Big|\frac{\partial b}{\partial\xi'}\Big|^2\right]}_{=:F(S,S',t,\xi,\xi',\tau)}\,d\tau\,dx\,dt.
\end{aligned}\tag{4.18}$$

The idea now is to show that $F(S,S',t,\xi,\xi',\tau)$ is bounded. This idea has been first used by Barles and Murat [**8**]. In the case (H7) with covariance matrices depending on u, we will make explicit use of the sign of $p-1$ to obtain the necessary estimates. A computation leads to

$$\begin{aligned}
&\frac{\partial b}{\partial v}(\hat{S},t,v,(\nabla v,\nabla' v))\\
&=-\frac{1}{2}\Big(\frac{\varphi''}{\varphi'}\Big)'(v)\Big(\frac{(\nabla\varphi(v))^\top C(\varphi(v))\nabla\varphi(v)}{\varphi'(v)^2}+\frac{(\nabla'\varphi(v))^\top C'(\varphi(v))\nabla'\varphi(v)}{\varphi'(v)^2}\Big)\\
&\quad-\frac{1}{2}\frac{\varphi''}{\varphi'}(v)\Big(\frac{(\nabla\varphi(v))^\top\frac{\partial C}{\partial v}(\varphi(v))\nabla\varphi(v)}{\varphi'(v)^2}+\frac{(\nabla'\varphi(v))^\top\frac{\partial C'}{\partial v}(\varphi(v))\nabla'\varphi(v)}{\varphi'(v)^2}\Big)\\
&\quad+\frac{\varphi''}{2\varphi'(v)^2}\Big(\frac{1}{p-1}(\nabla\varphi(v))^\top C(\varphi(v))\nabla\varphi(v)-(\nabla'\varphi(v))^\top C'(\varphi(v))\nabla'\varphi(v)\Big)
\end{aligned}\tag{4.19}$$

$$+\frac{1}{2}\Big(\frac{1}{p-1}(\nabla\varphi(v))^\top\frac{\partial C}{\partial u}(\varphi(v))\nabla\varphi(v)+(\nabla'\varphi(v))^\top\frac{\partial C'}{\partial u}(\varphi(v))\nabla'\varphi(v)\Big)$$
$$+\frac{1}{2}\Big(\frac{\partial(\operatorname{div}C)}{\partial u}(\varphi(v))\nabla\varphi(v)+\frac{\partial(\operatorname{div}'C')}{\partial u}(\varphi(v))\nabla'\varphi(v)\Big)$$
$$+\frac{q}{2}(\mu-rS)^\top\frac{\partial C^{-1}}{\partial u}(\varphi(v))(\mu-rS)-\frac{\varphi''}{\varphi'(v)^2}\Big(\frac{q}{2}\beta(\varphi(v))^2-pr\Big),$$

recalling the definition $\beta^2(\varphi(v))=(\mu-rS)^\top C^{-1}(\varphi(v))(\mu-rS)$, and

$$\begin{aligned}\frac{\partial b}{\partial\xi}(\hat S,t,v,\nabla v)=&-\frac{\varphi''}{\varphi'}(v)C(\varphi(v))\nabla v-\mu+q(\mu-rS)\\&+\frac{1}{p-1}C(\varphi(v))\nabla\varphi(v)+\frac{1}{2}\operatorname{div}C(\varphi(v)),\\\frac{\partial b}{\partial\xi'}(\hat S,t,v,\nabla' v)=&-\frac{\varphi''}{\varphi'}(v)C'(\varphi(v))\nabla' v-\mu'-C'(\varphi(v))\nabla'\varphi(v)\\&+\frac{1}{2}\operatorname{div}'C'(\varphi(v)).\end{aligned}\tag{4.20}$$

We want to obtain expressions in terms of the original variable u. Using

$$\varphi'(v)=K-e^{Au},\quad \frac{\varphi''}{\varphi'}(v)=-Ae^{Au},\quad \Big(\frac{\varphi''}{\varphi'}\Big)'(v)\frac{1}{\varphi'(v)}=-A^2e^{Au},$$

we obtain from (4.19)

$$\begin{aligned}&\frac{\partial b}{\partial v}(\hat S,t,v,(\nabla v,\nabla' v))\\=&\frac{e^{Au}}{2(K-e^{Au})}\Big[A^2\Big((\nabla u)^\top C(u)\nabla u+(\nabla' u)^\top C'(u)\nabla' u\Big)+A\Big((\nabla u)^\top\frac{\partial C}{\partial u}(u)\nabla u\\&+(\nabla' u)^\top\frac{\partial C'}{\partial u}(u)\nabla' u-\frac{1}{p-1}(\nabla u)^\top C(u)\nabla u+(\nabla' u)^\top C'(u)\nabla' u\Big)\Big]\\&+\frac{1}{2}\Big(\frac{1}{p-1}(\nabla u)^\top\frac{\partial C}{\partial u}(u)\nabla u+(\nabla' u)^\top\frac{\partial C'}{\partial u}(u)\nabla' u\Big)\\&+\frac{1}{2}\Big(\frac{\partial(\operatorname{div}C)}{\partial u}(u)\nabla u+\frac{\partial(\operatorname{div}'C')}{\partial u}(u)\nabla' u\Big)\\&+\frac{q}{2}(\mu-rS)^\top\frac{\partial C^{-1}}{\partial u}(u)(\mu-rS)+\frac{Ae^{Au}}{K-e^{Au}}\Big(\frac{q\beta^2}{2}(\varphi(v))-pr\Big)\\\geq&\frac{e^{Au}}{2(K-e^{Au})}\Big[A(A-\frac{1}{p-1})(\nabla u)^\top C(u)\nabla u+A(A+1)(\nabla' u)^\top C'(u)\nabla' u\end{aligned}$$

$$
\begin{aligned}
&+ A(\nabla u)^\top \frac{\partial C}{\partial u}(u)\nabla u + A(\nabla' u)^\top \frac{\partial C'}{\partial u}(u)\nabla' u - c\Big] \\
&+ \frac{1}{2}\frac{\partial(\operatorname{div} C)}{\partial u}(u)\nabla u + \frac{1}{2}\frac{\partial(\operatorname{div}' C')}{\partial u}(u)\nabla' u
\end{aligned}
$$

for some $c > 0$ and using (H7) (in particular, we use here $p > 1$ since then $1/(p-1) > 0$). For the last two terms we use Young's inequality, for some $\delta > 0$:

$$
\frac{1}{2}\frac{\partial(\operatorname{div} C)}{\partial u}(u)\nabla u + \frac{1}{2}\frac{\partial(\operatorname{div}' C')}{\partial u}(u)\nabla' u \geq -\frac{\delta}{4}(|\nabla u|^2 + |\nabla' u|^2) - c(\delta), \tag{4.21}
$$

where $c(\delta) > 0$ is a constant which depends on δ and the L^∞–norm of $\partial(\operatorname{div} C)/\partial u$ and $\partial(\operatorname{div}' C')/\partial u$. Now choose $A > 1/(p-1)$. In view of (H2) and (H6) or (H7), respectively, we can estimate for sufficiently large $K > 0$ and sufficiently small $\delta > 0$,

$$
\frac{\partial b}{\partial v}(\hat{S}, t, v, (\nabla v, \nabla' v)) \geq \eta |\nabla u|^2 + \eta' |\nabla' u|^2 - c \tag{4.22}
$$

for some $\eta = \eta(\alpha, K, A, \delta)$, $\eta' = \eta'(\alpha', K, A, \delta) > 0$ and $c > 0$. Notice that $u \in L^\infty(Q_T)$.

The derivatives (4.20) in the original variable

$$
\begin{aligned}
\frac{\partial b}{\partial \xi}(\hat{S}, t, v, \nabla v) &= \frac{A e^{Au}}{K - e^{Au}} C(u)\nabla u - \mu + q(\mu - rS) \\
&\qquad + \frac{1}{p-1} C(u)\nabla u + \frac{1}{2}(\operatorname{div} C)(u), \\
\frac{\partial b}{\partial \xi'}(\hat{S}, t, v, \nabla' v) &= \frac{A e^{Au}}{K - e^{Au}} C'(u)\nabla' u - \mu' - C'(u)\nabla' u + \frac{1}{2}(\operatorname{div}' C')(u)
\end{aligned}
$$

can be estimated as follows:

$$
\begin{aligned}
&\Big|\frac{\partial b}{\partial \xi}(\hat{S}, t, v, \hat{\xi})\Big|^2 + \Big|\frac{\partial b}{\partial \xi'}(\hat{S}, t, v, \hat{\xi})\Big|^2 \\
\leq &\Big(\frac{e^{2Au}}{(K - e^{Au})^2} + \frac{1}{(p-1)^2}\Big)||C(u)||_2^2 |\nabla u|^2 + \Big(\frac{e^{2Au}}{(K - e^{Au})^2} + 1\Big)||C'(u)||_2^2 |\nabla' u|^2 \\
&+ (|\mu|^2 + q^2|\mu - rS|^2 + \frac{1}{4}|\operatorname{div} C(u)|^2) + (|\mu'|^2 + \frac{1}{4}|\operatorname{div}' C'(u)|^2) \\
\leq &L_1(|\nabla u|^2 + |\nabla' u|^2 + 1)
\end{aligned} \tag{4.23}
$$

for some positive constant $L_1 = L_1(p, ||C||_2, ||C'||_2)$. Further we can estimate

$$\Big|\frac{\partial a}{\partial v}\Big|^2 = \Big|\begin{pmatrix} \frac{\partial C}{\partial u}\nabla u \\ \frac{\partial C'}{\partial u}\nabla' u \end{pmatrix}\Big|^2 \le \Big\|\frac{\partial C}{\partial u}\Big\|_2^2 |\nabla u|^2 + \Big\|\frac{\partial C'}{\partial u}\Big\|_2^2 |\nabla' u|^2.$$

This implies

$$\begin{aligned} F(S, S', t, \xi, \xi', \tau) \le& \\ &- \eta|\nabla u|^2 - \eta'|\nabla' u|^2 + c + \frac{n}{2\varepsilon}\Big(\Big\|\frac{\partial C}{\partial u}\Big\|_2^2 |\nabla u|^2 + \Big\|\frac{\partial C'}{\partial u}\Big\|_2^2 |\nabla' u|^2\Big) \\ &+ \frac{L_1}{2\varepsilon n}\Big(|\nabla u|^2 + |\nabla' u|^2 + 1\Big). \end{aligned}$$

Choosing

$$n > L_1/(2\varepsilon \min(\eta, \eta')) \tag{4.24}$$

and assuming that

$$\Big\|\frac{\partial C}{\partial u}\Big\|_2^2 \le \frac{2\varepsilon\eta}{n} - \frac{L_1}{n^2} \quad \text{and} \quad \Big\|\frac{\partial C'}{\partial u}\Big\|_2^2 \le \frac{2\varepsilon\eta'}{n} - \frac{L_1}{n^2}, \tag{4.25}$$

we conclude that $F(S, S', t, \xi, \xi', \tau) \le L_1/(2\varepsilon n) + c$. Applying Gronwall's lemma in (4.18) yields $v \le 0$ in Q_T. In a similar way, we can use the test function $(\min(0, v))^n$ for odd $n \in \mathbb{N}$ to prove that $v_1 - v_2 \ge 0$ in Q_T. Hence $v_1 = v_2$ in Q_T which completes the proof. □

Combining Lemma 4.9 and Theorem 4.5 yields the following theorem.

Theorem 4.10. *Let (H1)–(H4) and either (H6) or (H7) hold. If* $\hat{\Omega} = \mathbb{R}^d \times \mathbb{R}^{d'}$ *we assume additionally (H5). Then there exists a unique solution to* (1.30) *such that* $u - u_D \in L^2(0, T; H_0^1(\hat{\Omega})) \cap L^\infty(0, T; L^\infty(\hat{\Omega}))$, $u \in H^1(0, T; H^{-1}(\hat{\Omega}))$.

Remark 4.11. With additional regularity it is possible to prove a uniqueness result for viscosity solutions of (1.30). If C, C' do not depend on u (condition (H6)) similar uniqueness results for viscosity solutions have been shown in [**95**]. A more general result, again for viscosity solutions, has been proved in [**121**] assuming a relaxed monotonicity hypotheses on C, C' depending on u.

4.5 Application and numerical example

In this section we present a numerical example showing the influence of the non–tradable state variables on the value function. Examples for such state variables are weather variables or an employee's personal income, which usually cannot be traded. We consider the case $d = 2$ and $d' = 1$, i.e. two risky assets S_1, S_2 and one non–tradable state variable $S' = S_3$. Thus we have to solve a three–dimensional parabolic problem. We choose the following covariance matrices

$$C(S_1, S_2) = \begin{pmatrix} 0.04S_1^2 & -0.01S_1S_2 \\ -0.01S_1S_2 & 0.005S_2^2 \end{pmatrix}, \quad C'(S_3) = \begin{pmatrix} 0.05S_3^2 \end{pmatrix}.$$

The returns are defined as the so–called Ornstein–Uhlenbeck–type drifts

$$\begin{pmatrix} \mu_1(S_1) \\ \mu_2(S_2) \\ \mu_3(S_3) \end{pmatrix} = \begin{pmatrix} \big((6 - S_1) + 0.2\big)S_1 \\ \big((4 - S_2) + 0.1\big)S_2 \\ \big((4 - S_3) + 0.3\big)S_3 \end{pmatrix},$$

and the interest rate is assumed to be zero (for simplicity). As initial condition we choose $u_0(S_1, S_2, S_3) = 0$ which corresponds to the initial capital $x = 1$. The risk aversion parameter is taken to be $p = 0.5$. We use quadratic finite elements and a standard Runge–Kutta time discretization as provided by the FEMLAB package for MATLAB to compute the numerical solution. We choose our computational domain as $[2, 10] \times [2, 6] \times [2, 12]$ and the time horizon as $[0, 0.8]$. We used approximately 23,000 3D elements to solve problem (1.30).

Figure 4.1 shows the contour plots of the solution at times $t = 0.1, 0.4, 0.8$ for various values of the state variable S_3. The solution $(S_1, S_2) \mapsto u(S_1, S_2, S_3)$ has a local minimum at $S^* = (S_1^*, S_2^*) = (6.2, 4.1)$, since the expected return of investments in the two assets is zero at this point and since the interest rate is assumed to be zero. The qualitative behavior of the solution in the variables S_1, S_2 is similar for different values of S_3. The variation with respect to S_3 is noticeable . More precisely, for the values shown in Figure 4.1, the maximal relative difference to the minimum $S_3^* = 4.0$ at time $t = 0.8$ equals

$$\sup_{(S_1,S_2)\in(2,10)\times(2,6)} \frac{|u(S_1, S_2, S_3, 0.8) - u(S_1, S_2, S_3^*, 0.8)|}{|u(S_1, S_2, S_3^*, 0.8)|} \approx \begin{cases} 4.6 & : S_3 = 2, \\ 2.9 & : S_3 = 12. \end{cases}$$

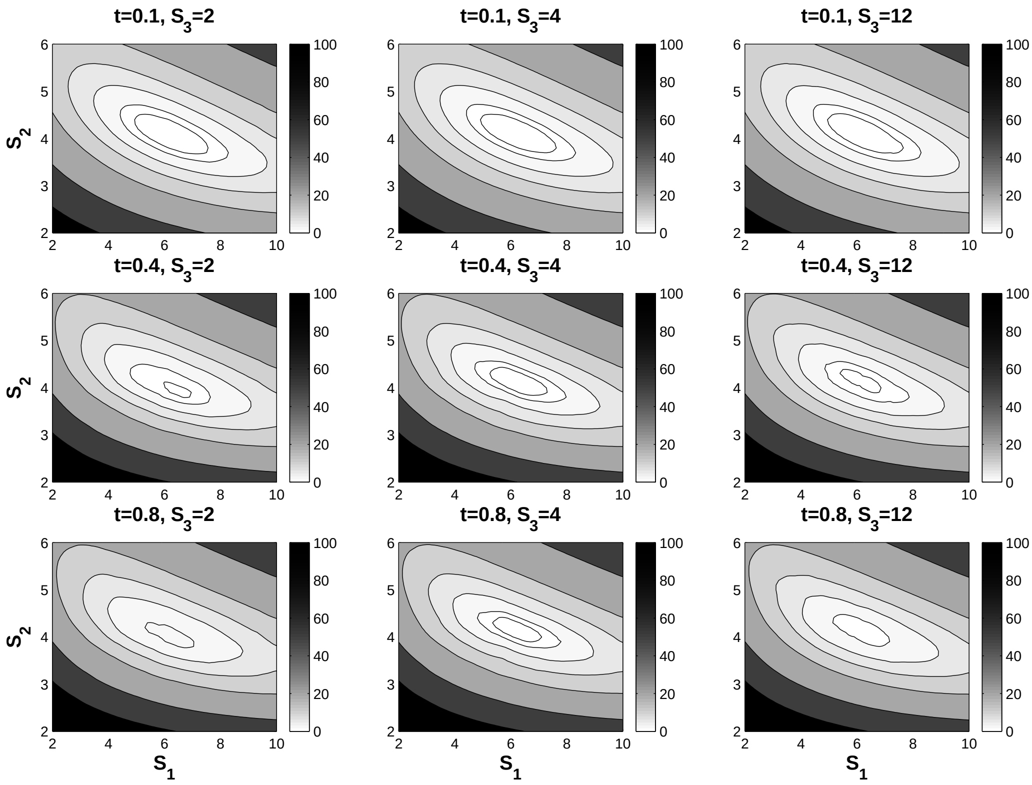

Figure 4.1: Contour plots of the numerical solution at times $t = 0.1, 0.4, 0.8$ and for fixed $S_3 = 2, 4, 12$.

For asset prices smaller than S^* the returns are increasing and hence the solution u, which relates to the utility of the optimal portfolio, too. In that region the optimal portfolio strategy $H(S_1, S_2, S_3) = (\lambda - \nabla u)/(p - 1)$ (the shares of the underlyings S_1 and S_2) has positive components. For sufficiently large asset prices the portfolio strategy has negative components. This indicates short selling for the optimal portfolio, which is permitted in the model. The above results give information on how to change the shares of the portfolio in order to achieve an optimal portfolio value.

Bibliography

[1] Y. Achdou, O. Pironneau. Volatility smile by multilevel least square. *Int. J. Theor. Appl. Finance* **5**(6), 619-643, 2002.

[2] L. Arnold. *Stochastische Differentialgleichungen.* Oldenbourg, München, 1973.

[3] H.W. Alt, S. Luckhaus. Quasilinear elliptic–parabolic differential equations. *Math. Z.* **183**, 311-341, 1983.

[4] M. Avellaneda, C. Friedman, R. Holmes, D. Samperi. Calibrating volatility surfaces via relative–entropy minimization. *Appl. Math. Finance* **4**(1), 37-64, 1997.

[5] L. Bachelier. Théorie de la spéculation. *Ann. Sci. Ecole Norm. Sup.* **17**, 21-86, 1900.

[6] V. Barbu. *Analysis and control of nonlinear infinite dimensional systems.* Math. Sci. Engrg. 190. Boston Academic Press, 1993.

[7] G. Barles, Ch. Daher, M. Romano. Convergence of numerical schemes for parabolic equations arising in finance theory. *Math. Models Meth. Appl. Sci.* **5**(1), 125-143, 1995.

[8] G. Barles, F. Murat. Uniqueness and the maximum principle for quasilinear elliptic equations with quadratic growth conditions. *Arch. Ration. Mech. Anal.* **133**(1), 77-101, 1995.

[9] G. Barles, B. Perthame. Discontinuous solutions of deterministic optimal stopping time problems. *Modélisation Math. Anal. Numér.* **21**, 557-579, 1987.

[10] G. Barles, B. Perthame. Exit time problems in optimal control and vanishing viscosity method. *SIAM J. Control Optim.* **26**(5), 1133-1148, 1988.

[11] G. Barles, H.M. Soner. Option Pricing with transaction costs and a nonlinear Black–Scholes equation. *Finance Stochast.* **2**, 369-397, 1998.

[12] G. Barles, P.E. Souganides. Convergence of approximation schemes for fully nonlinear second order equations. *Asympt. Anal.* **4**, 271-283, 1991.

[13] S. Benninga, J. Mayshar. Heterogeneity and option pricing, *Review Derivatives Res.* **4**, 7-27, 2000.

[14] B. Bensaid, J. Lesne, H. Pagès, J. Scheinkman. Derivative asset pricing with transaction costs. *Math. Finance* **2**, 63-86, 1992.

[15] A. Bensoussan, L. Boccardo, F. Murat. On a nonlinear partial differential equation having natural growth terms and unbounded solution. *Ann. Inst. Henri Poincaré, Anal. non linéaire* **5**(4), 347-364, 1988.

[16] G. Ben-Yu. Generalized stability of discretization and its applications to numerical solutions of nonlinear differential equations. American Mathematical Society. *Contemp. Math.* **163**, 33-54, 1994.

[17] H. Berestycki,J. Busca, I. Florent. Asymptotics and calibration of local volatility models. *Quant. Finance* **2**, 61-69, 2002.

[18] M. Bergounioux, K. Ito, K. Kunisch. Primal–dual strategy for constrained optimal control problems. *SIAM J. Contr. Optim.* **35**, 1524-1543, 1997.

[19] D. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, Massachusetts, 1995.

[20] F. Black, M. Scholes. The pricing of options and corporate liabilities. *J. Polit. Econ.* **81**, 637-659, 1973.

[21] L. Boccardo, T. Gallouët, F. Murat. A unified presentation of two existence results for problems with natural growth. In *Progress in Partial Differential Equations: The Metz Surveys 2*, M. Chipot (ed.). Pitman Res. Notes Math. Ser. **296**, 127-137, 1993.

[22] L. Boccardo, F. Murat, J.-P. Puel. Existence de solutions non bornées pour certaines équations quasi–linéaires. *Port. Math.* **41**, 507-534, 1982.

[23] L. Boccardo, F. Murat, J.-P. Puel. Existence de solutions faibles pour des équations elliptiques quasi–linéaires à croissance quadratique. *Nonlinear partial differential equations and their applications, Collège de France Semin., Vol. 4, Res. Notes Math.* **84**, 19-73, 1983.

[24] L. Boccardo, F. Murat, J.-P. Puel. Résultats d'existence pour certains problèmes elliptiques quasilinéaires. *Ann. Sc. Norm. Super. Pisa, IV. Ser.* **11**, 213-235, 1984.

[25] L. Boccardo, F. Murat, J.-P. Puel. Existence of bounded solutions for non linear elliptic unilateral problems. *Ann. Mat. Pura Appl., IV. Ser.* **152**, 183-196, 1988.

[26] L. Boccardo, F. Murat, J.-P. Puel. L^∞ estimate for some nonlinear elliptic partial differential equations and application to an existence result. *SIAM J. Math. Anal.* **23**, 326-333, 1992.

[27] R. Bodenmann, H.J. Schroll. Compact difference methods applied to initial–boundary value problems for mixed systems. *Numer. Math.* **73**(3), 291-309, 1996.

[28] P.T. Boggs, J.W. Tolle. Sequential quadratic programming. In *Acta Numerica 1995*, A. Iserles (ed.), 1-51, Cambridge University Press, 1995.

[29] J.F. Bonnans. *Optimisation numérique.* Springer–Verlag, Paris, 1997.

[30] J.F. Bonnans. Second–order analysis for control constrained optimal control problems of semilinear elliptic systems. *Appl. Math. Optim.* **38**, 303-325, 1998.

[31] I. Bouchouev, V. Isakov. Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets. *Inverse Probl.* **15**(3), R95-R116, 1999.

[32] S. Boyarchenko, S. Levendorskij. *Non–Gaussian Merton–Black–Scholes Theory.* World Scientific, Singapur, 2002.

[33] P. Boyle, T. Vorst. Option replication in discrete time with transaction costs. *J. Finance* **47**, 271-293, 1992.

[34] G.H. Bruce, D.W. Peaceman, H.H. Rachford, J.D. Rice. Calculations of Unsteady–State Gas Flow through Porous Media. *Trans. Am. Inst. Min. Engrs. (Petrol Div.)* **198**, 79-92, 1953.

[35] Q. Chen. A nonlinear parabolic system arising from the eddy currents problem. *Nonlin. Anal.* **42**, 759-770, 2000.

[36] G.M. Constantinides, T. Zariphopoulou. Bounds on process of contingent claims in an intertemporal economy with proportional transaction costs and general preferences. *Finance Stoch.* **3**(3), 345-369, 1999.

[37] J. Cox, J. Ingersoll, S. Ross. A theory of the term structure of interest rates. *Econometrica* **53**, 385-467, 1985.

[38] M. Crandall, H. Ishii, P.L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.* **27**, 1-67, 1992.

[39] M. Crandall, P.L. Lions. Viscosity solutions of Hamilton–Jacobi equations. *Trans. Amer. Math. Soc.* **277**, 1-42, 1983.

[40] S. Crépey. Calibration of the local volatility in a generalized Black–Scholes model using Tikhonov regularization. *SIAM J. Math. Anal.* **34**(5), 1183-1206, 2003.

[41] S. Crépey. Calibration of the local volatility in a trinomial tree using Tikhonov regularization. *Inverse Probl.* **19**(1), 91-127, 2003.

[42] R. Dautray, J.-P. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5: Evolution Problems.* Berlin, Springer–Verlag, 1992.

[43] M. Davis, V. Panis, T. Zariphopoulou. European option pricing with transaction fees. *SIAM J. Contr. Optim.* **31**, 470-493, 1993.

[44] J. Dewynne, S. Howison, P. Wilmott. *Option pricing: mathematical models and computation.* Oxford, Financial Press, 1995.

[45] D. Duffie. *Dynamic Asset Pricing Theory.* Princeton University Press, Princeton, 1996.

[46] B. Dumas, J. Fleming, R.E. Whaley. Implied Volatility Functions: Empirical Tests. J. Finance **53**(6), 2059-2106, 1998.

[47] B. Dupire. Pricing with a smile. *Risk* **7**, 18-20, 1994.

[48] B. Dupire. A Unified Theory of Volatility. In *Derivatives Pricing: The Classic Collection*, P. Carr (ed.), Risk Books, 2004.

[49] B. Düring, E. Lüders. Option prices under generalized pricing kernels. *Rev. Der. Research* **8**(2), 97-123, 2005.

[50] E. Eberlein, U. Keller. Hyperbolic distributions in finance. *Bernoulli* **1**, 281-299, 1995.

[51] H. Egger, H.W. Engl. Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates. *Inverse Problems* **21**, 1027-1045, 2005.

[52] N. Eisenbaum. On Itô's formula of Föllmer and Protter. In *Séminaire de Probabilités XXXV*, J. Azéma (ed.) et al. *Lect. Notes Math.* **1755**, 390-395, Springer, Berlin, 2001.

[53] H.W. Engl, M. Hanke, A. Neubauer. *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, 1996.

[54] L.C. Evans. *Partial Differential Equations.* AMS Graduate Studies in Mathematics, Providence, 1998.

[55] V. Ferone, M.R. Posteraro. On a class of quasilinear elliptic equations with quadratic growth in the gradient. *Nonlin. Anal.* **20**(6), 703-711, 1993.

[56] V. Ferone, M.R. Posteraro, J.M. Rakotoson. L^∞–estimates for nonlinear elliptic problems with p–growth in the gradient. *J. Inequal. Appl.* **3**(2), 109-125, 1999.

[57] H. Föllmer, P. Protter. On Itô's formula for multidimensional Brownian motion. *Probab. Theory Relat. Fields* **116**(1), 1-20, 2000.

[58] P. Forsyth, K. Vetzal, R. Zvan. A finite element approach to the pricing of discrete lookbacks with stochastic volatility. *Appl. Math. Finance* **6**, 87-106, 1999.

[59] J. Frehse. Existence and perturbation theorems for nonlinear elliptic systems. *Nonlinear Partial Diferential Equations and Their Applications*, Collège de France Semin., Vol. 4, Research Notes Math. **84**, 87-110, 1983.

[60] R. Frey. Perfect option hedging for a large trader. *Finance Stochast.* **2**, 115-141, 1998.

[61] R. Frey. Market illiquidity as a source of model risk in dynamic hedging. In *Model Risk*, R. Gibson (ed.), RISK Publications, London, 2000.

[62] G. Genotte, H. Leland. Market liquidity, hedging and crashes. *Amer. Econ. Rev.* **80**, 999-1021, 1990.

[63] V. Girault, P.-A. Raviart. *Finite Element Methods for Navier–Stokes Equations. Theory and Algorithms.* Springer–Verlag, Berlin, 1980.

[64] H. Goldberg, F. Tröltzsch. On a SQP–muligrid technique for nonlinear parabolic boundary problems. In *Optimal Control: Theory, Algorithms, and Applications*, W.W. Hager, P.M. Pardalos (eds.), Kluwer Academic Publishers, 1998.

[65] T. Goll, J. Kallsen. Optimal portfolios for logarithmic utility. *Stoch. Proc. Appl.* **89**, 31-48, 2000.

[66] C. Gourieroux, J.P. Laurent, H. Pham. Mean–variance heging and numéraire. *Math. Finance* **8**(3), 179-200, 1998.

[67] N. Grenon-Isselkou, J. Mossino. Existence de solutions bornées pour certaines équations elliptiques quasilinéaires. *C. R. Acad. Sci., Paris, Sér. I* **321**(1), 51-56, 1995.

[68] W.A. Gruver, E.W. Sachs. *Algorithmic Methods in Optimal Control.* Research Notes in Mathematics, vol. 47, Pitman, London, 1980.

[69] M. Günther, A. Jüngel. *Finanzderivate mit MATLAB.* Vieweg, Wiesbaden, 2003.

[70] M. Hanke, E. Rösler. Computation of local volatilities from regularized Dupire equations. *Int. J. Theor. Appl. Finance* **8**(2), 207-221, 2005.

[71] J.M. Harrison, S.R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes Appl.* **11**, 215-260, 1981.

[72] M. Heinkenschloss. Projected sequential quadratic programming methods. *SIAM J. Optim.* **6**, 373-417, 1996.

[73] M. R. Hestenes. Applications of the theory of quadratic forms in Hilbert space to the calculus of variations. *Pacific J. Math.* **1**, 525-581, 1951.

[74] S.L. Heston. Invisible Parameters in Option Prices. *Journal of Finance* **48**, 933-947, 1993.

[75] M. Hintermüller. On a globalized augmented Lagrangian–SQP algorithm for nonlinear optimal control problems with box constraints. In *Fast solution of discretized optimization problems*, K.-H. Hoffmann, R.H.W. Hoppe, V. Schulz (eds.), International Series of Numerical Mathematics 138, 139-153, 2001.

[76] M. Hintermüller. A primal–dual active set algorithm for bilaterally control constrained optimal control problems. *Quart. Appl. Math.* **61**(1), 131-161, 2003.

[77] M. Hintermüller, K. Ito, K. Kunisch. The primal–dual active set strategy as a semi–smooth Newton method. *SIAM J. Optimization* **13**, 865-888, 2003.

[78] S.D. Hodges, A. Neuberger. Optimal replication of contingent claims under transaction costs. *Rev. Future Markets* **8**, 222-239, 1989.

[79] G.A. Holton. *Value–at–Risk: Theory and Practice*, Academic Press, 2003.

[80] J. Hull, A. White. The pricing of options on assets with stochastic volatilities. *J. Finance* **42**, 281-300, 1987.

[81] V.P. Il'in. On high–order compact difference schemes. *Russ. J. Numer. Anal. Math. Model.* **15**(1), 29-46, 2000.

[82] H. Ishii. A boundary value problem of the Dirichlet type for Hamilton–Jacobi equations. *Ann. Scuola Norm. Sup. Pisa (4)* **16**(1), 105-135, 1989.

[83] K. Ito, K. Kunisch. The primal–dual active set method for nonlinear problems with bilateral constraints. *SIAM J. Control Optim.* **43**(1), 357-376, 2004.

[84] K. Ito, K. Kunisch. Augmented Lagrangian methods for nonsmooth, convex optimization in Hilbert spaces. *Nonl. Anal. TMA* **41**(5), 591-616, 2000.

[85] N. Jackson, E. Süli, S. Howison. Computation of deterministic volatility surfaces. *J. Comp. Finance* **2**, 5-32, 1999.

[86] J.C. Jackwerth. Recovering risk aversion from option prices and realized returns. *Review of Financial Studies* **13**, 433-451, 2000.

[87] J.C. Jackwerth, M. Rubinstein. Recovering Probability Distributions from Contemporary Security Prices, *Journal of Finance* **51**, 347-369, 1996.

[88] R. Jarrow. Market manipulation, bubbles, corners and short squeezes. *J. Financial Quant. Anal.* **27**, 311-336, 1992.

[89] P. Kangro, R. Nicolaides. Far field boundary conditions for Black–Scholes equations. *SIAM J. Numer. Anal.* **38**, 1357-1368, 2000.

[90] I. Kani, E. Derman, M. Kamal. Trading and Hedging Local Volatility. *J. Financ. Engeneering* **6**, 1233-1268, 1997.

[91] I. Karatzas, J.P. Lehoczky, S.E. Shreve, G.-L. Xu. Martingale and duality methods for utility maximization in an incomplete market. *SIAM J. Contr. Optim.* **29**(3), 702-730, 1991.

[92] I. Karatzas, S.E. Shreve. *Methods of Mathematical Finance*, Springer, New York, 1999.

[93] H.B. Keller. Approximation methods for nonlinear problems with application to two–point boundary value problems. *Math. Comp.* **29**, 464-476, 1975.

[94] J.L. Kelly. A new interpretation of the information rate. *Bell Systems Technical J.* **35**, 917-926, 1956.

[95] M. Kobylanski, Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.* **28**(2), 558-602, 2000.

[96] R. Korn, E. Korn. *Optionsbewertung und Portfoliooptimierung.* Vieweg, Braunschweig, 1999.

[97] D. Kramkov, W. Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Ann. Appl. Probab.* **9**(3), 904-950, 1999.

[98] S. Kusuoka. Option replication cost with transaction costs. In *Nonlinear and convex analysis in economic theory*, T. Maruyama (ed.), *Proceedings of the T.I. Tech/ K.E.S. conference on nonlinear and convex analysis in economic theory, Keio University, Japan, July 2-4, 1993.* Lect. Notes Econ. Math. Syst. **419**, 197-202, 1995.

[99] K. Kunisch, E.W. Sachs. Reduced SQP Methods for Parameter Identification Problems. *SIAM J. Numer. Anal.* **29**(6), 1793-1820, 1992.

[100] O.A. Ladyzenskaja, V.A. Solonnikov, N.N. Ural'ceva. *Linear and Quasilinear Equations of Parabolic Type*, AMS Transl., Vol. 23, Providence, 1968.

[101] R. Lagnado, S. Osher. A technique for calibrating derivative security pricing models: numerical solution of an inverse problem. *J. Comput. Finance* **1**, 13-25, 1997.

[102] Y. Lai, J. Spanier. Applications of Monte Carlo/quasi–Monte Carlo methods in finance: Option pricing. In *Monte Carlo/quasi–Monte Carlo Methods*, H. Niederreiter (ed.), *Proceedings of a conference held at the Claremont Graduate Univ., CA, USA, June 1998.* Springer, Berlin, 2000.

[103] D. Lamberton, B. Lapeyre. *Introduction au Calcul Stochastique Appliqué à la Finance.* 2ème éd. Ellipses, Paris, 1997.

[104] D. Lamberton. Error estimates for the binomial approximation of American put options. *Ann. Appl. Probab.* **8**, 206-233, 1998.

[105] J.P. Laurent, H. Pham. Dynamic programming and mean–variance hedging. *Finance Stoch.* **3**(1), 83-110, 1999.

[106] F. Leibfritz, E.W. Sachs. Inexact SQP interior point methods and large scale optimal control problems. *SIAM J. Contr. Optim.* **38**(1), 272-299, 1999.

[107] J. Leitner. Continous time CAPM, price for risk and utility maximization. In *Mathematical Finance. Workshop of the Mathematical Finance Research Project, Konstanz, Germany*, M. Kohlmann (ed.) et al. Birkhäuser, Basel, 2001.

[108] H. Leland. Option pricing and replication with transaction costs. *J. Finance* **40**, 1283-1301, 1985.

[109] P. Lévy. *Processus stochastiques et mouvement brownian.* Gauthiers–Villars, Paris, 1965.

[110] J.-L. Lions. *Control of Distributed Singular Systems.* Gauthier–Villars, Paris, 1983.

[111] D.G. Luenberger. *Linear and Nonlinear Progamming.* Addison-Wesley Publishing Company, Reading, Massachusetts, 1984.

[112] C. Maderna, C.D. Pagani, S. Salsa. Quasilinear elliptic equations with quadratic growth in the gradient. *J. Diff. Eqs.* **97**(1), 54-70, 1992.

[113] H. Markowitz. Portfolio selection. *J. Finance* **7**, 77-91, 1952.

[114] R.C. Merton. Lifetime portfolio selection under uncertainty: the continuous–time case. *Rev. Econom. Statist.* **51**, 247-257, 1969.

[115] R.C. Merton. Optimum consumption and portfolio rules in a time–continuous world. *J. Econ. Theory* **3**, 373-413, 1971.

[116] R.C. Merton. Theory of rational option pricing. *Bell J. Econ. Manag. Sci.* **4**, 141-183, 1973.

[117] R.C. Merton. Option pricing when the underlying stock returns are discontinuous. *J. Financ. Econom.* **3**, 125-144, 1976.

[118] R.C. Merton. *Continuous–Time Finance*, Blackwell, Oxford, 1990.

[119] D. Michelson. Convergence theorem for difference approximations of hyperbolic quasi–linear initial–boundary value problems. *Math. Comput.* **49**, 445-459, 1987.

[120] B. Øksendal. *Stochastic Differential Equations.* Springer, Berlin, 1998.

[121] M. Papi. A generalized Osgood condition for viscosity solutions to fully nonlinear parabolic degenerate equations. *Adv. Differ. Equ.* **7**(9) (2002), 1125-1151.

[122] A. Parás, M. Avellaneda. Dynamic hedging portfolios for derivative securities in the presence of large transaction costs. *Appl. Math. Finance* **1**, 165-193, 1994.

[123] E. Platen, M. Schweizer. On feedback effects from hedging derivatives. *Math. Finance* **8**, 67-84, 1998.

[124] O. Pironneau, F. Hecht. Mesh adaption for the Black and Scholes equations. *East–West J. Numer. Math.* **8**, 25-35, 2000.

[125] H. Pham, T. Rheinländer, M. Schweizer. Mean–variance hedging for continuous processes: new proofs and examples. *Finance Stoch.* **2**, 173-198, 1998.

[126] J.-M. Rakotoson. Réarrangement relatif dans les équations elliptiques quasi–linéaires avec un second membre distribution: Application à un théorème d'existence et de régularité. *J. Diff. Eqs.* **66**, 391-419, 1987.

[127] R.D. Richtmyer, K.W. Morton. *Difference Methods for Initial Value Problems.* Interscience, New York, 1967.

[128] A. Rigal. High Order Difference Schemes for Unsteady One–Dimensional Diffusion–Convection Problems. *J. Comp. Phys.* **114**(1), 59-76, 1994.

[129] A. Rigal. Numerical Analysis of three–time–level Finite Difference Schemes for Unsteady Diffusion–Convection Problems. *J. Num. Meth. Engineering* **30**, 307-330, 1990.

[130] K. Schittkowski. On the convergence of a sequential quadratic programming method with an augmented Lagrangian line search function.

Math. Operationsforschung und Statistik, Ser. Optimization **14**, 197-216, 1983.

[131] Z. Schuss. *Theory and Applications of Stochastical Differential Equations.* John Wiley, New York, 1980.

[132] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl., IV. Ser.* **146**, 65-96, 1987.

[133] K.R. Sircar, G. Papanicolaou. General Black–Scholes models accounting for increased market volatility from hedging strategies. *Appl. Math. Finance* **5**(1), 45-82, 1998.

[134] H.M. Soner, S.E. Shreve, J. Cvitanic. There is no nontrivial hedging portfolio for option pricing with transaction costs. *Ann. Appl. Probab.* **5**(2), 327-355, 1995.

[135] T. Sottinen. Fractional Brownian motion, random walks and binary market models. *Finance Stoch.* **5**, 343-355, 2001.

[136] H.J. Stetter. *Analysis of discretization methods for ordinary differential equations.* Springer Tracts in Natural Philosophy. Vol. **23**, 1973.

[137] G. Strang. Accurate partial difference methods. II: Non–linear problems. *Numer. Math.* **6**, 37-46, 1964.

[138] J.C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations.* Wadsworth & Brooks/Cole, Mathematics series, 1989.

[139] D. Tavella, C. Randall. *Pricing Financial Instruments: The Finite Difference Method.* John Wiley & Sons, 2000.

[140] F. Tröltzsch. On the Lagrange–Newton–SQP Method for the Optimal Control of Semilinear Parabolic Equations, *SIAM J. Contr. Optim.* **38**), 294-312, 1999.

[141] C. Wang, J. Liu. Fourth order convergence of compact finite difference solver for 2D incompressible flow. *Commun. Appl. Anal.* **7**(2), 171-191.

[142] R.F. Warming, B.J. Hyett. The modified equation approach to the stability and accuracy analysis of finite–difference methods. *J. Comput. Phys.* **14**, 159-179, 1974.

[143] A. Whalley, P. Wilmott. An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Math. Finance* **7**, 307-324, 1997.

[144] P. Wilmott. *Paul Wilmott on Quantitative Finance.* John Wiley, Chichester, 2000.

[145] P. Wilmott, J. Dewynne, S. Howison. *Option Pricing: Mathematical Models and Computation.* Oxford Financial Press, Oxford, 1993.

[146] H. Windcliff, P.A. Forsyth, K.R. Vetzal. Analyis of the Stability of the Linear Boundary Condition for the Black-Scholes equation. *J. Comp. Finance* **8**(1), 65-92, 2004.

[147] X. Xu. Existence and uniqueness for the nonstationary problem of the electrical heating of a conductor due to the Joule–Thomson Effect. *Int. J. Math. Math. Sci.* **16**(1), 125-138, 1993.

[148] E. Zeidler. *Nonlinear Functional Analysis and Its Applications II/A*, Springer, New York, 1990.