ADAPTIVE FINITE ELEMENT APPROXIMATION OF THE BLACK-SCHOLES EQUATION BASED ON RESIDUAL-TYPE A POSTERIORI ERROR ESTIMATORS

A Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Huifang Li

December 2009

ADAPTIVE FINITE ELEMENT APPROXIMATION OF THE BLACK-SCHOLES EQUATION BASED ON RESIDUAL-TYPE A POSTERIORI ERROR ESTIMATORS

Huifang Li

APPROVED:

Prof. Ronald H.W. Hoppe,
Chairman

Prof. Edward Kao

Prof. Jiwen He

Prof. William Fitzgibbon

Prof. Tim Warburton

Dean, College of Natural Sciences and Mathematics

ADAPTIVE FINITE ELEMENT APPROXIMATION OF THE BLACK-SCHOLES EQUATION BASED ON RESIDUAL-TYPE A POSTERIORI ERROR ESTIMATORS

An Abstract of a Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Huifang Li

December 2009

Abstract

For the pricing of options on equity shares, the Black-Scholes equation has become an indispensable tool for agents on the financial market. Under the assumption that the value of the underlying share evolves in time according to a stochastic differential equation and some further assumptions on the financial market, the equation can be derived by an application of Itô's calculus. It represents a deterministic second order parabolic differential equation backward in time with the price of the option as the unknown and the interest rate and the volatility entering the equation as coefficient functions. Since analytical solutions in explicit form are only available in special cases, in general the equation must be solved by numerical methods based on appropriate discretizations in time and in space where the spatial variable is the value of the share. This can be done by finite difference techniques or finite element methods with respect to suitable partitions of the time interval and the spatial domain. If the volatility depends on the independent variables, sudden changes of the volatility may imply rapid local changes of the solution as well so that a solution-dependent time-stepping and space-meshing is appropriate in order to keep the computational work at a moderate level while maintaining the accuracy of the computed approximate solution. During the past thirty years, such an adaptive choice of the discretizations in time and in space based on reliable a posteriori estimators of the global discretization error has been developed for finite element methods and achieved some state of maturity for standard partial differential equations. This thesis is devoted to an application of adaptive finite element methods to the numerical solution of the Black-Scholes equation.

Contents

Chapter 1

Introduction

1.1 Preface

For the pricing of options on equity shares, the Black-Scholes equation has become an indispensable tool for agents on the financial market. Under the assumption that the value of the underlying share evolves in time according to a stochastic differential equation and some further assumptions on the financial market, the equation can be derived by an application of Itô's calculus. It represents a deterministic second order parabolic differential equation backward in time with the price of the option as the unknown and the interest rate and the volatility entering the equation as coefficient functions. Since analytical solutions in explicit form are only available in special cases, in general the equation must be solved by numerical methods based on appropriate discretizations in time and in space where the spatial variable is the value of the share. This can be done by finite difference techniques or finite element methods with respect to suitable partitions of the time interval and the spatial domain. If the volatility depends on the independent variables, sudden changes of the volatility may imply rapid local changes of the solution as well so that a solution-dependent time-stepping and space-meshing is appropriate in order to keep the computational work at a moderate level while maintaining the accuracy of the computed approximate solution. During the past thirty years, such an adaptive choice of the discretizations in time and in space based on reliable a posteriori estimators of the global discretization error has been developed for finite element methods and achieved some state of maturity for standard partial differential equations. This thesis is devoted to an application of adaptive finite element methods to the numerical solution of the Black-Scholes equation.

The thesis is organized as follows: In the remaining part of this introductory first chapter, we will briefly sketch the issue of option pricing for plain vanilla European options and review the classical Black-Scholes model as well as the derivation of the Black-Scholes equation. Moreover, we will introduce to the basic concepts of adaptive finite element methods including a discussion of the crucial properties of reliability and efficiency of a posteriori error estimators.

The second chapter is devoted to a more detailed exposition of the Black-Scholes model followed by the variational formulation of the Black-Scholes equation in a suitable Sobolev space setting which provides the basis for its numerical solution by finite element methods.

In the third chapter, we will be concerned with the discretization of the Black-Scholes equation using an implicit discretization in time and standard P1 conforming finite elements in space with respect to a simplicial triangulation of the spatial domain.

The main part of this thesis is the fourth chapter where we present a residual-type a posteriori error estimator consisting of a time error estimator and a space error estimator which will take care of the combined space-time adaptivity. We will establish both the reliability and the efficiency of the estimator as well as its local efficiency in the sense that the local contributions of the space error estimator can be bounded from above by appropriate norms of the discretization error on local patches associated with the elements of the spatial triangulation.

The fifth chapter deals with a description of the remaining basic ingredients of the adaptive cycle which are - besides the a posteriori error estimation - the solution of the fully discretized problem, the marking of the time intervals as well as of the elements of the triangulation for refinement, and the technical realization of the refinement strategy.

The following sixth chapter provides a detailed documentation of the numerical results for selected text examples illustrating the performance of the suggested error estimator.

Some concluding remarks are given in the final seventh chapter.

1.2 Pricing of Options

1.2.1 Vanilla European Options - An Economic Model

A European vanilla call option (put option) is a contract giving its owner theright to buy (sell) a fixed number of shares of a specific common stock at a fixed price K at a certain date T.

The specific stock is called the underlying asset or the underlying security. The price of the underlying asset will be referred to as the spot price and will be denoted by S or S_t.

Since an option gives the holder a right, it has a value which is called the option price. Denoting by $C_t = C(t)(P_t = P(t))$ the value of the call-option (put-option) at time t, we are interested in evaluating $C_t(P_t)$ for $0 \leq t \leq T$.

In order to do that we make the following assumptions:

(i) There is no-arbitrage, i.e. an immediate risk-free profit is not possible.

(ii) The transactions have no cost and are instantaneous.

(iii) The market is liquid and trade is possible at any times.

For simplicity we assume further that:

(i) There is no dividend on the basic asset.

(ii)There is a fixed interest rate $r > 0$ for bonds/credits with proportional yield.

Pricing the option at maturity is easy. If S_T is the spot price at maturity, and if

(i) $S_T > K$, then the owner of the call option will make a benefit of $S_T - K$ by exercising the option and immediately selling the asset,

(ii) $S_T \leq K$, then the owner of the call option will do nothing.

In summary, at maturity the value of the call is given by the payoff function

$$C_T = (S_T - K)_+ := max(S_T - K, 0).$$

Similarly, we obtain the value of a put at maturity. Here the payoff function is given by

$$P_T = (K - S_T)_+ := max(K - S_T, 0).$$

Theorem 1.1. *(Put-Call Parity) Let $K, S_t, P(S_t, t)$ and $C(S_t, t)$ be the value of a bond (with constant interest rate $r > 0$ and proportional yield), an asset, a put option and a call option. Under the previous assumptions, for $0 \leq t \leq T$ there holds*

$$S_t + P(S_t, t) - C(S_t, t) = Kexp(-r(T - t)).$$

Proof. A proof is given in [37]. □

This formula tells us that as soon we know the price of the call option we know as well the price of the put option, or vice versa.

Without any additional assumptions, e.g. on the behaviour of the underlying share price, we have the following structural result.

Theorem 1.2. *(Lower and Upper Bounds for Call/Put Options). Let* $K, S_t, P(S_t, t)$ *and* $C(S_t, t)$ *be the value of a bond (with constant interest rate* $r > 0$ *and proportional yield), an asset, a put option and a call option. Under the previous assumptions, for* $0 \leq t \leq T$ *there holds*

$$i) \quad 0 \leq (S_t - Kexp(-r(T-t)))_+ \leq C(S_t, t) \leq S_t,$$

$$ii) \quad 0 \leq (Kexp(-r(T-t) - S_t))_+ \leq P(S_t, t) \leq Kexp(-r(T-t)).$$

Proof. A proof is given in [37]. □

1.2.2 The (Classical) Black-Scholes Model

The Black-Scholes model is a continuous-time model involving one riskless asset and one risky asset. We take the time dynamics of the price β_t of the riskless asset to be given by the ordinary differential equation

$$d\beta_t = r\beta_t dt, \tag{1.1}$$

and the time dynamics of the price S_t of the risky asset to be given by the stochastic differential equation

$$dS_t = S_t(\mu dt + \sigma dB_t), \tag{1.2}$$

where r, μ and σ are some constants, $\sigma > 0$ and B_t stands for a (standard) Brownian motion. For an accurate mathematical meaning of (1.2) we refer to [30, 31, 36, 41].

We can interpret the drift parameter μ as an average rate of growth and σ as the volatility of the asset price. Setting $\beta_0 = 1$ we find that

$$\beta_t = e^{rt},$$

and set $S(0) = S_0$ we obtain

$$S_t = S_0 exp(\mu t - \frac{\sigma^2}{2} t + \sigma B_t).$$

We remark that S_t is a geometric Brownian motion, i.e. $log S_t$ is a (not necessarily standard) Brownian motion.

1.2.3 Option Pricing

The basic idea in the computation of the option price is to consider a replicating portfolio consisting of a_t units of the risky asset and b_t units of the riskless asset. The value of our portfolio is then given by

$$V_t = a_t S_t + b_t \beta_t. \tag{1.3}$$

Denoting by $h(S_T)$ the payoff function we have the terminal replication constraint $V_T = h(S_T)$. Because the option has no cash flow until the terminal time, the replicating portfolio must be continuously rebalanced in such a way that there is no cash flowing into or out of the portfolio until the terminal time T. In terms of stochastic differentials, this requirement is given by the equation

$$\text{self-financing condition: } dV_t = a_t dS_t + b_t d\beta_t. \tag{1.4}$$

This equation imposes a strong constraint on the possible values for a_t and b_t. When coupled with the termination constraint $V_T = h(S_T)$, the self-financing condition turns out to be enough to determine a_t and b_t uniquely.

Assuming the portfolio valueV_t can be written as $V_t = u(S_t, t)$ where u is an appropriately smooth function, applying Itôs formula respectively u and using the self-financing condition we obtain by a coefficient matching argument the backward-in-time parabolic boundary value problem with terminal Cauchy-Condition

$$\frac{\partial u}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 u}{\partial S^2} + rS\frac{\partial u}{\partial S} - ru = 0, \quad \text{in } \mathbb{R}_+ \times [0.T) \tag{1.5a}$$

$$u|_{t=T} = u_0, \quad \text{in } \mathbb{R}_+ \tag{1.5b}$$

where $u_0 = h(S_T)$. We shall refer to (1.5a) as the Black-Scholes equation. The derivation of the Black-Scholes equation reveals further that the portfolio weights are given by

$$a_t = \frac{\partial u}{\partial S} \quad \text{and } b_t = \frac{1}{r\beta_t}(\frac{\partial u}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 u}{\partial S^2}). \tag{1.6}$$

The Black-Scholes equation yields a formula for pricing the option at $t < T$:

Theorem 1.3. *The price of the European vanilla call (put) option is given by*

$$C(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2), \tag{1.7}$$

and

$$P(S,t) = -SN(-d_1) + Ke^{-r(T-t)}N(-d_2) \tag{1.8}$$

where

$$d_1 = \frac{log(\frac{S}{K}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \text{and} \quad d_2 = d_1 - \sigma\sqrt{T-t} \tag{1.9}$$

and

$$N(d) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d} e^{-\frac{x^2}{2}}dx. \tag{1.10}$$

Proof. A proof is given in [28]. □

One of the main features of the Black-Scholes model is the fact that the pricing formulas (1.7) and (1.8) as well as the hedging formulas (1.6) depend on only one non-observable parameter - the volatility σ.

As the heading already suggests there are more general stock and bond models. It is possible to work with

$$dS_t = \mu_t S_t dt + \sigma_t S_t dB_t \quad \text{and} \quad d\beta_t = r_t \beta_t dt \tag{1.11}$$

where one assumes nothing about the coefficients in the SDEs except that the processes $(r_t)_{t\geq 0}$ and $(\sigma_t)_{t\geq 0}$ are both nonnegative and $(S_t)_{t\geq 0}$ and $(\beta_t)_{t\geq 0}$ are both diffusion processes.

This leads one to the pricing formula

$$V_t = \beta_t \mathbb{E}_{\mathbb{Q}}(X/\beta_T | \mathcal{F}_T) \tag{1.12}$$

which establishes a counterpart to the Black-Scholes equation. Here $\mathbb{Q}$ is the unique probability measure equivalent to $\mathbb{P}$ such that the discounted asset price $(S_t/\beta_t)_{t\geq 0}$ is a $\mathbb{Q}$ -martingal on $[0, T]$ and X is a random variable which can be interpreted as the value on the underlying portfolio. The existence of $\mathbb{Q}$ is a consequence of Girsanovs Theorem. In the special case of constant volatility σ, constant drift parameter μ and constant interest rate r one can derive again the pricing and hedging formulas seen above.

1.3 The Black-Scholes Equation

In the early seventies of the last century, F. Black, M. Scholes, and R. Merton achieved a major breakthrough in the history of modern financial economics: they published their groundbreaking papers *The pricing of options and corporate liabilities* [3] and *Theory of* *rational option pricing pricing* [32], where they developed an option pricing formula which later became known as the *Black-Scholes formula*. The foundation of their research relied on works developed by scientists such as L. Bachelier, A. J. Boness, S.T. Kassouf, E.O. Thorp, and P. Samuelson. The Black-Scholes formula had a huge influence on pricing derivatives and hedging risks. It also gave rise to the growth of financial engineering in the eighties and nineties. Merton and Scholes received the 1997 Nobel Prize in Economics for this and related works. Though ineligible for the prize because of his death, Black was mentioned as a contributor by the Swedish academy.

There are several assumptions underlying the Black-Scholes model of calculating options pricing. The most significant one is that the volatility, a measure of how much a stock can be expected to move in the near-term, and the risk-free interest rate are constant over time and the underlying assets. The Black-Scholes model also assumes that stock prices follow a log-normal random walk in continuous time, and that stocks pay no dividends until expiration. The assumptions on the market conditions include no arbitrage and no transaction costs or taxes in buying or selling the stock or the options. It is possible to borrow and lend cash at a constant risk-free interest rate and to short sell underlying stocks.

As time went on, the Black-Scholes model had been found as being too simple to fit the market prices in practice. Much research has been conducted to modify the Black-Scholes model based on geometric Brownian motion in order to incorporate two empirical features of the stock prices:

- asymmetric leptokurtic features, i.e., the return distribution is skewed to the left and has a higher peak and two heavier tails than those of the normal distribution.,
- volatility smile. The volatility of the stock price is assumed to be a constant in the

Black-Scholes model, but it has been observed that the implied volatility curve which is a function of strike and maturity resembles a 'smile' with respect to strike price [25].

Several more elaborate models have been proposed to fit the empirical features. Popular ones include

- a Black-Scholes variant known as ARCH (Autoregressive Conditional Heteroskedasticity). This variant replaces constant volatility with stochastic (random) volatility. A number of different models was developed after that like GARCH, E-GARCH, N-GARCH, H-GARCH [9],

- a generalization of the Black-Scholes approach by assuming the spot price is a Levy process (financial modeling with jump processes [23]),

- the use of local volatility, i.e., assuming that the volatility in the Black-Scholes model is a function of time and of the prices of the underlying assets.

1.4 Adaptive Finite Element Methods

1.4.1 The Adaptive Cycle

Adaptive Finite Element Methods (AFEMs) for Partial Differential Equations (PDEs) on the basis of a posteriori error estimates have been intensively studied during the past decades and successfully applied to technologically relevant problems (cf., e.g., the monographs [2, 4, 5, 26, 34, 40] and the references therein). A convergence analysis of AFEMs in case of standard Lagrangian type finite elements for linear second order elliptic boundary value problems has been initiated in [24] and further studied in [33], whereas the issue of optimal complexity has been addressed in [7, 19, 38]. A convergence analysis of nonstandard finite element methods such as nonconforming and mixed elements as well as edge elements has been provided in [17, 18] and [29]. As far as parabolic initial-boundary value problems are concerned, adaptivity in space has to be combined with an automatic time-stepping. We refer to [6, 10, 11, 12] and [20] for details.

An adaptive edge finite element method (AEFEM) consists of successive loops of the cycle

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE} .$$

Here, SOLVE means the numerical solution of the fully discretized problem. The following step ESTIMATE involves the efficient and reliable a posteriori error estimation of the global discretization error. This area has reached some state of maturity documented by a bundle of monographs and numerous research articles published during the past decade (cf. [2, 4, 5, 26, 40] and the references therein). The third step MARK deals with the selection of the next time step and the selection of elements of the triangulation for coarsening and refinement based on the information provided by the error estimators. The final step REFINE is devoted to the technical realization of the coarsening and refinement of elements selected in MARK.

1.4.2 Reliability and Efficiency of Error Estimators

Given some fully discrete approximation $u_{h,\Delta t}$ of the solution u of a time-dependent partial differential equation, we want to gain information on the error $e_u(t_n) := (u - u_{h,\Delta t})(t_n)$ at the time instant t_n in some suitable norm $\|\cdot\|$ in order to improve the quality of the approximation by an appropriate choice of the next time step and by eventually refining or coarsening the finite element mesh. An *a posteriori error estimator* $\eta_{h,\Delta t}$ is a computable quantity that may depend on the data of the problem (computational domain, coefficients of the equation, right-hand side, boundary conditions), on the underlying triangulation, and on the available approximate solution $u_{h,\Delta t}$ and that provides information on the error in terms of upper and/or lower bounds.

In particular, an error estimator $\eta_{h,\Delta t}$ is called *reliable*, if it provides an upper bound for the error up to possible data oscillations $osc^{rel}_{h,\Delta t}$, i.e., if there exists a constant $C_{rel} > 0$, independent of the time-steps and mesh size of the underlying triangulation, such that

$$\|e_u(t_n)\| \;\leq\; C_{rel}\;\eta_{h,\Delta t} \;+\; osc^{rel}_{h,\Delta t}. \tag{1.13}$$

On the other hand, an estimator $\eta_{h,\Delta t}$ is said to be *efficient*, if up to possible data oscillations $osc^{eff}_{h,\Delta t}$ it gives a lower bound for the error, i.e., if there exists a constant $C_{eff} > 0$, independent of the time-steps and mesh size of the underlying triangulation, such that

$$\eta_{h,\Delta t} \leq C_{eff}\|e_u(t_n)\| + osc^{eff}_{h,\Delta t}. \tag{1.14}$$

Finally, an estimator $\eta_{h,\Delta t}$ is called *asymptotically exact*, if it is both reliable and efficient with $C_{rel} = C^{-1}_{eff}$.

The notion *reliability* is motivated by the use of the error estimator in *error control*. Given a *tolerance tol*, an idealized *termination criterion* would be

$$\|e_u(t_n)\| \leq tol. \tag{1.15}$$

Since the error $\|e_u(t_n)\|$ is unknown, we replace it by the upper bound in (1.13), i.e.,

$$C_{rel}\eta_{h,\Delta t} + osc^{rel}_{h,\Delta t} \leq tol. \tag{1.16}$$

We remark that the termination criterion (1.16) both requires the knowledge of C_{rel} and the incorporation of the data oscillation term $osc^{rel}_{h,\Delta t}$. In the special case $C_{rel} = 1$ and

$osc^{rel}_{h,\Delta t} \equiv 0$, it reduces to

$$\eta_{h,\Delta t} \leq tol.$$

Due to (1.13), the termination criterion (1.16) guarantees the error control (1.15) which justifies to call the error estimator reliable.

An alternative, but less used termination criterion is based on the lower bound (1.14), i.e., we require

$$\frac{1}{C_{eff}}\Big(\eta_{h,\Delta t} \;-\; osc^{eff}_{h,\Delta t}\Big) \leq tol. \tag{1.17}$$

Typically, the criterion (1.17) requires less computational time than (1.16) which motivates to call the estimator efficient.

Chapter 2

The Black-Scholes Model

2.1 Derivation of the Black-Scholes Model

Before deriving the generalized Black-Scholes model, we recall some notions of probability theory:

Given a set Ω, let $\mathcal{A}$ be a σ- algebra of subsets of Ω and $\mathbb{P}$ a nonnegative measure on Ω such that $\mathbb{P}(\Omega) = 1$. Then, the triple $(\Omega, \mathcal{A}, \mathbb{P})$ is called a *probability space*.

A *real-valued random variable* X on $(\Omega, \mathcal{A}, \mathbb{P})$ is an $\mathcal{A}$-measurable real-valued function on Ω; i.e., for each Borel subset B of $\mathbb{R}$, $X^{-1}(B) \in \mathcal{A}$. A *real-valued stochastic process* $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{A}, \mathbb{P})$ assigns to each time t a random variable X_t on $(\Omega, \mathcal{A}, \mathbb{P})$.

A *filtration* $F_t = (\mathcal{A}_t)_{t\geq 0}$ is an increasing family of σ- algebras $\mathcal{A}_t$, i.e., for $t > \tau$ we have $\mathcal{A}_\tau \subset \mathcal{A}_t \subset \mathcal{A}$. The σ-algebras $\mathcal{A}_t$ usually represents a certain past history available at time t.

A stochastic process $(X_t)_{t\geq 0}$ is said to be F_t*-adapted*, if X_t is $\mathcal{A}_t$-measurable for any $t \geq 0$.

A *stochastic process* ω is called a *Wiener process*, if the following conditions hold true:

(i) $\omega(0) = 0$.

(ii) The process ω has independent increments, i.e. if $r < s \leqslant t < u$, then $\omega(u) - \omega(t)$ and $\omega(s) - \omega(r)$ are independent stochastic variables.

(iii) For $s < t$, the stochastic variable $\omega(t) - \omega(s)$ has the Gaussian distribution $N[0, \sqrt{t-s}]$.

(iv) ω has continuous trajectories.

Theorem 2.1. *Let g be a process satisfying*

$$\int_a^b E[g^2(s)]ds < \infty,$$

i.e., g is adapted to the $\mathcal{F}_t^W$-filtration. Then, there holds

$$E[\int_a^b g(s)d\omega_s] = 0.$$

Theorem 2.2. $[It\hat{o}'s formula]$ *Assume that the process X has a stochastic differential given by*

$$dX(t) = \mu(t)dt + \sigma(t)d\omega_t, \tag{2.1}$$

where μ and σ are adapted processes. For a $C^{1,2}$-function f consider the process Z defined by

$$Z(t) = f(t, X(t)).$$

Then Z has the stochastic differential

$$df(t, X(t)) = (\frac{\partial f}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x_2} + \mu\frac{\partial f}{\partial x})dt + \sigma\frac{\partial f}{\partial x}d\omega_t. \tag{2.2}$$

We consider a security, henceforth called a *stock*, with price process

$$dS(t) = S(t)r(t)dt + S(t)\sigma(t, S_t)d\omega_t$$

under risk-neutral probability $\mathbb{P}$. The price B is the price of a risk free asset, if it has the dynamics

$$dB(t) = r(t)B(t)dt,$$

where r, σ are any adapted continuous, bounded, and nonnegative functions.

Lemma 2.3. *For a European put option with price*

$$u(S,t) = E(e^{-\int_t^T r(\tau)d\tau} u_0(S(T))|F_t) \tag{2.3}$$

and payoff $u_0(S(T)) := (S(T) - K)_-$ *at maturity time* T*, where* K *stands for the strike, if function* $P = P(S,t)$ *satisfies the Cauchy problem*

$$\frac{\partial}{\partial t}P(S,t) + \frac{\sigma^2(S,t)S^2}{2}\frac{\partial^2}{\partial S^2}P(S,t) + r(t)S\frac{\partial}{\partial S}P(S,t) - r(t)P(S,t) = 0, \tag{2.4a}$$

$$P(S,T) = u_0(S(T)), \tag{2.4b}$$

an easy application of Itô's formula shows

$$u(S,t) = P(S,t). \tag{2.5}$$

Proof. For $f(S(t),t) = e^{\int_t^T r(\tau)d\tau}P(S(t),t)$, Ito's formula gives

$$df(S(t),t) = dPe^{\int_t^T r(\tau)d\tau} - P(S(t),t)e^{\int_t^T r(\tau)d\tau)}r(t)dt \tag{2.6}$$

Since

$$\begin{aligned} dP(S(t),t) &= (\frac{\partial}{\partial t}P(S,t) + \frac{\sigma^2(S,t)S^2}{2}\frac{\partial^2}{\partial S^2}P(S,t) + r(t)S\frac{\partial}{\partial S}P(S,t))dt + \\ &\quad \sigma(S,t)\frac{\partial}{\partial S}P(S,t)d\omega_t, \end{aligned}$$

we have

$$
\begin{aligned}
df(S(t),t) = \quad & e^{\int_t^T r(\tau)d\tau}\Big(\frac{\partial}{\partial t}P(S,t) + \frac{\sigma^2(S,t)S^2}{2}\frac{\partial^2}{\partial S_2}P(S,t) \\
& + r(t)S\frac{\partial}{\partial S}P(S,t) - P(S,t)r(t)\Big)dt + \\
& \sigma(S,t)\frac{\partial}{\partial S}P(S,t)e^{\int_t^T r(\tau)d\tau}d\omega_t \qquad (2.7)
\end{aligned}
$$

By assumption, $P(S,t)$ actually satisfies (2.4) and hence, the drift part of (2.7) vanishes. If $\sigma(S,t)\frac{\partial}{\partial S}P(S,t)e^{\int_t^T r(\tau)d\tau}d\omega_t$ is sufficiently integrable, we obtain

$$
P(S(T),T) = e^{\int_t^T r(\tau)d\tau}P(S,t) + \int_t^T \sigma(S,t)\frac{\partial}{\partial S}P(S,t)e^{\int_t^T r(\tau)d\tau}d\omega_t. \qquad (2.8)
$$

Taking the expectation of $P(S(T),T)$, the stochastic integral will also vanish, whence $P(S,t) = \mathbb{E}(e^{-\int_t^T r(\tau)d\tau}P(S(T),T))$. Comparing with (2.3) yields (2.5). □

For convenience, we replace t by $T-t$ which transforms the final time to an initial value problem which reads as follows

$$
\frac{\partial}{\partial t}u(S,t) - \frac{\sigma^2(S,t)S^2}{2}\frac{\partial^2}{\partial S^2}u(S,t) - r(t)S\frac{\partial}{\partial S}u(S,t) + r(t)u(S,t) = 0, \qquad (2.9a)
$$

$$
u(S,0) = u_0(S). \qquad (2.9b)
$$

The problem (2.9a),(2.9b) has a unique *strong solution* $u \in \mathcal{C}^0(\mathbb{R}_+ \times [0,T])$ which is $\mathcal{C}^1$ -regular with respect to t and $\mathcal{C}^2$ -regular with respect to S and satisfies $0 \leqslant u(S,t) \leqslant C(1+S)$ for some constant $C \in \mathbb{R}_+$, if the following assumptions are satisfied (cf., e.g., [35])

($\mathbf{A_1}$) The function $(S,t) \mapsto S\sigma(S,t)$ is Lipschitz continuous on $\mathbb{R}_+ \times [0,T]$, bounded from above on $\mathbb{R}_+ \times [0,T]$ and bounded from below by a positive constant.

($\mathbf{A_2}$) The function r is bounded and Lipschitz continuous.

($\mathbf{A_3}$) The Cauchy data u_0 satisfies $0 \leqslant u_0(S) \leqslant C(1+S)$ for some $C \in \mathbb{R}_+$.

For later discretization purposes, we truncate the domain in the variable S and consider (2.9a),(2.9b) on $\Omega \times (0,T)$, where $\Omega := (0, \bar{S})$

$$\frac{\partial}{\partial t}u(S,t) - \frac{\sigma^2(S,t)S^2}{2}\frac{\partial^2}{\partial S^2}u(S,t) - r(t)S\frac{\partial}{\partial S}u(S,t) + r(t)u(S,t) = 0, \tag{2.10a}$$

$$u(\bar{S},t) = 0, \tag{2.10b}$$

$$u(S,0) = u_0(S). \tag{2.10c}$$

2.2 Variational Formulation of the Black-Scholes Equation

We use standard notation from Lebesgue and Sobolev space theory and denote by $\mathcal{D}(\Omega)$ the space of infinitely often differentiable functions with compact support in $\Omega \subset \mathbb{R}_+$ and by $L^2(\Omega), \Omega \subseteq \mathbb{R}_+$, the Hilbert space of square integrable functions on Ω with inner product $(\cdot,\cdot)_{0,\Omega}$ and associated norm $\|\cdot\|_{0,\Omega}$. We further refer to $H^1(\Omega)$ as the Hilbert space of square integrable functions with square integrable weak derivatives equipped with the norm $\|\cdot\|_{1,\Omega}$. The Hilbert spaces $L^2((0,T))$ and $H^1((0,T))$ are defined analogously.

In order to derive an appropriate variational formulation of (2.10a)-(2.10c), we introduce the *weighted Sobolev space*

$$V = \{v : v \in L^2(\Omega), S\frac{\partial v}{\partial S} \in L^2(\Omega)\}, \tag{2.11}$$

endowed with the inner product

$$(v,w)_V := \int_\Omega (v(S)w(S) + S^2\frac{\partial v}{\partial S}(S)\frac{\partial w}{\partial S}(S))dS, \tag{2.12}$$

where $\dfrac{\partial v}{\partial S}$ stands for the weak derivative, and we refer to $\|\cdot\|_V$ as the associated norm.

We define V_0 as the closure of $\mathcal{D}(\Omega)$ in V. Then, it is easy to see that V_0 is a closed subspace of V with $v(\bar{S}) = 0$ for $v \in V_0$. Moreover, the following *Poincaré-Friedrichs inequality* holds true:

Lemma 2.4. *(Poincaré-Friedrichs inequality) For all $v \in V_0$ there holds*

$$\|v\|_{L^2(\Omega)} \leq 2|v|_V. \tag{2.13}$$

Proof. Since $\mathcal{D}(\Omega)$ is dense in V_0, it suffices to prove (2.13) for $v \in \mathcal{D}(\Omega)$. Obviously, we have

$$\|v\|^2_{L^2(\Omega)} = \int_\Omega v^2 dS = -2 \int_\Omega Sv \frac{\partial v}{\partial S}(S) dS.$$

An application of the Cauchy-Schwarz inequality to the right-hand side gives

$$|\int_\Omega Sv\frac{\partial v}{\partial S}(S)dS| \leq (\int_\Omega (S\frac{\partial v}{\partial S}(S))^2 dS)^{1/2} (\int_\Omega v(S)^2 dS)^{1/2}$$

from which we deduce the desired result. □

Consequently, the semi-norm

$$|v|_V = (\int_\Omega S^2 (\frac{\partial v}{\partial S})^2 dS)^{1/2},$$

is in fact a norm on V_0 equivalent to $\| \cdot \|_V$. We refer to V_0^* as the dual of V_0 with norm $\| \cdot \|_{V_0^*}$ and to $\langle \cdot, \cdot \rangle_{V_0^*, V_0}$ as the dual pairing between V_0 and V_0^*.

We further denote by $L^2((0,T); L^2(\Omega))$ the Hilbert space equipped with the norm

$$\|u\|^2_{L^2((0,T);L^2(\Omega))} := \int_0^T \|u(t)\|^2_{0,\Omega} \; dt$$

and define $L^2((0,T); V_0)$ and $\|\cdot\|_{L^2((0,T);V_0)}$ analogously. Moreover, we introduce $H^1((0,T); V_0^*)$ as the Hilbert space with the norm

$$\|u\|^2_{H^1((0,T);V_0^*)} := \int_0^T \left(\|u(t)\|^2_{V_0^*} + \|u_t(t)\|^2_{V_0^*} \right) \, dt.$$

where $\|u\|_{V_0^*} = \sup_{v \in V_0} \frac{(u,v)}{|v|_V}$.

Now, multiplying (2.10a) by $v \in V_0$ and integrating over Ω, we obtain

$$\begin{aligned} 0 = \quad & \int_\Omega \frac{\partial}{\partial t} u(S,t)v(S)dS - \int_\Omega \frac{\sigma^2(S,t)S^2}{2} \frac{\partial^2}{\partial S^2} u(S,t)v(S)dS \\ & -r(t) \int_\Omega S \frac{\partial}{\partial S} u(S,t)v(S)dS + r(t) \int_\Omega u(S,t)v(S)dS. \end{aligned} \tag{2.14}$$

Integrating by parts and applying the fact that $v(\bar{S}) = 0$ results in

$$\begin{aligned} 0 = \quad & \int_\Omega \frac{\partial}{\partial t} u(S,t)v(S)dS + \int_\Omega \frac{\sigma^2(S,t)S^2}{2} \frac{\partial u}{\partial S}(S,t) \frac{\partial v}{\partial S}(S)dS \\ & + \int_\Omega (\sigma^2(S,t)S\sigma(S,t) + \frac{\partial \sigma}{\partial S}(S,t) - r(t)) S \frac{\partial}{\partial S} u(S,t)v(S)dS \\ & + r(t) \int_\Omega u(S,t)v(S)dS. \end{aligned} \tag{2.15}$$

In view of (2.15), we introduce the bilinear form $a_t(\cdot,\cdot) : V_0 \times V_0 \to \mathbb{R}$ according to

$$a_t(u,v) = (\frac{\sigma^2}{2} S \frac{\partial u}{\partial S}, S \frac{\partial v}{\partial S}) + ((-r + \sigma^2 + S\sigma \frac{\partial \sigma}{\partial S}) S \frac{\partial u}{\partial S}, v) + r(u,v). \tag{2.16}$$

Consequently, the boundary value problem (2.10a)-(2.10c) has the following variational formulation: Find $u \in H^1((0,T);V_0^*) \cap L^2((0,T);V_0)$ such that for all $v \in V_0$

$$\langle \frac{\partial u}{\partial t}, v \rangle_{V_0^*,V_0} + a_t(u,v) = 0, \tag{2.17a}$$

$$(u(\cdot,0),v)_{0,\Omega} = (u_0,v)_{0,\Omega}. \tag{2.17b}$$

We note that $H^1((0,T);V_0^*) \cap L^2((0,T);V_0)$ is continuously embedded in $C^0([0,T];L^2(\Omega))$ (cf., e.g., [35]).

In order to prove existence and uniqueness of a solution of (2.17a),(2.17b), we impose the following assumptions on σ and r

($\mathbf{A_4}$) The function σ is continuously differentiable, and there exist constants

$0 < \sigma_{min} \leq \sigma_{max}$ and $C_\sigma > 0$, such that for all $(S,t) \in \Omega \times [0,T]$ there holds

$$\sigma_{min} \leq \sigma(S,t) \leq \sigma_{max}, \tag{2.18a}$$

$$|S\frac{\partial \sigma}{\partial S}(S,t)| \leq C_\sigma. \tag{2.18b}$$

($\mathbf{A_5}$) The function r is continuous and nonnegative on $[0,T]$.

We first prove that under assumptions ($\mathbf{A_4}$) and ($\mathbf{A_5}$) the bilinear form $a_t(\cdot,\cdot)$ is bounded and satisfies a Gårding's-type inequality.

Lemma 2.5. *Under assumptions* ($\mathbf{A_4}$) *and* ($\mathbf{A_5}$) *there exists a constant* $\mu > 0$ *such that for all* $v, w \in V_0$ *there holds*

$$|a_t(v,w)| \leq \mu |v|_V |w|_V. \tag{2.19}$$

Proof. In view of ($\mathbf{A_4}$) we obtain

$$|(\frac{\sigma^2}{2}S\frac{\partial u}{\partial S}, S\frac{\partial v}{\partial S})| \leq \frac{\sigma_{max}^2}{2}|v|_V|u|_V,$$

Let $R = max_{t\in[0,T]}r(t)$. Then, by ($\mathbf{A_4}$), ($\mathbf{A_5}$) and the *Poincaré-Friedrichs inequality*

$$\begin{aligned}|((-r + \sigma^2 + S\sigma\frac{\sigma}{S})S\frac{\partial u}{\partial S}, v)| &\leq |R + \sigma_{max}^2 + C_\sigma\sigma_{max}||u|_V\|v\|_{L^2(\Omega)} \\ &\leq 2|R + \sigma_{max}^2 + C_\sigma\sigma_{max}||u|_V|v|_V.\end{aligned}$$

The *Cauchy-Schwarz inequality* and the *Poincaré-Friedrichs inequality* imply

$$|r(u,v)| \leq |R|\|u\|_{0,\Omega}\|v\|_{0,\Omega} \leq 4R|u|_V|v|_V$$

Finally, (2.19) follows with $\mu = \dfrac{\sigma_{max}^2}{2} + 2|R + \sigma_{max}^2 + C_\sigma\sigma_{max}| + 4R$.

□

Lemma 2.6. *(Gårding's inequality)*

Under assumptions $(\mathbf{A_4})$ *and* $(\mathbf{A_5})$ *there exists a nonnegative constant* λ *such that for all* $v \in V_0$ *there holds*

$$a_t(v, v) \geq \frac{1}{4}\sigma_{min}^2 |v|_V^2 - \lambda \|v\|^2. \tag{2.20}$$

Proof. Using $(\mathbf{A_4}), (\mathbf{A_5})$ and $R = max_{t\in[0,T]} r(t)$, straightforward computations reveal the following three inequalities

$$\begin{aligned} |(\frac{\sigma^2}{2} S \frac{\partial v}{\partial S}, S\frac{\partial v}{\partial S})| &\geq \frac{\sigma_{min}^2}{2}|v|_V^2, \\ |((-r + \sigma^2 + S\sigma\frac{\partial \sigma}{\partial S})S\frac{\partial u}{\partial S}, v)| &\leq |R + \sigma_{max}^2 + C_\sigma \sigma_{max}||u|_V \|v\|_{L^2(\Omega)} \\ &\leq \frac{\sigma_{min}^2}{4}|v|_V^2 + \hat{\lambda}|_V \|v\|_{L^2(\Omega)}, \end{aligned}$$

where $\hat{\lambda} = (R + \sigma_{max}^2 + C_\sigma \sigma_{max})^2/\sigma_{min}^2$. Since

$$|r(v, v)| \leq 4R|v|_V^2,$$

(2.20) follows with $\lambda = max(0, \hat{\lambda} - 4R)$. □

The previous results immediately give rise to the existence and uniqueness of the solution of (2.17a), (2.17b).

Theorem 2.7. *Suppose that the assumptions* $(\mathbf{A_4}), (\mathbf{A_5})$ *are satisfied and* $u_0 \in L^2(\Omega)$*. Then, the variational formulation* (2.17a),(2.17b) *has a unique solution. Moreover, for all* $0 < t < T$ *there holds*

$$e^{-2\lambda t}\|u(t)\|_{0,\Omega}^2 + \frac{1}{2}\sigma_{min}^2 \int_0^t e^{-2\lambda s}|u(s)|_V^2 \, ds \leq \|u_0\|_{0,\Omega}^2. \tag{2.21}$$

Proof. Existence can be shown by the Galerkin method, i.e., by constructing a sequence $u_n \in C^1((0,T); V_n), n \in \mathcal{N}$, of solutions of (2.17a),(2.17b) in finite dimensional subspaces

$V_n \subset V_0$ that are limit-dense in V_0 and then passing to the limit. For details of the existence proof we refer to [35]. Uniqueness readily follows from (2.21). For the proof of (2.21), we choose $v = u(t)e^{-2\lambda t}$ in (2.17a) and integrate over $(0,t)$ which gives

$$\int_0^t (\frac{\partial u}{\partial t}, u(\tau)e^{-2\lambda\tau})d\tau + \int_0^t a_\tau(u(\tau), u(\tau)e^{-2\lambda\tau})d\tau = 0. \tag{2.22}$$

Integrating by parts, we obtain

$$\|u_0\|^2 = \|u(t)\|^2 e^{-2\lambda t} - \int_0^t (u, \frac{\partial u}{\partial t}e^{-2\lambda\tau} - 2\lambda u e^{-2\lambda\tau})d\tau + \int_0^t e^{-2\lambda\tau} a_t(u(\tau), u(\tau))d\tau.$$

Now, setting

$$[[v]](t) := (e^{-2\lambda t}\|v(t)\|^2 + \frac{1}{2}\sigma_{min}^2 \int_0^t e^{-2\lambda t}|v(\tau)|_V^2 d\tau)^{\frac{1}{2}},$$

an application of Gårding's inequality yields

$$\begin{aligned}
&\|u_0\|^2 \\
&\geq \|u(t)\|^2 e^{-2\lambda t} - \int_0^t (u, \frac{\partial u}{\partial t}e^{-2\lambda\tau} - 2\lambda u e^{-2\lambda\tau})d\tau + \int_0^t e^{-2\lambda\tau}(\frac{1}{4}\sigma_{min}^2|u|_V^2 - \lambda\|u\|^2)d\tau \\
&= [[v]]^2(t) - \frac{1}{4}\int_0^t e^{-2\lambda\tau}\sigma_{min}^2|u|_V^2 d\tau - \int_0^t (u, \frac{\partial u}{\partial t}e^{-2\lambda\tau} - 2\lambda u e^{-2\lambda\tau})d\tau \\
&\geq [[v]]^2(t) - \int_0^t e^{-2\lambda\tau} a_\tau d\tau - \int_0^t (u(\tau)e^{-2\lambda\tau}, \frac{\partial u}{\partial t})d\tau \\
&= [[v]]^2(t),
\end{aligned}$$

from which we deduce (2.21). □

The stability estimate (2.21) motivates to consider the norm

$$[[v]](t) = (e^{-2\lambda t}\|v(t)\|^2 + \frac{1}{2}\sigma_{min}^2 \int_0^t e^{-2\lambda t}|v(\tau)|_V^2 d\tau)^{\frac{1}{2}}, \tag{2.23}$$

so that (2.21) reads

$$[[u]](t) \leq \|u_0\|. \tag{2.24}$$

Similar techniques as in the proof of (2.21) allow to establish the following estimate.

Lemma 2.8. *For any* $u \in H^1([0,T];V_0^*) \cap L^2(0,T;V_0) \subset C^0([0,T];L^2(\Omega))$ *there holds*

$$\|e^{-\lambda t}\frac{\partial u}{\partial t}\|_{L^2((0,T);V_0^*)} \leq \sqrt{2}\frac{\mu}{\sigma_{min}}\|u_0\|. \tag{2.25}$$

Proof. In view of Lemma 2.5 and (2.22) we get

$$\begin{aligned}
&|\int_0^T (\frac{\partial u}{\partial t}, u(\tau)e^{-2\lambda\tau})d\tau| \\
&\leq \int_0^t |a_\tau(u(\tau), u(\tau)e^{-2\lambda\tau})|d\tau \\
&\leq \sqrt{2}\frac{\mu}{\sigma_{min}}[[u]](T)|v|,
\end{aligned}$$

whence

$$\|e^{-\lambda t}\frac{\partial u}{\partial t}\|_{L^2(0,T;V_0^*)} \leq \sqrt{2}\frac{\mu}{\sigma_{min}}\|u_0\|.$$

□

Chapter 3

Discretization of the Black-Scholes Equation

For the discretization of the variational formulation (2.17a),(2.17b) of the Black-Scholes equation we use *Rothe's method*, i.e., we first consider a semidiscretization in time by the implicit Euler scheme which amounts to the solution of an elliptic subproblem for each time step. The elliptic subproblems are then approximated by continuous, piecewise linear finite elements with respect to simplicial triangulations of the spatial domain Ω.

3.1 Semidiscretization in Time

We consider a partition of the interval $[0, T]$ into subintervals $[t_{n-1}, t_n]$, $1 \leq n \leq N$, such that

$$0 = t_0 < t_1 < \cdots < t_N = T.$$

Set $\Delta t_n := t_n - t_{n-1}, \Delta t := \max\{\Delta t_n \mid 1 \leq n \leq N\}$ and

$$\rho_{\Delta t} := \max_{2 \leq n \leq N} \frac{\Delta t_n}{\Delta t_{n-1}}. \tag{3.1}$$

For continuous function f on $[0,T]$, we introduce the notation $f^n = f(t_n)$. The semidiscrete problem arising from the implicit Euler scheme is as follows: Find $(u^n)_{0 \leq n \leq N} \in V_0$ such that

$$(u^n - u^{n-1}, v)_{0,\Omega} + \Delta t_n a_{t_n}(u^n, v) = 0, \quad v \in V_0 \ , \ 1 \leq n \leq N, \tag{3.2a}$$

$$u^0 \ = u_0. \tag{3.2b}$$

The existence and uniqueness of the solution $u^n \in V_0$ of (3.2a),(3.2b) can be shown for sufficiently small time step Δt_n.

Theorem 3.1. *Under the assumptions* $(\mathbf{A_4}), (\mathbf{A_5})$ *and the time step restriction*

$$\Delta t_n < \frac{1}{2\lambda} \tag{3.3}$$

the semidiscrete problem (3.2a),(3.2b) *admits a unique solution.*

Proof. We note that (3.2a) can be equivalently written as

$$c_n(u^n, v) = (u^{n-1}, v)_{0,\Omega} \quad , \quad v \in V_0,$$

where the bilinear form $c_n(\cdot,\cdot) : V_0 \times V_0 \to \mathbb{R}$ is given by

$$c_n(v, w) = \Delta t_n a_{t_n}(v, w) + (v, w)_{0,\Omega} \quad , \quad v, w \in V_0.$$

Due to $(\mathbf{A_4}), (\mathbf{A_5})$, the bilinear form $c_n(\cdot,\cdot) : V_0 \times V_0 \to \mathbb{R}$ is bounded. Moreover, taking additionally (3.3) into account, it is V_0-elliptic as well, i.e., there exists a constant $\alpha > 0$ such that

$$c_n(v, v) \geq \alpha \|v\|_V^2 \quad , \quad v \in V_0.$$

Hence, the assertion follows from the Lax-Milgram Lemma (cf., e.g., [13]). □

For the sequence $(u^m)_{1\leq m\leq n}, n \leq N$, we introduce a discrete norm $[[u^m]]_n$ as the discrete analogue of $[[u]](t)$ (cf. (2.23)) according to

$$[[u^m]]_n := \Big(\Big(\prod_{i=1}^{n} (1 - 2\lambda \Delta t_i) \Big) \|u^n\|_{0,\Omega}^2 + \tag{3.4}$$

$$\frac{1}{2}\sigma_{min}^2 \sum_{m=1}^{n} \Delta t_m \Big(\prod_{i=1}^{m-1} (1 - 2\lambda \Delta t_i) \Big) |u^m|_V^2 \Big)^{1/2}. \tag{3.5}$$

As a counterpart of (2.24) we obtain:

Lemma 3.2. *Under the assumptions of Theorem 3.1 there holds*

$$[[u^m]]_n \leq \|u^0\|_{0,\Omega}. \tag{3.6}$$

Proof. By Young's inequality we have

$$(1 - 2\lambda \Delta t_n) \|u^n\|_{0,\Omega}^2 + \frac{1}{2} \Delta t_n \sigma_{min}^2 |u^n|_V^2 \leq \|u^{n-1}\|_{0,\Omega}^2. \tag{3.7}$$

Multiplication of (3.7) by $\prod_{i=1}^{n-1}(1 - 2\lambda\Delta t_i)$ and summation over n gives the assertion. □

Given the sequence $(u_n)_{1\leq n\leq N}$ of solutions of (3.2a),(3.2b), we introduce the function $u_{\Delta t}$ on $[0, T]$ by

$$u_{\Delta t}(t)|_{[t_{n-1},t_n]} := u_{n-1} + (\Delta t_n)^{-1}(t - t_{n-1})(u_n - u_{n-1}), \quad 1 \leq n \leq N, \tag{3.8}$$

which obviously is affine on each interval $[t_{n-1}, t_n], 1 \leq n \leq N$.

The following result establishes the equivalence of $[[u^m]]_n$ and $[[u_{\Delta t}]](t_n)$ which will be used later in chapter 4.

Lemma 3.3. *Suppose that* $(\mathbf{A_4}), (\mathbf{A_5})$ *hold true. Then, there exists a positive constant* $\alpha \leq \frac{1}{2}$ *such that for*

$$\Delta t \leq \frac{\alpha}{\lambda} \tag{3.9}$$

and for any family $(v^n)_{0\le n\le N} \in V_0^{N+1}$ *there holds*

$$\frac{1}{8}[[(v^m)]]_n^2 \le [[v_{\Delta t}]]^2(t_n) \le max(2, 1+\rho_{\Delta t})[[(v^m)]]_n^2 + \frac{1}{2}\sigma_{min}^2\Delta t_1|v^0|_V^2, \tag{3.10}$$

where $\rho_{\Delta t}$ *is given by* (3.1).

Proof. (i) Proof of the left inequality: In view of (3.8) we have

$$\frac{e^{2\lambda t_{m-1}}}{\Delta t_m}\int_{t_{m-1}}^{t_m} e^{-2\lambda\tau}|v_{\Delta t}|_V^2(\tau)d\tau = \tag{3.11}$$
$$\int_0^1 e^{-2\lambda\Delta t_m\tau}(|v^m|_V^2\tau^2 + |v^{m-1}|_V^2(1-\tau)^2 + 2(v^{m-1}, v^m)_V^*\tau(1-\tau))d\tau.$$

where

$$(v^{m-1}, v^m)_V^* := \int_\Omega S\frac{\partial v^{m-1}}{\partial S}S\frac{\partial v^m}{\partial S}dS.$$

If $\Delta t_m = 0$, for the right-hand side in (3.11) we obtain

$$\frac{1}{3}(|v^m|_V^2 + ||_V^2 + (v^m, v^{m-1})_V),$$

which can be estimated from below by $\frac{1}{4}|v^m|_V^2$ due to the inequality $ab \ge -\frac{a^2}{4} - b^2$.

If $\Delta t_m \ne 0$, we use that $e^{-2\lambda\Delta t_m\tau}$ is continuous with respect to τ. The mean value theorem of integral calculus implies the existence of a positive constant $\hat\tau \in (0,1]$ such that the right-hand side in (3.11) is equal to

$$e^{-2\lambda\Delta t_m\hat\tau}\int_0^1(|u^m|_V^2\tau^2 + |u^{m-1}|_V^2(1-\tau)^2 + 2(u^{m-1}, u^m)_V^*\tau(1-\tau))d\tau.$$

Consequently, there exists a constant $\alpha \le \frac{1}{2}$ such that for $\Delta t \le \frac{\alpha}{\lambda}$ there holds

$$\frac{e^{2\lambda t_{m-1}}}{\Delta t_m}\int_{t_{m-1}}^{t_m} e^{-2\lambda\tau}|u_{\Delta t}|_V^2(\tau)d\tau \ge \frac{1}{8}||_V^2,$$

whence

$$\int_{t_{m-1}}^{t_m} e^{-2\lambda\tau}|u_{\Delta t}|_V^2(\tau)d\tau \le \frac{\Delta t_m}{8}e^{-2\lambda t_{m-1}|v^m|_V^2}. \tag{3.12}$$

By Taylor Expansion, for $2\lambda\Delta t < 1$ we get

$$\prod_{i=1}^{m-1}(1-2\lambda\Delta t_i) \le e^{-2\lambda t_{m-1}},$$

which together with (3.12) yields

$$\frac{1}{2}\sigma_{min}^2\sum_{m=1}^{n}\Delta t_m(\prod_{i=1}^{m-1}(1-2\lambda\Delta t_i))|v^m|_V^2 \le 8(\frac{1}{2}\sigma_{min}^2\int_0^{t_n}e^{-2\lambda\tau}|u_{\Delta t}|_V^2(\tau)d\tau). \tag{3.13}$$

Since $\|v_{\Delta t}^{t_n}\| = \|v^n\|$, it follows that

$$(\prod_{i=1}^{m-1}(1-2\lambda\Delta t_i))\|v^n\|^2 \le e^{-2\lambda t_n}\|v_{\Delta t}^n\|. \tag{3.14}$$

The inequalities (3.13) and (3.14) give rise to the upper bound for $[[(v^m)]]_n$.

(ii) Proof of the right inequality: In view of the identity (3.11), we obtain the estimate

$$\begin{aligned}
&\frac{e^{2\lambda t_{m-1}}}{\Delta t_m}\int_{t_{m-1}}^{t_m}e^{-2\lambda\tau}|u_{\Delta t}|_V^2(\tau)d\tau \\
&\le |v^m|_V^2\int_0^1 e^{-2\lambda\Delta t_m}\tau d\tau + |v^{m-1}|_V^2\int_0^1 e^{-2\lambda\Delta t_m}(1-\tau)d\tau \\
&\le \frac{1}{2}(|v^{m-1}|_V^2 + |v^m|_V^2),
\end{aligned}$$

from which it follows that

$$\int_0^{t_n}e^{-2\lambda\tau}|u_{\Delta t}|_V^2(\tau)d\tau \le \frac{1}{2}\sum_{m=1}^{n}\Delta t_m e^{2\lambda t_{m-1}}(|v^{m-1}|_V^2 + |v^m|_V^2)$$

We can find a constant $\alpha_2 < \frac{1}{2}$ such that

$$\Delta t \le \frac{\alpha_2}{\lambda}.$$

Taylor expansion for $e^{-2\lambda\Delta t_i}$ $(i = 1, \ldots, m-1)$ gives

$$e^{-2\lambda t_{m-1}} \le 2\prod_{i=1}^{m-1}(1-2\lambda\Delta t_i), \tag{3.15}$$

from which we deduce

$$\int_0^{t_n} e^{-2\lambda\tau}|v_{\Delta t}|_V^2(\tau)d\tau \leq \sum_{n}^{m-1} \Delta t_m \prod_{i=1}^{m-1}(1-2\lambda\Delta t_i)(|v^{m-1}|_V^2+|v^m|_V^2)$$

$$\leq (1+\rho_{\Delta t})\sum_{m=1}^{n} \Delta t_m \prod_{i=1}^{m-1}(1-2\lambda\Delta t_i)|v^m|_V^2 + \Delta t_1|v^0|_V^2.$$

For $\Delta t \leq \frac{\alpha_2}{\lambda}$ we have

$$e^{-2\lambda t_{m-1}}\|v^n\|^2 \leq 2\prod_{i=1}^{m-1}(1-2\lambda\Delta t_i)\|v^n\|^2.$$

We conclude by choosing $\alpha := \min(\alpha_1, \alpha_2)$. □

From (3.6) and (3.10), we get the following relation for all n, $1 \leq n \leq N$,

$$[[u_{\Delta t}]](t_n) \leq C(u_0), \tag{3.16}$$

where

$$C(u_0) = (max(2, 1+\rho_{\Delta t})\|u_0\|^2 + \frac{1}{2}\sigma_{min}^2\Delta t_1|u_0|_V^2)^{\frac{1}{2}}. \tag{3.17}$$

3.2 Fully Discretized Problem

Given a null sequence $\mathcal{H}$ of positive real numbers, for the discretization of the semidiscrete problems (3.2a),(3.2b) in space, we use continuous, piecewise linear finite elements with respect to a family of simplicial triangulations $\mathcal{T}_{nh}, 1 \leq n \leq N$, of Ω. For $T \in \mathcal{T}_{nh}$, we denote by $S_{min}(T), S_{max}(T)$ the endpoints of T and refer to $h_T := S_{max}(T) - S_{min}(T)$ as the length of T and to $h_n := \max\{h_T \mid T \in \mathcal{T}_{nh}\}$ as the maximal size of the intervals in $\mathcal{T}_{nh}$. Moreover, for $D \subseteq \Omega$ we refer to $\mathcal{N}_{nh}(D)$ as the set of nodes of $\mathcal{T}_{nh}$ in D and associate with each $T \in \mathcal{T}_{nh}$ the patch ω_T according to

$$\omega_T := \bigcup\{T' \in \mathcal{T}_{nh} \mid \mathcal{N}_{nh}(T') \cap \mathcal{N}_{nh}(T) \neq \emptyset\}. \tag{3.18}$$

We assume that the family of triangulations is locally quasi-uniform in the sense that there exists a constant $\rho > 0$ such that for two adjacent elements $T, T' \in \mathcal{T}_{nh}$ there holds

$$h_T \le \rho h_{T'} \quad , \quad h \in \mathcal{H}. \tag{3.19}$$

For each $h \in \mathcal{H}$, we define the finite element spaces by

$$V_{nh} := \{v_h^n \in C^0(\Omega) \mid v_h^n|_T \in P^1(T), T \in \mathcal{T}_{nh}\}, \tag{3.20a}$$

$$V_{nh}^0 := V_{nh} \cap V_0, \tag{3.20b}$$

where $P^1(T)$ stands for the linear space of polynomials of degree 1 on T.

Assuming that $u_0 \in V_{1h}$, the fully discrete problem reads as follows: Find $(u_h^n)_{1\le n\le N}$, $u_h^n \in V_{nh}^0, 1 \le n \le N$, such that

$$(u_h^n - u_h^{n-1}, v_h)_{0,\Omega} + \Delta t_n a_{t_n}(u_h^n, v_h) = 0 \quad , \quad v_h \in V_{nh}^0, \tag{3.21a}$$

$$u_h^0 = u_0, \tag{3.21b}$$

Theorem 3.4. *Assume that* $(\mathbf{A_4}), (\mathbf{A_5})$ *and* (3.3) *hold true. Then, the fully discrete problem admits a unique solution. Moreover, for the sequence* $(u_h^m)_{1\le m\le n}, 1 \le n \le N$, *we have the stability estimate*

$$[[(u_h^m)]]_n \le \|u^0\|_{0,\Omega}. \tag{3.22}$$

Proof. Existence and uniqueness follow from the Lax-Milgram Lemma, since $V_{nh}^0 \subset V_0, 1 \le n \le N$. The estimate is an immediate consequence of Lemma 3.2. □

As in section 3.1 (cf. (3.8)) we define $u_{h,\Delta t}$ as the piecewise affine function

$$u_{h,\Delta t}(t)|_{[t_{n-1},t_n]} := P_h^n u_h^{n-1} + (\Delta t_n)^{-1}(t - t_{n-1})(u_h^n - P_h^n u_h^{n-1}), \quad 1 \le n \le N, \tag{3.23}$$

where $P_h^n u_h^{n-1}$ is the L^2-projection of u_h^{n-1} onto V_{nh}^0.

Chapter 4

A Posteriori Error Analysis

In this chapter, we will provide a *residual-type a posteriori error estimation* for the error

$$[[u - u_{h,\Delta t}]](t_n) \quad , \quad 1 \leq n \leq N,$$

where $[[\cdot]](t)$ is the norm given by (2.23). It consists of computable error estimators reflecting the contributions to the error due to the discretizations in time by the implicit Euler scheme and in space by the finite element approximation described in the previous sections 3.1 and 3.2. The error estimators will be presented in section 4.1 followed by an *a posteriori error analysis* which establishes the *reliability* of the estimators in section 4.2 and their *efficiency* in section 4.3.

The a posteriori error analysis requires more assumptions on the data of the problem:

($\mathbf{A_6}$) The functions σ and $S\frac{\partial \sigma}{\partial S}$ are Lipschitz continuous on $[0, T]$ uniformly in $S \in \bar{\Omega}$, and the function r is Lipschitz continuous on $[0, T]$. In particular, there exists positive constants C_i, $1 \le i \le 3$, such that for all t, $t' \in [0,T]$ there holds

$$\|\rho^2(\cdot,t) - \rho^2(\cdot,t')\|_{L^\infty(0,\bar{S})} \le C_1 |t-t'|, \tag{4.1}$$

$$\| - r(t) + r(t') + \frac{\rho^2(\cdot,t) - \rho^2(\cdot,t')}{2} + S(\rho(\cdot,t)\frac{\partial \rho}{\partial S}(\cdot,t) - \rho(\cdot,t')\frac{\partial \rho}{\partial S}(\cdot,t'))\|_{L^\infty(0,\bar{S})}$$

$$\le C_2 |t-t'|, \tag{4.2}$$

$$|r(t) - r(t')| \le C_3 |t-t'|. \tag{4.3}$$

Throughout this chapter, for quantities $A, B \in \mathbb{R}_+$ we will use the notation $A \lesssim B$, if there exists a constant $c > 0$, independent of Δt_n and $h_T, T \in \mathcal{T}_{nh}, 1 \le n \le N, h \in \mathcal{H}$, such that $A \le cB$.

4.1 The A Posteriori Error Estimator

For the fully discretized Black-Scholes equations (3.21a) and (3.21b), the global discretization error $u - u_{h,\Delta t}$ can be assessed by a *time error estimator* and a *price error estimator*. The *time error estimator* is local in time and global in price. It is given by

$$\eta_n := \sqrt{\Delta t_n}\; e^{-\lambda t_{n-1}}\; \frac{\sigma_{min}}{\sqrt{2}}\; |u_h^n - u_h^{n-1}|_V \quad , \quad 1 \le n \le N, \tag{4.4}$$

where $\sigma_{min} > 0$ and $\lambda \ge 0$ are the constants from the ellipticity assumption (2.18a) and Gårding's inequality (2.20).

On the other hand, the *price error estimator* is local both in time and price. It is given by

$$\eta_{n,T} := \frac{h_T}{S_{max}(T)}\; \|R_T(u_h^{n-1}, u_h^n)\|_{0,T} \quad , \quad T \in \mathcal{T}_{nh}\ ,\ 1 \le n \le N, \tag{4.5}$$

where $R_T(u_h^{n-1}, u_h^n)$ stands for the *residual* with respect to the strong form (2.10a)-(2.10c) of the Black-Scholes equation

$$R_T(u_h^{n-1}, u_h^n) := \frac{u_h^n - u_h^{n-1}}{\Delta t_n} - rS\frac{\partial u_h^n}{\partial S} + ru_h^n. \tag{4.6}$$

Remark 4.1. *Since $u_h^n \in V_{nh}^0$ is piecewise linear, we have $\partial^2 u_h^n/\partial S^2|_T \equiv 0, T \in \mathcal{T}_{nh}(\Omega)$, and hence, this term does not occur in* (4.6)*. However, if higher-order finite elements are used for the discretization in space, this term has to be included in the residual, i.e., the residual then reads*

$$R_T(u_h^{n-1}, u_h^n) := \frac{u_h^n - u_h^{n-1}}{\Delta t_n} - \frac{\sigma^2 S^2}{2} \frac{\partial^2 u_h^n}{\partial S^2} - rS\frac{\partial u_h^n}{\partial S} + ru_h^n. \tag{4.7}$$

Remark 4.2. *Compared to residual-type a posteriori error estimators derived for parabolic initial-boundary value problems on bounded domains of dimension ≥ 2 (cf., e.g., [20]), the price error term $\eta_{n,T}$ does not contain jumps $[\partial u_h^n/\partial S]_{S_i}$ of the derivatives $\partial u_h^n/\partial S$ in the nodal points $S_i \in \mathcal{N}_{nh}(\Omega), 1 \leq i \leq N_{nh}$. The reason is as follows: For conforming finite element discretizations, residual-type a posteriori error estimators are usually derived by taking advantage of Galerkin orthogonality (cf.* (4.25) *below) and by using suitable interpolation or quasi-interpolation operators $P_h^n : V_0 \to V_{nh}^0$ with specific stability and local approximation properties such as the* Scott-Zhang interpolation operator *[16] or the* Clément quasi-interpolation operator *[40]. The standard* interpolation operator *$I_h^n : V_0 \to V_{nh}^0$ from finite element analysis [16] can not be used, since V_0 is not continuously embedded in $C^0(\bar{\Omega})$. The situation is different, however, in one space dimension, where due to the* Sobolev embedding theorem *[16] the embedding $V_0 \to C^0(\bar{\Omega})$ is continuous indeed. In particular, it will be shown in section 4.2 below (cf. Lemma 4.5 and Proposition 4.6) that an interpolation operator can be constructed such that the jump terms vanish in the evaluation of the residuals (cf.* (4.27)*). However, if the* Scott-Zhang interpolation operator *or the* Clément quasi-interpolation operator *is used instead, the jump terms do not vanish and enter the price error term $\eta_{n,T}$ according to*

$$\eta_{n,T} := \frac{h_T}{S_{max}(T)} \ \|R_T(u_h^{n-1}, u_h^n)\|_{0,T} + \frac{h_T^{1/2}}{4} \sum_{S_i \in \mathcal{N}_{nh}(T) \cap \mathcal{N}_{nh}(\Omega)} \sigma^2(t_n, , S_i) S_i^2 |[\frac{\partial u_h^n}{\partial S}]_{S_i}|. \tag{4.8}$$

We emphasize that for parabolic problems in spatial domains of dimension ≥ 2 such as *the Black-Scholes equation for European basket options [37] the jumps* $[\mathbf{n} \cdot \nabla u_h^n]_F$ *of the normal derivatives* $\mathbf{n} \cdot \nabla u_h^n$ *across interior faces* F *of the simplicial triangulation of the computational domain always enter the price error term.*

4.2 Reliability of the Estimators

The main result of this section is the reliability of the error estimators.

Theorem 4.3. *Under the assumptions* $(\mathbf{A_4}) - (\mathbf{A_6})$ *and* $u_0 \in V_{1h}^0$ *let* $u \in H^1((0,T); V_0^*) \cap L^2((0,T); V_0)$ *be the solution of* (2.17a),(2.17b) *and let* $u_{h,\Delta t}$ *be given by* (3.23) *in terms of the solution* $(u_h^n)_{0 \leq n \leq N}$ *of the fully discrete problem* (3.21),(3.21a). *Moreover, let* η_n *and* $\eta_{n,\omega}$ *be the time error and price error estimators given by* (4.4) *and* (4.5), *respectively. Then, there exists a positive constant* $\alpha \leq \frac{1}{2}$ *such that for* $\lambda \Delta t \leq \alpha$ *there holds*

$$[[u - u_{h,\Delta t}]](t_n) \lesssim \tag{4.9}$$
$$\Big(\frac{C}{\sigma_{min}^2} C(u_0 \Delta t) + \frac{\mu}{\sigma_{min}^2} \Big(\sum_{m-1}^{n} \eta_m^2 + \frac{\Delta t_m}{\sigma^2 min} \kappa(\rho_{\Delta t}) \prod_{i=1}^{m-1} (1 - 2\lambda \Delta t_i) \sum_{T \in \mathcal{T}_{nh}} \eta_{m,T}^2 \Big)^{1/2} \Big),$$

where $C := 4C_1 + 2C_2 + C_3$, $C(u_0)$ *is given by* (3.17) *and*

$$\kappa(\rho_{\Delta t}) := (1 + \rho_{\Delta t})^2 ||u_0||^2 + max(2, 1 + \rho_{\Delta t}). \tag{4.10}$$

The proof of Theorem 4.3 will be provided by splitting the error $[[u - u_{h,\Delta t}]](t_n)$ according to

$$[[u - u_{h,\Delta t}]](t_n) \leq [[u - u_{\Delta t}]](t_n) + [[u_{\Delta t} - u_{h,\Delta t}]](t_n) \tag{4.11}$$

and to estimate the two terms on the right-hand side separately.

Proposition 4.4. *Under the assumptions of Theorem 4.3 there exists a constant* $\alpha \leq \frac{1}{2}$

such that for $\Delta t \le \frac{\alpha}{\lambda}$ *there holds*

$$[[u - u_{\Delta t}]](t_n) \lesssim \Big(\frac{L}{\sigma^2_{min}}\ C(u_0)\Delta t + [[u_{\Delta t} - u_{h,\Delta t}]](t_n) + \frac{\mu}{\sigma^2_{min}} (\sum_{m=1}^{n} \eta^2_m \Big)^{1/2} \Big). \tag{4.12}$$

Proof. For any $t \in (t_{n-1}, t_n]$ and $v \in V_0$ we have

$$\begin{aligned}
&(\frac{\partial}{\partial t} u_{\Delta t}(t), v) + a_t(u_{\Delta t}(t), v) \\
&= (\frac{u^n - u^{n-1}}{\Delta t_n}, v) + a_t(u_{\Delta t}(t), v) - a_{t_n}(u_{\Delta t}(t), v) + \\
&a_{t_n}(u_{\Delta t}(t) - u^n, v) + a_{t_n}(u^n, v).
\end{aligned}$$

Subtracting the previous equation from (2.17a) it follows that

$$\begin{aligned}
&(\frac{\partial}{\partial t}(u - u_{\Delta t})(t), v) + a_t((u - u_{\Delta t})(t), v) = -(\frac{u^n - u^{n-1}}{\Delta t_n}, v) - \\
&a_t(u_{\Delta t}(t), v) + a_{t_n}(u_{\Delta t}(t), v) - a_{t_n}(u_{\Delta t}(t) - u^n, v) - a_{t_n}(u^n, v).
\end{aligned}$$

In view of (3.2a) we obtain

$$\begin{aligned}
&(\frac{\partial}{\partial t}(u - u_{\Delta t})(t), v) + a_t((u - u_{\Delta t})(t), v) \\
&= -a_t(u_{\Delta t}(t), v) + a_{t_n}(u_{\Delta t}(t), v) - a_{t_n}(u_{\Delta t}(t) - u^n, v).
\end{aligned} \tag{4.13}$$

We choose $v(t) = (u - u_{\Delta t})(t)e^{-2\lambda t}$ and integrate the first term on the right-hand side of (4.13) over (t_{m-1}, t_m) which results in

$$\begin{aligned}
&\int_{t_{m-1}}^{t_m} \frac{\partial}{\partial t}(u - u_{\Delta t})(t)(u - u_{\Delta t})(t)e^{-2\lambda} dt = \\
&\frac{1}{2}\Big((u - u_{\Delta t})^2(t)e^{-2\lambda t}|_{t_{m-1}}^{t_m} + 2\lambda \int_{t_{m-1}}^{t_m} (u - u_{\Delta t})^2(t)e^{-2\lambda t} dt \Big).
\end{aligned}$$

Summation over m from $m = 1$ to $m = n$ and observing $(u - u_{\Delta t})(t_0) = 0$ yields

$$\begin{aligned}
&\sum_{m=1}^{n} \int_{t_{m-1}}^{t_m} (\frac{\partial}{\partial t}(u - u_{\Delta t})(t), (u - u_{\Delta t})(t)e^{-2\lambda}) dt = \\
&\frac{1}{2}\|u - u_{\Delta t}\|^2(t_n)e^{-2\lambda t_n} + \lambda \sum_{m=1}^{n} \int_{t_{m-1}}^{t_m} \|u - u_{\Delta t}\|^2(t)e^{-2\lambda t} dt].
\end{aligned} \tag{4.14}$$

Integrating both sides of (4.13) over (t_{m-1}, t_m) and summing up, we get

$$\sum_{m=1}^{n} \int_{t_{m-1}}^{t_m} (\frac{\partial}{\partial t}(u - u_{\Delta t})(t), v) dt = \tag{4.15}$$
$$\sum_{m=1}^{n} (- \int_{t_{m-1}}^{t_m} a_\tau((u - u_{\Delta t})(\tau), v) d\tau -$$
$$\int_{t_{m-1}}^{t_m} (a_\tau(u_{\Delta t}(\tau), v) - a_\tau(u_{\Delta t}(\tau), v)) d\tau - \int_{t_{m-1}}^{t_m} a_{t_m}(u_{\Delta t}(\tau) - u^m, v) d\tau.$$

Combining (4.14) and (4.15), it follows that

$$\frac{1}{2}\|u - u_{\Delta t}\|^2(t_n)e^{-2\lambda t_n} + \lambda \sum_{m=1}^{n} \int_{t_{m-1}}^{t_m} \|u - u_{\Delta t}\|^2(t)e^{-2\lambda t} dt] = \tag{4.16}$$
$$\sum_{m=1}^{n} (- \int_{t_{m-1}}^{t_m} a_\tau((u - u_{\Delta t})(\tau), v) d\tau - \int_{t_{m-1}}^{t_m} (a_\tau(u_{\Delta t}(\tau), v) - a_\tau(u_{\Delta t}(\tau), v)) d\tau -$$
$$\int_{t_{m-1}}^{t_m} a_{t_m}(u_{\Delta t}(\tau) - u^m, v) d\tau.$$

Due to *Gårding's inequality* there holds

$$\sum_{m=1}^{n} \int_{t_{m-1}}^{t_m} a_\tau((u - u_{\Delta t})(\tau), v) d\tau \geq \tag{4.17}$$
$$\sum_{m=1}^{n} (\frac{1}{4} \int_{t_{m-1}}^{t_m} \sigma_{min}^2 |u - u_{\Delta t}|_V^2 e^{-2\lambda\tau} d\tau - \lambda \int_{t_{m-1}}^{t_m} \|u - u_{\Delta t}\|^2 e^{-2\lambda\tau} d\tau).$$

Using (4.16) and (4.17), we get

$$[[u - u_{\Delta t}]]^2(t_n) \leq \tag{4.18}$$
$$- 2 \sum_{m=1}^{n} \int_{t_{m-1}}^{t_m} (a_\tau(u_{\Delta t}(\tau), v) - a_{t_m}(u_{\Delta t}(\tau), v)) d\tau$$
$$- 2 \sum_{m=1}^{n} \int_{t_{m-1}}^{t_m} a_{t_m}(u_{\Delta t}(\tau) - u^m, v) d\tau.$$

Now, we will evaluate each term on the right-hand side of (4.18). Using the assumptions

$(\mathbf{A_6}), (\mathbf{A_7})$ and $(\mathbf{A_8})$, it follows that

$$
\begin{aligned}
&|\int_{t_{m-1}}^{t_m} (a_\tau(u_{\Delta t}(\tau), v) - a_{t_m}(u_{\Delta t}(\tau), v))d\tau| \qquad (4.19)\\
&\le \Delta t_m[\frac{2C_1}{\sigma_{min}^2}\int_{t_{m-1}}^{t_m} \frac{\sigma_{min}^2}{2}|(S\frac{\partial u_{\Delta t}}{\partial S}, S\frac{\partial v}{\partial S})|\\
&+ \frac{2C_2}{\sigma_{min}^2}\int_{t_{m-1}}^{t_m} \sigma_{min}^2|(S\frac{\partial u_{\Delta t}}{\partial S}, v)| + \frac{2L_3}{\sigma_{min}^2}\int_{t_{m-1}}^{t_m} \sigma_{min}^2|(u_{\Delta t}, v)|]\\
&\le \frac{2C_1 + C_2 + C_3/2}{\sigma_{min}^2}\int_{t_{m-1}}^{t_m} \frac{\sigma_{min}^2}{2}|u_{\Delta t}|_V|u - u_{\Delta t}|_V e^{-2\lambda\tau}d\tau\\
&\le \frac{2C_1 + C_2 + C_3/2}{\sigma_{min}^2}(\int_{t_{m-1}}^{t_m} \frac{\sigma_{min}^2}{2}|u_{\Delta t}|_V^2 e^{-2\lambda\tau}d\tau)^{1/2}\\
&\cdot (\int_{t_{m-1}}^{t_m} \frac{\sigma_{min}^2}{2}|u - u_{\Delta t}|_V^2 e^{-2\lambda\tau}d\tau)^{1/2}.
\end{aligned}
$$

Setting $C := 4C_1 + 2C_2 + C_3$ and taking

$$
C(u_0) \ge [[u_{\Delta t}]](t_n) \ge (\frac{1}{2}\sigma_{min}^2\int_0^{t_n} e^{-2\lambda\tau}|u_{\Delta t}(\tau)|_V^2 d\tau)^{1/2}
$$

into account, we obtain

$$
\begin{aligned}
&2|\sum_{m=1}^{n}\int_{t_{m-1}}^{t_m} (a_\tau(u_{\Delta t}(\tau), v) - a_{t_m}(u_{\Delta t}(\tau), v))d\tau| \qquad (4.20)\\
&\le \frac{L}{\sigma_{min}^2}C(u_0)\Delta t(\sum_{m=1}^{n}\int_{t_{m-1}}^{t_m} \frac{\sigma_{min}^2}{2}|u - u_{\Delta t}|_V^2 e^{-2\lambda\tau}d\tau)^{1/2}\\
&\le \frac{L}{\sigma_{min}^2}C(u_0)\Delta t[[u - u_{\Delta t}]]t_n.
\end{aligned}
$$

For the second term on the right-hand side of (4.19), Lemma 2.5 and the Cauchy-Schwarz inequality give

$$
\begin{aligned}
&|\int_{t_{m-1}}^{t_m} a_{t_m}(u_{\Delta t}(\tau) - u^m, v)d\tau|\\
&\le \mu(\int_{t_{m-1}}^{t_m} |u_{\Delta t}(\tau) - u^m|_V^2 e^{-2\lambda\tau}d\tau)^{1/2}(\int_{t_{m-1}}^{t_m} |v|_V^2 e^{2\lambda\tau}d\tau)^{1/2}\\
&\le \frac{\sqrt{2}\mu}{\sigma_{min}}(\int_{t_{m-1}}^{t_m} |u_{\Delta t}(\tau) - u^m|_V^2 e^{-2\lambda\tau}d\tau)^{1/2}(\int_{t_{m-1}}^{t_m} \frac{\sigma_{min}^2}{2}|u - u_{\Delta t}|_V^2 e^{2\lambda\tau}d\tau)^{1/2}.
\end{aligned}
$$

Since

$$u_{\Delta t}(\tau) - u^m = \frac{t_m - \tau}{\Delta t_m}(u^{m-1} - u^m),$$

we have

$$\begin{aligned}
&(\int_{t_{m-1}}^{t_m} |u_{\Delta t}(\tau) - u^m|_V^2 e^{-2\lambda\tau} d\tau)^{\frac{1}{2}} \\
&= |u^{m-1} - u^m|_V (\int_{t_{m-1}}^{t_m} (\frac{t_m - \tau}{\Delta t_m})^2 e^{-2\lambda\tau} d\tau)^{\frac{1}{2}} \\
&\leq (\frac{\Delta t_m}{3})^{\frac{1}{2}} e^{-\lambda t_{m-1}} |u^{m-1} - u^m|_V,
\end{aligned}$$

whence

$$\begin{aligned}
&\Big(\int_{t_{m-1}}^{t_m} |u_{\Delta t}(\tau) - u^m|_V^2 e^{-2\lambda\tau} d\tau\Big)^{\frac{1}{2}} \\
&\leq \Delta t_m e^{-2\lambda t_{m-1}} \Big(|u_h^{m-1} - u_h^m|_V^2 + |u^{m-1} - u_h^{m-1}|_V^2 + |u^m - u_h^m|_V)^2\Big).
\end{aligned} \tag{4.21}$$

Using (3.15) and (3.10), for the sum over m of the last two terms on the right-hand side in (4.21) we find

$$\begin{aligned}
&2(1 + \rho_{\Delta t}) \sum_{m=1}^{n} \Delta t_m \prod_{i=1}^{m-1} (1 - 2\lambda \Delta t_i) |u^m - u_h^m|_V^2 \\
&\leq \frac{32}{\sigma_{min}^2}(1 + \rho_{\Delta t})[[u_{\Delta t} - u_{h,\Delta t}]]^2(t_n),
\end{aligned}$$

from which we derive the following upper bound for the last term of the right-hand side of (4.18),

$$\begin{aligned}
&|2 \sum_{m=1}^{n} \int_{t_{m-1}}^{t_m} a_{t_m}(u_{\Delta t}(\tau) - u^m, v) d\tau| \\
&\leq \frac{4\mu}{\sigma_{min}^2}(16(1 + \rho_{\Delta t})[[u_{\Delta t} - u_{h,\Delta t}]]^2(t_n) + \\
&\sum_{m=1}^{n} \Delta t_m e^{-2\lambda t_{m-1}} \frac{\sigma_{min}^2}{2} |u_h^m - u_h^{m-1}|_V^2)^{1/2} [[u - u_{\Delta t}]](t_n).
\end{aligned} \tag{4.22}$$

Finally, the desired estimate follows from (4.20) and (4.22). □

Referring to the discussion in Remark 4.2 above, for the estimation of the discretization error $u_{\Delta t} - u_{h,\Delta t}$ from above, an interpolation operator $I_h^n : V^0 \to V_{nh}^0$ with appropriate local approximation properties will be provided by the following lemma.

Lemma 4.5. *Let $S_i, 0 \le i \le N_{nh}$, be the grid points of the triangulation $\mathcal{T}_{nh}$ such that $0 = S_0 < S_1 < \cdots < S_{N_{nh}} = \bar{S}$ and define $I_h^n : V^0 \to V_{nh}^0$ according to*

$$I_h^n v(S_i) = v(S_i) \ , \ 1 \le i \le N_{nh} \quad , \quad \int_{S_0}^{S_1} (v - I_h^n v) dS = 0.$$

Then, for $T \in \mathcal{T}_{nh}$ there holds

$$\|v - I_h^n v\|_{0,T} \lesssim \frac{h_T}{S_{max}(T)} \ \|S \frac{\partial v}{\partial S}\|_{0,T}, \tag{4.23a}$$

$$\|S(v - I_h^n v)\|_{\infty,T} \lesssim h_T^{1/2} \ \|S \frac{\partial v}{\partial S}\|_{0,T}. \tag{4.23b}$$

Proof. The local approximation properties follow from standard linear interpolation in Sobolev spaces [16]. □

Proposition 4.6. *Under the assumptions of Theorem 4.3 there holds*

$$[[u_{\Delta t} - u_{h,\Delta t}]]^2(t_n) \lesssim \tag{4.24}$$
$$\sigma_{min}^{-2} \ max(2, 1 + \rho_{\Delta t}) \sum_{m=1}^{n} \Delta t_m \prod_{i=1}^{m-1} (1 - 2\lambda \Delta t_i) \sum_{\omega \in \mathcal{T}_{nh}} \eta_{m,\omega}^2.$$

Proof. Taking advantage of the Galerkin orthogonality

$$(u^n - u_h^n - (u^{n-1} - u_h^{n-1}), v_h)_{0,\Omega} + \Delta t_n a_{t_n}(u^n - u_h^n, v_h) = 0 \quad , \quad v_h \in V_{nh}^0, \tag{4.25}$$

for $v \in V^0$ and $v_h \in V_{nh}^0$ we deduce

$$(u^n - u_h^n, v - v_h)_{0,\Omega} + \Delta t_n a_{t_n}(u^n - u_h^n, v - v_h)$$
$$= (u^{n-1} - u_h^{n-1}, v - v_n)_{0,\Omega} + (u_h^n - u_h^{n-1}, v - v_h)_{0,\Omega} - \Delta t_n a_{t_n}(u_h^n, v - v_h).$$

Integration by parts yields

$$
\begin{aligned}
&(u^n - u_h^n, v - v_h)_{0,\Omega} + \Delta t_n a_{t_n}(u^n - u_h^n, v - v_h) \qquad (4.26)\\
&= (u^{n-1} - u_h^{n-1}, v - v_n)_{0,\Omega} + (u_h^n - u_h^{n-1}, v - v_h)_{0,\Omega} + \\
&\Delta t_n \sum_{T\in\mathcal{T}_{n,h}} \Big(\int_T (rS\frac{\partial u_h^n}{\partial S} - ru_h^n)(v - v_h)dS - \\
&\frac{1}{4} \sum_{S_i\in\mathcal{N}_{nh}(T)\cap\mathcal{N}_{nh}(\Omega)} \sigma^2(S_i, t_n) S_i^2 [\frac{\partial u_h^n}{\partial S}]_{S_i}(v - v_h)(S_i)\Big),
\end{aligned}
$$

Choosing $v_h = I_h^n v$ with the interpolation operator I_h^n from Lemma 4.5, we have

$$
-\frac{1}{4} \sum_{S_i\in\mathcal{N}_{nh}(T)\cap\mathcal{N}_{nh}(\Omega)} \sigma^2(S_i, t_n) S_i^2 [\frac{\partial u_h^n}{\partial S}]_{S_i}(v - v_h)(S_i) = 0, \qquad (4.27)
$$

whence

$$
\begin{aligned}
&| - \int_T \frac{u_h^n - u_h^{n-1}}{\Delta t_n}(v - v_h)dS + \int_w (rS\frac{\partial u_h^n}{\partial S} - ru_h^n)(v - v_h)dS \\
&- \frac{1}{4}\sum_{i=1}^{2} \sigma^2(\xi_i, t_n)\xi_i^2[\frac{\partial u_h^n}{\partial S}](\xi_i)(v - v_h)(\xi_i)| \\
&\lesssim (\frac{h_T}{S_{max}(T)} \| \frac{u_h^n - u_h^{n-1}}{\Delta t_n} - rS\frac{\partial u_h^n}{\partial S} + ru_h^n\|_{0,T}) \|S\frac{\partial v}{\partial S}\|_{0,T} \\
&\lesssim \eta_{n,T} \|S\frac{\partial v}{\partial S}\|_{0,T}.
\end{aligned}
$$

In particular, for $v = (u^n - u_h^n)$ we obtain

$$
\begin{aligned}
&(1 - \lambda\Delta t_n)\|u^n - u_h^n\|^2 + \frac{1}{4}\Delta_n\sigma_{min}^2|u^n - u_h^n|_V^2 \\
&\lesssim \frac{1}{2}\|u^{n-1} - u_h^{n-1}\|^2 + \frac{1}{2}\|u^n - u_h^n\|^2 + \frac{1}{8}\Delta t_n\sigma^2{}_{min}|u^n - u_h^n|_V^2) \\
&+ 2\sigma_{min}^{-2}\Delta t_n \sum_{T\in\mathcal{T}_{nh}} \eta_{n,T}^2 + \frac{1}{8}\Delta_{min}^2|u^n - u_h^n|_V^2.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&(1 - 2\lambda\Delta t_n)\|u^n - u_h^n\|^2 + \frac{1}{4}\Delta_n\sigma_{min}^2|u^n - u_h^n|_V^2 \\
&\lesssim \|u^{n-1} - u_h^{n-1}\|^2 + 4\sigma_{min}^{-2}\Delta t_n \sum_{T\in\mathcal{T}_{nh}} \eta_{n,T}^2.
\end{aligned}
$$

Multiplication of the previous equation by $\prod_{i=1}^{n-1}(1-2\lambda\Delta t_i)$ and summation over n yields

$$[[u^m - u_h^m]]_n^2 \lesssim \sigma_{min}^{-2} \sum_{m=1}^{n} \Delta t_m \prod_{i=1}^{m-1}(1-2\lambda\Delta t_i) \sum_{T\in\mathcal{T}_{mh}} \eta_{m,T}^2.$$

Finally, utilizing (3.10) we arrive at

$$[[(u_{\Delta t} - u_{h,\Delta t}]]^2(t_n) \lesssim \sigma_{min}^{-2}\max(2, 1+\rho_{\Delta t}) \sum_{m=1}^{n} \Delta t_m \prod_{i=1}^{m-1}(1-2\lambda\Delta t_i) \sum_{T\in\mathcal{T}_{mh}} \eta_{m,T}^2.$$

□

4.3 Efficiency of the Estimators

This section is devoted to the proofs of the efficiency of the estimator η_n and the local efficiency of the estimator $\eta_{n,T}$. The efficiency of η_n will be shown in subsection 4.3.1, whereas subsection 4.3.2 is devoted to the proof of the local efficiency of $\eta_{n,T}$.

4.3.1 Efficiency of the estimator η_n.

For $(v^n)_{1\le n\le N}, v^n \in V_0$ we introduce the norm

$$[[v^n]]^2 = \frac{\sigma_{min}^2}{2}\Delta t_n \prod_{i=1}^{n-1}(1-2\lambda\Delta t_i)|v^n|_V^2. \tag{4.28}$$

By means of this norm we now establish the efficiency of η_n.

Theorem 4.7. *Suppose that $u^0 \in V_0$ and $\lambda\Delta t \le \alpha$ as in Lemma 3.1. Then, for $2 \le n \le N$ there holds*

$$\begin{aligned}\eta_n^2 \lesssim\ & [[u^n - u_h^n]]^2 + \rho_{\Delta t}[[u^{n-1} - u_h^{n-1}]]^2 + \\ & \frac{e^{-2\lambda t_{n-1}}}{\sigma_{min}^2}\Big(||\frac{\partial}{\partial t}(u-u_{\Delta t})||^2_{L^2((t_{n-1},t_n);V_0^*)} + ||u-u_{\Delta t}||^2_{L^2((t_{n-1},t_n);V_0)}\Big) + \\ & \frac{max(1,\sigma_{\Delta t})}{\sigma_{min}^4}\ (\Delta t_n)^2||u^0||^2\Big),\end{aligned} \tag{4.29}$$

whereas for $n = 1$ we have

$$\eta_1^2 \lesssim [[u^1 - u_h^1]]^2 + \frac{1}{\sigma_{min}^2}\Big(||\frac{\partial}{\partial t}(u - u_{\Delta t})||^2_{L^2((0,t_1);V_0^*)} + ||u - u_{\Delta t}||^2_{L^2((0,t_1);V_0)}\Big) + \frac{(\Delta t_1)^2}{\sigma_{min}^6}\Big(\|u^0\|^2 + \sigma_{min}^2\ \Delta t_1 |u^0|_V^2\Big). \tag{4.30}$$

Proof. We split η_n according to

$$\eta_n \leq \sqrt{\Delta t_n} e^{-\lambda t_{n-1}} \frac{\sigma_{min}}{\sqrt{2}}\Big(|u^n - u_h^n|_V + |u^{n-1} - u_h^{n-1}|_V + |u^n - u^{n-1}|_V\Big), \tag{4.31}$$

whence

$$\eta_n^2 \leq 3\Delta t_n\ e^{-2\lambda t_{n-1}}\ \frac{\sigma_{min}^2}{2}\Big(|u^n - u_h^n|_V^2 + |u^{n-1} - u_h^{n-1}|_V^2 + |u^n - u^{n-1}|_V^2\Big), \tag{4.32}$$

Taking advantage of (3.15), for the first term on the right-hand side we obtain

$$3\Delta t_n e^{-2\lambda t_{n-1}} \frac{\sigma_{min}^2}{2} |u^n - u_h^n|_V^2 \leq 3\Delta t_n \prod_{i=1}^{n}(1 - 2\lambda t_i)\sigma_{min}^2 |u^n - u_h^n|_V^2 \leq 6\ [[u^n - u_h^n]]^2. \tag{4.33}$$

Likewise, for the second term it follows that

$$3\Delta t_n e^{-2\lambda t_{n-1}} \frac{\sigma_{min}^2}{2} |u^{n-1} - u_h^{n-1}|_V^2 \leq 6\rho_{\Delta t}[[u^{n-1} - u_h^{n-1}]]^2. \tag{4.34}$$

Using Gårding's inequality, for the third term we find

$$\begin{aligned} &3\Delta t_n e^{-2\lambda t_{n-1}} \frac{\sigma_{min}^2}{2} |u^n - u^{n-1}|_V^2 \\ &\leq 6\Delta t_n e^{-2\lambda t_{n-1}}\Big(a_{t_n}(u^n - u^{n-1}, u^n - u^{n-1}) + \lambda\|u^n - u^{n-1}\|^2\Big) \end{aligned}$$

In view of

$$u_{\Delta t}(\tau) = u^{n-1}\frac{t_n - \tau}{t_n - t_{n-1}} + u^n \frac{\tau - t_{n-1}}{t_n - t_{n-1}} \quad , \quad t_n \leq \tau \leq t_{n-1},$$

we see that

$$\int_{t_{n-1}}^{t_n} a_{t_n}(u_{\Delta t}(\tau) - u^n, u^n - u^{n-1}) d\tau = -\frac{\Delta t_n}{2} a_{t_n}(u^n - u^{n-1}, u^n - u^{n-1}).$$

Moreover, there holds

$$(\frac{\partial}{\partial\tau}u(\tau), u^n - u^{n-1}) + a_\tau(u(\tau), u^n - u^{n-1}) = 0,$$
$$(\frac{\partial}{\partial\tau}u_{\Delta t}(\tau), u^n - u^{n-1}) + a_{t_n}(u^n, u^n - u^{n-1}) = 0,$$

whence

$$\begin{aligned}
&6\Delta t_n e^{-2\lambda t_{n-1}} \frac{\sigma_{min}^2}{2} |u^n - u^{n-1}|_V^2 \leq \\
&\underbrace{24 e^{-2\lambda t_{n-1}} \int_{t_{n-1}}^{t_n} \frac{\partial}{\partial t}(u - u_{\Delta t})(\tau)(u^n - u^{n-1}) d\tau}_{=I} + \\
&\underbrace{24 e^{-2\lambda t_{n-1}} \int_{t_{n-1}}^{t_n} a_\tau(u - u_{\Delta t}, u^n - u^{n-1}) d\tau}_{=II} + \\
&\underbrace{24 e^{-2\lambda t_{n-1}} \int_{t_{n-1}}^{t_n} a_\tau(u_{\Delta t}, u^n - u^{n-1}) - a_{t_n}(u_{\Delta t}, u^n - u^{n-1}) d\tau}_{=III} + \\
&\underbrace{12\lambda\Delta t_n e^{-2\lambda t_{n-1}} \|u^n - u^{n-1}\|^2}_{=IV}.
\end{aligned} \tag{4.35}$$

We will estimate the four terms on the right-hand side of (4.35) separately. This will be done in the following four Lemmas. Using the results these lemmas in (4.35) together with (4.33) and (4.34) gives the assertions. □

In the derivation of upper bounds for the terms I, II, III and IV from (4.35), we will frequently make use of Young's inequality

$$ab \leq 1/(4\varepsilon)a^2 + \varepsilon b^2. \tag{4.36}$$

Lemma 4.8. *Under the assumptions of Theorem 4.7, for the term I from* (4.35) *we obtain* *the upper bound*

$$|I| \le \frac{1}{10}\ \eta_n^2 + 6[[u^n - u_h^n]]^2 + 6\rho_{\Delta t}[[u^{n-1} - u_h^{n-1}]]^2 + \frac{192}{\sigma_{min}^2} e^{-2\lambda t_{n-1}} \Big(1 + 60 e^{-2\lambda t_{n-1}}\Big) \| \frac{\partial}{\partial t}(u - u_{\Delta t})\|^2_{L^2(t_{n-1},t_n;V_0^*)}. \tag{4.37}$$

Proof. We have

$$\begin{aligned}|I| \le &\frac{24\sqrt{2\Delta t_n}}{\sigma_{min}} e^{-2\lambda t_{n-1}}\ \| \frac{\partial}{\partial t}(u - u_{\Delta t})\|_{L^2(t_{n-1},t_n;V_0^*)}\ \frac{\sigma_{min}}{\sqrt{2}} |u^n - u^{n-1}|_V \le \\ &\underbrace{\frac{8\sqrt{6}}{\sigma_{min}}\ e^{-\lambda t_{n-1}}\ \| \frac{\partial}{\partial t}(u - u_{\Delta t})\|_{L^2((t_{n-1},t_n);V_0^*)}\ \sqrt{3\Delta t_n}\ e^{-\lambda t_{n-1}}\ \frac{\sigma_{min}}{\sqrt{2}}\ |u^n - u_h^n|_V}_{=\,I_1} + \\ &\underbrace{\frac{8\sqrt{6}}{\sigma_{min}}\ e^{-\lambda t_{n-1}}\ \| \frac{\partial}{\partial t}(u - u_{\Delta t})\|_{L^2((t_{n-1},t_n);V_0^*)}\ \sqrt{3\Delta t_n}\ e^{-\lambda t_{n-1}}\ \frac{\sigma_{min}}{\sqrt{2}}\ |u^{n-1} - u_h^{n-1}|_V}_{=\,I_2} + \\ &\underbrace{\frac{24\sqrt{2}}{\sigma_{min}}\ e^{-2\lambda t_{n-1}}\ \| \frac{\partial}{\partial t}(u - u_{\Delta t})\|_{L^2((t_{n-1},t_n);V_0^*)}\ \sqrt{\Delta t_n}\frac{\sigma_{min}}{\sqrt{2}}\ |u_h^n - u_h^{n-1}|_V}_{=\,I_3}\ .\end{aligned} \tag{4.38}$$

Applying Young's inequality (4.36) with $\varepsilon = 1$ as well as (4.33), we get

$$\begin{aligned}I_1 &\le \frac{96}{\sigma_{min}^2}\ e^{-2\lambda t_{n-1}}\ \| \frac{\partial}{\partial t}(u - u_{\Delta t})\|^2_{L^2(t_{n-1},t_n;V_0^*)} + 3\Delta t_n\ e^{-2\lambda t_{n-1}}\ \frac{\sigma_{min}^2}{2}\ |u^n - u_h^n|_V^2 \\ &\le \frac{96}{\sigma_{min}^2}\ e^{-2\lambda t_{n-1}}\ \| \frac{\partial}{\partial t}(u - u_{\Delta t})\|^2_{L^2(t_{n-1},t_n;V_0^*)} + 6\ [[u^n - u_h^n]]^2.\end{aligned} \tag{4.39}$$

Likewise, but with (4.34) instead of (4.33) we obtain

$$\begin{aligned}I_2 &\le \frac{96}{\sigma_{min}^2}\ e^{-2\lambda t_{n-1}}\ \| \frac{\partial}{\partial t}(u - u_{\Delta t})\|^2_{L^2(t_{n-1},t_n;V_0^*)} + 3\Delta t_n\ e^{-2\lambda t_{n-1}}\ \frac{\sigma_{min}^2}{2}\ |u^{n-1} - u_h^{n-1}|_V^2 \\ &\le \frac{96}{\sigma_{min}^2}\ e^{-2\lambda t_{n-1}}\ \| \frac{\partial}{\partial t}(u - u_{\Delta t})\|^2_{L^2(t_{n-1},t_n;V_0^*)} + 6\ \rho_{\Delta t}\ [[u^{n-1} - u_h^{n-1}]]^2.\end{aligned} \tag{4.40}$$

Observing (4.4) and applying Young's inequality (4.36) with $\varepsilon = 1/10$, we get

$$I_3 \le \frac{1}{8}\ \eta_n^2 + \frac{11520}{\sigma_{min}^2}\ e^{-4\lambda t_{n-1}}\ \| \frac{\partial}{\partial t}(u - u_{\Delta t})\|^2_{L^2(t_{n-1},t_n;V_0^*)}. \tag{4.41}$$

Summing up the upper bounds in (4.39),(4.40) and (4.41) allows to conclude. □

Lemma 4.9. *Let the assumptions of Theorem 4.7 hold true. Then, for the term II from* (4.35) *we have*

$$|II| \leq \frac{1}{10}\ \eta_n^2 + 6[[u^n - u_h^n]]^2 + 6\rho_{\Delta t}[[u^{n-1} - u_h^{n-1}]]^2 + \frac{16\mu}{3\sigma_{min}^2} e^{-2\lambda t_{n-1}}\left(1 + 30e^{-2\lambda t_{n-1}}\right)\|u - u_{\Delta t}\|^2_{L^2(t_{n-1},t_n;V_0)}. \tag{4.42}$$

Proof. Straightforward estimation yields

$$\begin{aligned} |II| &\leq \frac{4\sqrt{2\mu\Delta t_n}}{\sigma_{min}} e^{-2\lambda t_{n-1}}\|u - u_{\Delta t}\|_{L^2(t_{n-1},t_n;V_0)} \frac{\sigma_{min}}{\sqrt{2}}|u^n - u^{n-1}|_V \qquad (4.43)\\ &\leq \underbrace{\frac{4\sqrt{\frac{2}{3}\mu}}{\sigma_{min}} e^{-\lambda t_{n-1}}\|u - u_{\Delta t}\|_{L^2(t_{n-1},t_n;V_0)}\sqrt{3\Delta t_n} e^{-\lambda t_{n-1}}\frac{\sigma_{min}}{\sqrt{2}}|u^n - u_h^n|_V}_{=:\, II_1} + \\ &\underbrace{\frac{4\sqrt{\frac{2}{3}\mu}}{\sigma_{min}} e^{-\lambda t_{n-1}}\|u - u_{\Delta t}\|_{L^2(t_{n-1},t_n;V_0)}\sqrt{3\Delta t_n} e^{-\lambda t_{n-1}}\frac{\sigma_{min}}{\sqrt{2}}|u^{n-1} - u_h^{n-1}|_V}_{=:\, II_2} + \\ &\underbrace{\frac{4\sqrt{2\mu}}{\sigma_{min}} e^{-2\lambda t_{n-1}}\|u - u_{\Delta t}\|_{L^2(t_{n-1},t_n;V_0)}\sqrt{\Delta t_n}\frac{\sigma_{min}}{\sqrt{2}}|u_h^n - u_h^{n-1}|_V}_{=:\, II_3}. \end{aligned}$$

The same estimates as for I_1, I_2 and I_3 result in

$$II_1 \leq \frac{8\mu}{3\sigma_{min}^2} e^{-2\lambda t_{n-1}}\|u - u_{\Delta t}\|^2_{L^2(t_{n-1},t_n;V_0)} + 6[[u^n - u_h^n]]^2, \tag{4.44}$$

$$II_2 \leq \frac{8\mu}{3\sigma_{min}^2} e^{-2\lambda t_{n-1}}\|u - u_{\Delta t}\|^2_{L^2(t_{n-1},t_n;V_0)} + 6\rho_{\Delta t}[[u^n - u_h^n]]^2, \tag{4.45}$$

$$II_3 \leq \frac{1}{10}\eta_n^2 + \frac{160\mu}{\sigma_{min}^2} e^{-4\lambda t_{n-1}}\|u - u_{\Delta t}\|^2_{L^2(t_{n-1},t_n;V_0)}. \tag{4.46}$$

The upper bound (4.42)follows by summing up (4.44),(4.45) and (4.46). ∎

Lemma 4.10. *Suppose that the assumptions of Theorem 4.7 hold true. Then, for $n > 1$*

the term III from (4.35) *can be bounded from above according to*

$$|III| \leq \frac{1}{10}\eta_n^2 + 6[[u^n - u_h^n]]^2 + 6\rho_{\Delta t}[[u^{n-1} - u_h^{n-1}]]^2 + \frac{48C^2}{\sigma_{min}^4}\ max(1, \rho_{\Delta t})\ \Big(1 + 60e^{-2\lambda t_{n-1}}\Big)(\Delta t_n)^2\|u^0\|^2. \tag{4.47}$$

In case $n = 1$, we have

$$|III| \leq \frac{1}{10}\eta_1^2 + 6[[u^1 - u_h^1]]^2 + \frac{23232C^2}{\sigma_{min}^6}\ (\Delta t_1)^2\ \Big(\|u^0\|^2 + \frac{\sigma_{min}^2}{2}\ \Delta t_1 |u^0|_V^2\Big). \tag{4.48}$$

Proof. **Case n > 1:** As in (4.20) it follows that

$$\begin{aligned}
|III| &\leq \frac{6\sqrt{2}C}{\sigma_{min}}\Delta t_n e^{-2\lambda t_{n-1}}\|u_{\Delta t}\|_{L^2(t_{n-1},t_n;V_0)}|u^n - u^{n-1}|_V \\
&\leq \underbrace{\frac{4C\sqrt{3\Delta t_n}}{\sigma_{min}^2}e^{-\lambda t_{n-1}}\|u_{\Delta t}\|_{L^2(t_{n-1},t_n;V_0)}\sqrt{3\Delta t_n}e^{-\lambda t_{n-1}}\frac{\sigma_{min}}{\sqrt{2}}|u^n - u_h^n|_V}_{=:\ III_1} + \\
&\quad \underbrace{\frac{4C\sqrt{3\Delta t_n}}{\sigma_{min}^2}e^{-\lambda t_{n-1}}\|u_{\Delta t}\|_{L^2(t_{n-1},t_n;V_0)}\sqrt{3\Delta t_n}e^{-\lambda t_{n-1}}\frac{\sigma_{min}}{\sqrt{2}}|u^{n-1} - u_h^{n-1}|_V}_{=:\ III_2} + \\
&\quad \underbrace{\frac{12C\sqrt{\Delta t_n}}{\sigma_{min}^2}e^{-2\lambda t_{n-1}}\|u_{\Delta t}\|_{L^2(t_{n-1},t_n;V_0)}\sqrt{\Delta t_n}\frac{\sigma_{min}}{\sqrt{2}}|u_h^n - u_h^{n-1}|_V}_{=:\ III_3}.
\end{aligned} \tag{4.49}$$

Estimating III_1, III_2 and III_3 as before, we obtain

$$III_1 \leq \frac{12C^2\Delta t_n}{\sigma_{min}^4}e^{-2\lambda t_{n-1}}\|u_{\Delta t}\|_{L^2(t_{n-1},t_n;V_0)}^2 + 6[[u^n - u_h^n]]^2, \tag{4.50}$$

$$III_2 \leq \frac{12C^2\Delta t_n}{\sigma_{min}^4}e^{-2\lambda t_{n-1}}\|u_{\Delta t}\|_{L^2(t_{n-1},t_n;V_0)}^2 + 6\rho_{\Delta t_n}[[u^{n-1} - u_h^{n-1}]]^2, \tag{4.51}$$

$$III_3 \leq \frac{1440C^2\Delta t_n}{\sigma_{min}^4}e^{-4\lambda t_{n-1}}\|u_{\Delta t}\|_{L^2(t_{n-1},t_n;V_0)}^2 + \frac{1}{10}\eta_n^2. \tag{4.52}$$

For the estimation of the first terms on the right-hand sides of (4.50),(4.51) and (4.52) we use (3.15) and obtain

$$
\begin{aligned}
& e^{-2\lambda t_{n-1}} \Delta t_n \|u_{\Delta t}\|^2_{L^2(t_{n-1},t_n;V_0)} && (4.53)\\
& \leq 2 \prod_{i=1}^{n-1} (1-2\lambda\Delta t_i)\ \Delta t_n \|u_{\Delta t}\|^2_{L^2(t_{n-1},t_n;V_0)} && (4.54)\\
& \leq \max(1,\rho_{\Delta t})\ \Delta t_n \Big(\Delta t_n \prod_{i=1}^{n-1} (1-2\lambda\Delta t_i)\ \Delta t_n |u^n|^2_V) \\
& + \Delta t_{n-1} \Big(\Delta t_{n-1} \prod_{i=1}^{n-2} (1-2\lambda\Delta t_i)\ \Delta t_{n-1} |u^{n-1}|^2_V \Big) \\
& \leq 2\max(1,\rho_{\Delta t})\ (\Delta t_n)^2 [[(u^m)]]^2_n \\
& \leq 2\max(1,\rho_{\Delta t})\ (\Delta t_n)^2 \|u^0\|^2.
\end{aligned}
$$

Using (4.53) in (4.50),(4.51),(4.52) and summing up, we deduce (4.47).

Case n = 1: As in the case $n > 1$, we obtain

$$
\begin{aligned}
|III| & \leq \frac{6\sqrt{2}C}{\sigma^2_{min}} \Delta t_1 \|u_{\Delta t}\|_{L^2(0,t_1;V_0)} |u^1 - u^0|_V && (4.55)\\
& \leq \underbrace{\frac{4C\sqrt{3\Delta t_1}}{\sigma^2_{min}} \|u_{\Delta t}\|_{L^2(0,t_1;V_0)} \sqrt{3\Delta t_n} \frac{\sigma_{min}}{\sqrt{2}} |u^1 - u^1_h|_V}_{=:\ III_1} + \\
& \quad \underbrace{\frac{12C\sqrt{\Delta t_1}}{\sigma^2_{min}} \|u_{\Delta t}\|_{L^2(0,t_1;V_0)} \sqrt{\Delta t_1} \frac{\sigma_{min}}{\sqrt{2}} |u^1_h - u^0_h|_V,}_{=:\ III_2}
\end{aligned}
$$

from which we deduce the upper bounds

$$
\begin{aligned}
III_1 & \leq \frac{12C^2\Delta t_1}{\sigma^4_{min}} \|u_{\Delta t}\|^2_{L^2(0,t_1;V_0)} + 6[[u^1 - u^1_h]]^2, && (4.56)\\
III_2 & \leq \frac{1440C^2\Delta t_1}{\sigma^4_{min}} \|u_{\Delta t}\|^2_{L^2(0,t_1;V_0)} + \frac{1}{10}\eta^2_1. && (4.57)
\end{aligned}
$$

On the other hand, we have

$$\begin{aligned}\|u_{\Delta t}\|^2_{L^2(0,t_1;V_0)} &\le 2(\Delta t_1)^2\Big(|u^1|_V^2 + |u^0|_V^2\Big) \\ &\le \frac{16}{\sigma^2_{min}}\ \Delta t_1\ \Big([[(u^m)]]_1^2 + \frac{\sigma^2_{min}}{2}\ \Delta t_1 |u^0|_V^2\Big) \\ &\le \frac{16}{\sigma^2_{min}}\ \Delta t_1\ \Big(\|u^0\|^2 + \frac{\sigma^2_{min}}{2}\ \Delta t_1 |u^0|_V^2\Big).\end{aligned} \tag{4.58}$$

Inserting (4.58) into (4.56),(4.57) and observing (4.55) gives (4.48). □

Lemma 4.11. *Under the assumptions of Theorem 4.7, the term IV from* (4.35) *can be bounded from above by means of*

$$\begin{aligned}|IV| \le\ & \frac{1}{5}\eta_n^2 + 12[[u^n - u_h^n]]^2 + 12\rho_{\Delta t}[[u^n - u_h^n]]^2\ + \\ & \frac{64\lambda^2}{3\sigma^2_{min}}\ e^{-2\lambda t_{n-1}}\ \Big(1 + \frac{5}{4}\ e^{-2\lambda t_{n-1}}\Big)\ (\Delta t_n)^2\ \|\frac{\partial}{\Delta t}(u - u_{\Delta t})\|^2_{L^2(t_{n-1},t_n;V_0^*)}\ + \\ & \frac{64\lambda^2\mu^2}{3\sigma^4_{min}}\ e^{2\lambda T}\Big(1 + \frac{15}{4}e^{-2\lambda T}\Big)\ (\Delta t_n)^2\ \|u^0\|^2.\end{aligned} \tag{4.59}$$

Proof. We have

$$\begin{aligned}|IV| \le\ & 2\Delta t_n\ \lambda e^{-2\lambda t_{n-1}} \|u^n - u^{n-1}\|^2 \\ \le\ & \underbrace{2\lambda e^{-2\lambda t_{n-1}} |\int_{t_{n-1}}^{t_n} (u^n - u^{n-1} - \Delta t_n \frac{\partial u}{\partial t})(u^n - u^{n-1}) d\tau|}_{=:\ IV_1}\ + \\ & \underbrace{2\lambda e^{-2\lambda t_{n-1}}\ \Delta t_n |\int_{t_{n-1}}^{t_n} \frac{\partial u}{\partial t}(u^n - u^{n-1}) d\tau|}_{=:\ IV_2}.\end{aligned} \tag{4.60}$$

Straightforward estimation of IV_1 gives

$$
\begin{aligned}
IV_1 =\ & 2\lambda e^{-2\lambda t_{n-1}} \Delta t_n \ | \int_{t_{n-1}}^{t_n} \frac{\partial}{\Delta t}(u_{\Delta t} - u)(u^n - u^{n-1}) d\tau | \\
\leq\ & 2\lambda e^{-2\lambda t_{n-1}} (\Delta t_n)^{3/2} \ \| \frac{\partial}{\Delta t}(u - u_{\Delta t}) \|_{L^2(t_{n-1},t_n;V_0^*)} |u^n - u^{n-1}|_V \\
\leq\ & \underbrace{\frac{2\sqrt{2}\lambda}{\sqrt{3}\sigma_{min}} \ e^{-\lambda t_{n-1}} \ \Delta t_n \ \| \frac{\partial}{\Delta t}(u - u_{\Delta t}) \|_{L^2(t_{n-1},t_n;V_0^*)} \ \sqrt{3\Delta t_n} e^{-\lambda t_{n-1}} \frac{\sigma_{min}}{\sqrt{2}} \ |u^n - u_h^n|_V}_{=: \ IV_{1,1}} \ + \\
& \underbrace{\frac{2\sqrt{2}\lambda}{\sqrt{3}\sigma_{min}} \ e^{-\lambda t_{n-1}} \ \Delta t_n \ \| \frac{\partial}{\Delta t}(u - u_{\Delta t}) \|_{L^2(t_{n-1},t_n;V_0^*)} \ \sqrt{3\Delta t_n} e^{-\lambda t_{n-1}} \frac{\sigma_{min}}{\sqrt{2}} \ |u^{n-1} - u_h^{n-1}|_V}_{=: \ IV_{1,2}} \\
& + \underbrace{\frac{2\sqrt{2}\lambda}{\sigma_{min}} \ e^{-2\lambda t_{n-1}} \ \Delta t_n \ \| \frac{\partial}{\Delta t}(u - u_{\Delta t}) \|_{L^2(t_{n-1},t_n;V_0^*)} \ \sqrt{\Delta t_n} \frac{\sigma_{min}}{\sqrt{2}} \ |u_h^n - u_h^{n-1}|_V}_{=: \ IV_{1,3}} .
\end{aligned}
\tag{4.61}
$$

As before, by applying (4.36), we obtain the following upper bounds for $IV_{1,1}, IV_{1,2}$ and $IV_{1,3}$

$$
IV_{1,1} \leq \frac{32\lambda^2}{3\sigma_{min}^2} \ e^{-2\lambda t_{n-1}} \ (\Delta t_n)^2 \ \| \frac{\partial}{\Delta t}(u - u_{\Delta t}) \|^2_{L^2(t_{n-1},t_n;V_0^*)} + 6[[u^n - u_h^n]]^2, \tag{4.62}
$$

$$
IV_{1,2} \leq \frac{32\lambda^2}{3\sigma_{min}^2} \ e^{-2\lambda t_{n-1}} \ (\Delta t_n)^2 \ \| \frac{\partial}{\Delta t}(u - u_{\Delta t}) \|^2_{L^2(t_{n-1},t_n;V_0^*)} + 6\rho_{\Delta t}[[u^{n-1} - u_h^{n-1}]]^2, \tag{4.63}
$$

$$
IV_{1,3} \leq \frac{1}{10}\eta_n^2 + \frac{80\lambda^2}{3\sigma_{min}^2} \ e^{-4\lambda t_{n-1}} \ (\Delta t_n)^2 \ \| \frac{\partial}{\Delta t}(u - u_{\Delta t}) \|^2_{L^2(t_{n-1},t_n;V_0^*)}. \tag{4.64}
$$

Using (4.62),(4.63),(4.64) in (4.61) results in

$$
\begin{aligned}
IV_1 \leq\ & \frac{1}{10}\eta_n^2 + 6[[u^n - u_h^n]]^2 + 6\rho_{\Delta t}[[u^{n-1} - u_h^{n-1}]]^2 \ + \\
& \frac{64\lambda^2}{3\sigma_{min}^2} \ e^{-2\lambda t_{n-1}} \ \left(1 + \frac{5}{4} \ e^{-2\lambda t_{n-1}}\right) \ (\Delta t_n)^2 \ \| \frac{\partial}{\Delta t}(u - u_{\Delta t}) \|^2_{L^2(t_{n-1},t_n;V_0^*)}.
\end{aligned}
\tag{4.65}
$$

As far as IV_2 is concerned, in view of (2.25) we find

$$
\begin{aligned}
IV_2 &\le 2\lambda\ (\Delta t_n)^{3/2}\ e^{\lambda \Delta t_n}\ \|e^{-\lambda t}\frac{\partial u}{\partial t}\|_{L^2(t_{n-1},t_n;V_0^*)}\ |u^n - u^{n-1}|_V \qquad (4.66)\\
&\le \frac{2\sqrt{2}\lambda\mu}{\sigma_{min}}\ (\Delta t_n)^{3/2}\ e^{\alpha}\|u^0\|\ |u^n - u^{n-1}|_V\\
&\le \underbrace{\frac{2\sqrt{2}\lambda\mu}{\sqrt{3}\sigma_{min}^2}\ e^{\lambda t_{n-1}}\ \Delta t_n\|u^0\|\ \sqrt{3\Delta t_n}\ e^{-\lambda t_{n-1}}\ \frac{\sigma_{min}}{\sqrt{2}}\ |u^n - u_h^n|_V}_{=:\ IV_{2,1}}\ +\\
&\quad \underbrace{\frac{2\sqrt{2}\lambda\mu}{\sqrt{3}\sigma_{min}^2}\ e^{\lambda t_{n-1}}\ \Delta t_n\|u^0\|\ \sqrt{3\Delta t_n}\ e^{-\lambda t_{n-1}}\ \frac{\sigma_{min}}{\sqrt{2}}\ |u^{n-1} - u_h^{n-1}|_V}_{=:\ IV_{2,2}}\ +\\
&\quad \underbrace{\frac{2\sqrt{2}\lambda\mu}{\sigma_{min}^2}\ \Delta t_n\|u^0\|\ \sqrt{\Delta t_n}\ \frac{\sigma_{min}}{\sqrt{2}}\ |u_h^n - u_h^{n-1}|_V}_{=:\ IV_{2,3}}\,.
\end{aligned}
$$

Applying Young's inequality (4.36) gives

$$IV_{2,1} \le \frac{32\lambda^2\mu^2}{3\sigma_{min}^4}\ e^{2\lambda T}\ (\Delta t_n)^2\ \|u^0\|^2 + 6[[u^n - u_h^n]]^2, \qquad (4.67)$$

$$IV_{2,2} \le \frac{32\lambda^2\mu^2}{3\sigma_{min}^4}\ e^{2\lambda T}\ (\Delta t_n)^2\ \|u^0\|^2 + 6\rho_{\Delta t}[[u^n - u_h^n]]^2, \qquad (4.68)$$

$$IV_{2,3} \le \frac{1}{10}\eta_n^2 + \frac{80\lambda^2\mu^2}{\sigma_{min}^4}\ (\Delta t_n)^2\ \|u^0\|^2. \qquad (4.69)$$

Using (4.67),(4.68) and (4.69) in (4.66), we obtain

$$
\begin{aligned}
IV_2 \le\ & \frac{1}{10}\eta_n^2 + 6[[u^n - u_h^n]]^2 + 6\rho_{\Delta t}[[u^n - u_h^n]]^2\ + \qquad (4.70)\\
& \frac{64\lambda^2\mu^2}{3\sigma_{min}^4}\ e^{2\lambda T}\Big(1 + \frac{15}{4}e^{-2\lambda T}\Big)\ (\Delta t_n)^2\ \|u^0\|^2.
\end{aligned}
$$

The estimate (4.59) now follows from (4.65) and (4.70). □

4.3.2 Local efficiency of the estimator $\eta_{n,T}$.

Local efficiency means that we can estimate the local contributions of the error estimator from above by some local error norms. To this end, for $T \in \mathcal{T}_{nh}(\Omega)$ we denote by

$$\omega_T := \bigcup \{T' \in \mathcal{T}_{nh}(\Omega) \mid \mathcal{N}_{nh}(T') \cap \mathcal{N}_{nh}(T) \neq \emptyset\} \tag{4.71}$$

the patch around T consisting of all neighboring elements that share at least one vertex with T. We introduce the local space

$$V(\omega_T) := \{v : \omega_T \to \mathbb{R} \mid v \in L^2(\omega_T)\ ,\ S_i \frac{\partial v}{\partial S_i} \in L^2(\omega_T), 1 \leq i \leq 2\}, \tag{4.72}$$

equipped with the norm $\|\cdot\|_{V(\omega_T)}$ and semi-norm $|\cdot|_{V(\omega_T)}$. We further refer to $V_0(\omega_T)$ as the closure of $C_0^\infty(\omega_T)$ in $V(\omega_T)$ and to $V_0^*(\omega_T)$ as the associated dual space.

Under these prerequisites, we can prove the following local efficiency of $\eta_{n,T}$.

Theorem 4.12. *Let $\eta_{n,T}, T \in \mathcal{T}_{nh}(\Omega), 1 \leq n \leq N$, be the given by (4.5). Then, under the assumptions of Theorem 4.7 there holds*

$$\eta_{n,T}^2 \lesssim \Big(\|(\Delta t_n)^{-1}\Big(u^n - u_h^n - (u^{n-1} - u_h^{n-1})\Big)\|_{V_0^*(\omega_T)} + \mu\ |u^n - u_h^n|_{V_0(\omega_T)} \Big). \tag{4.73}$$

Proof. The local efficiency can be shown by standard arguments from the a posteriori error analysis of residual-type a posteriori error estimators [40]. In particular, for $T \in \mathcal{T}_{nh}(\Omega)$ we refer to $\lambda_i^T, 1 \leq i \leq 3$, as the barycentric coordinates and introduce

$$\psi_T := \prod_{i=1}^{3} \lambda_i^T$$

as the associated edge bubble function having its support in T. Then, we use (4.26) in the proof of Proposition 4.6 with $v = 0$ and $v_h = z_h^n \psi_T^{1/2}$ where $z_h^n \in P_1(T)$ is given by

$$z_h^n := (\Delta t_n)^{-1}(u_h^n - u_h^{n-1}) - S \frac{\partial u_h^n}{\partial S} + r u_h^n,$$

satisfying the inverse inequalities

$$\|z_h^n\|_{0,T} \lesssim \|z_h^n \psi_T^{1/2}\|_{0,T} \quad , \quad \|S\frac{\partial z_h^n}{\partial S}\|_{0,T} \lesssim \frac{S_{max}(T)}{h_T} \ \|z_h^n\|_{0,T}. \tag{4.74}$$

With this choice, (4.26) results in

$$(\Delta t_n)^{-1}((u^{n-1}-u_h^{n-1})-(u^n-u_h^n),v)_{0,T}+a_{t_n}(u_h^n-u^n,v)=\|v\|_{0,T}^2=\|z_h^n\psi_T^{1/2}\|_{0,T}^2.$$

The inverse inequalities (4.74) imply

$$\|z_h^n\|_{0,T}^2 \lesssim \frac{S_{max}(T)}{h_T} \ \Big(\|(\Delta t_n)^{-1}\Big(u^n-u_h^n-(u^{n-1}-u_h^{n-1})\Big)\|_{V_0^*(\omega_T)}+\mu \ |u^n-u_h^n|_{V_0(\omega_T)}\Big)\|z_h^n\|_{0,T},$$

which gives the assertion. □

Chapter 5

The Adaptive Cycle

As we have pointed out in chapter 2, adaptive finite element methods consist of successive loops of a cycle involving the basic steps $SOLVE$, $ESTIMATE$, $MARK$, and $REFINE$. While the concepts for an a posteriori error estimation based on residual-type estimators have been dealt with in detail in the previous chapter, here we focus on the numerical solution of the fully discretized Black-Scholes equation. As we shall see in section 5.1, the structure of the resulting linear algebraic systems to be solved is such that we are faced with tridiagonal $N \times N$ coefficient matrices. Hence, the solution by LU decomposition is the method of choice, since it is of optimal arithmetic complexity $O(N)$.

In Section 5.2, we will describe two different strategies for the realization of the adaptive process. The first strategy has been suggested in [1]. For each adaptive step, either a refinement in time or a refinement in space is done based on the information provided by the time and space error estimators. This strategy is the one which has been implemented in this thesis. Alternatively, we will describe a second strategy due to [20]. In contrast to the first strategy, a progressive time-stepping is realized with an adaptive choice of the next time step combined with a simultaneous refinement and coarsening of the spatial mesh. This strategy requires an additional estimator for coarsening which will be described in subsection 5.2.2.

5.1 The steps SOLVE and ESTIMATE

We obtain the fully discretized problem by applying the Crank-Nicolson scheme in time:

Find $(P_h^n)_{0\leq n\leq N}$, $P_h^n \in V_{nh}^0$ satisfying

$$P_h^0 = P_0, \tag{5.1}$$

and for all n, $1 \leq n \leq N$,

$$\forall v_h \in V_{nh}^0, \quad (P_h^n - P_h^{n-1}, v_h) + \frac{\Delta t_n}{2}(a_n(P_h^n, v_h) + a_{n-1}(P_h^{n-1}, v_h)) = 0. \tag{5.2}$$

where $a_n = a_{t_n}$, and $a_t(v,w), v, w \in V_{nh}^0$ is given by

$$a_t(v,w) = -\sum_{i=1}^{N} \frac{S_i^2\sigma^2(S_i,t)}{2}[\frac{\partial v}{\partial S}](S_i)w(S_i) - r(t)\int_0^{\bar{S}} S\frac{\partial v}{\partial S}w + r(t)\int_0^{\bar{S}} vw. \tag{5.3}$$

Here, $[\frac{\partial v}{\partial S}](S_i)$ denotes the jump of $\frac{\partial v}{\partial S}$ at S_i.

Let $(w^i)_{i=0,\cdots N}$ be the nodal basis of V_h, i.e.,

$$w^i = \begin{cases} 0, & S \leq S_{i-1} \\ \frac{S-S_{i-1}}{S_i-Si-1}, & S_{i-1} \leq S \leq S_i \\ \frac{S_{i+1}-S}{S_{i+1}-Si}, & S_i \leq S \leq S_{i+1} \\ 0, & S \geq S_{i+1} \end{cases} \tag{5.4}$$

Moreover, let $\mathbf{M}$ and $\mathbf{A}^n$ in $\mathbb{R}^{(N+1)\times(N+1)}$ be the mass and stiffness matrices defined by

$$\mathbf{M}_{i,j} = (w^i, w^j) \quad , \quad \mathbf{A}^n_{i,j} = a_{t_n}(w^j, w^i) \ , \ 0 \le i, j \le N.$$

Setting $P^n = (u_h^n(S_0), \cdots, u_h^n(S_N))^T$ and $P^0 = (u_0(S_0), \cdots, u_0(S_N))^T$, (5.2) is equivalent to

$$\mathbf{M}(P^n - P^{n-1}) + \Delta t_n \mathbf{A}^n P^n = 0. \tag{5.5}$$

The shape function w^i defined in (5.4) has its support in $[S_{i-1}, S_{i+1}]$. This implies that the matrices $\mathbf{M}$ and $\mathbf{A}^n$ are tridiagonal. Furthermore, we have

$$w^i(S) = \frac{S - S_{i-1}}{h_i}, \quad \frac{\partial w^i}{\partial S} = \frac{1}{h_i}, \qquad \forall S \in (S_{i-1}, S_i), \tag{5.6}$$

$$w^i(S) = \frac{S_{i+1} - S}{h_{i+1}}, \quad \frac{\partial w^i}{\partial S} = -\frac{1}{h_{i+1}}, \qquad \forall S \in (S_i, S_{i+1}), \tag{5.7}$$

which results in

$$\begin{aligned}
\int_0^{\bar{S}} w^{i-1} w^i = \frac{h_i}{6}, &\qquad \int_0^{\bar{S}} S w^i \frac{\partial w^{i-1}}{\partial S} = -\frac{S_{i-1}}{6} - \frac{S_i}{3}, \\
\int_0^{\bar{S}} (w^i)^2 = \frac{h_i + h_{i+1}}{3}, &\qquad \int_0^{\bar{S}} S w^i \frac{\partial w^i}{\partial S} = -\frac{h_i + h_{i+1}}{6} \quad \text{if} \quad i > 0, \\
\int_0^{\bar{S}} (w^0)^2 = \frac{h_1}{3}, &\qquad \int_0^{\bar{S}} S w^0 \frac{\partial w^0}{\partial S} = -\frac{h_1}{6}, \\
\int_0^{\bar{S}} w^{i+1} w^i = \frac{h_{i+1}}{6}, &\qquad \int_0^{\bar{S}} S w^i \frac{\partial w^{i+1}}{\partial S} = \frac{S_{i+1}}{6} + \frac{S_i}{3}.
\end{aligned} \tag{5.8}$$

From this, straightforward calculations show that the entries of $\mathbf{A}^n$ are given by

$$
\begin{aligned}
\mathbf{A}^n_{i,i-1} &= -\frac{S_i^2\sigma^2(t_n,S_i)}{2h_i} + \frac{r(t_n)S_i}{2}. \quad 1 \le i \le N, \\
\mathbf{A}^n_{i,i} &= \frac{S_i^2\sigma^2(t_n,S_i)}{2}\left(\frac{1}{h_i} + \frac{1}{h_{i+1}}\right) + \frac{r(t_n)}{2}(h_{i+1}+h_i), \quad 1 \le i \le N, \\
\mathbf{A}^n_{0,0} &= \frac{r(t_n)}{2}h_1, \\
\mathbf{A}^n_{i,i+1} &= -\frac{S_i^2\sigma^2(t_n,S_i)}{2h_{i+1}} - \frac{r(t_n)S_i}{2}. \quad 0 \le i \le N-1,
\end{aligned} \tag{5.9}
$$

whereas the entries of $\mathbf{M}^m$ are as follows

$$
\begin{aligned}
\mathbf{M}_{i,i-1} &= \frac{h_i}{6}. \quad 1 \le i \le N, \\
\mathbf{M}_{i,i} &= \frac{h_{i+1}+h_i}{3}. \quad 1 \le i \le N, \\
\mathbf{M}_{0,0} &= \frac{h_1}{3}. \\
\mathbf{M}_{i,i+1} &= \frac{h_{i+1}}{6}. \quad 0 \le i \le N-1,
\end{aligned} \tag{5.10}
$$

Hence, the fastest method to solve (5.5) is by LU decomposition.

As far as the ingredients of the step ESTIMATE are concerned, we refer to section 4.1.

5.2 The Steps MARK and REFINE

In this section, we briefly describe the marking and refinement strategy from [1] as well as the strategy suggested in [20] for the adaptive finite element solution of initial-boundary value problems for parabolic problems.

5.2.1 Refinement in space and time

We use the following refinement strategy:

Step 1: We need to decide whether we want to refine in the S or in the t variable. We will make this decision by a comparison of the quantities η_n^2 and $\sum_{T\in T_{nh}(\Omega)} \eta_{n,T}^2$:

If

$$\eta_n^2 > \sqrt{2}(\frac{\Delta t_n}{\sigma_{min}^2} \sum_{T\in T_{nh}(\Omega)} \eta_{n,T}^2),$$

which means that the time error indicator dominates the global S-discretization error indicator, we will refine the mesh in the t variable.

On the other hand, if

$$\eta_n^2 < \frac{1}{\sqrt{2}}(\frac{\Delta t_n}{\sigma_{min}^2} \sum_{T\in T_{nh}(\Omega)} \eta_{n,T}^2),$$

which means that the global S-discretization error estimator dominates the time error estimator, we will refine the mesh in the S variable. Otherwise, we refine the mesh in t, if the current level is even, and refine the mesh in S, if the level is odd.

Step 2: We define the refinement with respect to t as follows: First, we compute

$$\bar{\zeta} := max_n \eta_n \quad \text{and} \quad \underline{\zeta} := min_n \eta_n.$$

If

$$\zeta_n > \frac{\bar{\zeta} + \underline{\zeta}}{2},$$

we divide the time interval $[t_{n-1}, t_n]$ by 2. Otherwise, we do nothing.

The refinement in the S variable is done analogously.

5.2.2 Progressive time stepping and refinement/coarsening

When doing progressive adaptive time stepping combined with adaptive refinement/coarsening in space for parabolic partial differential equations, there is a need to introduce an additional coarsening strategy in the a posteriori error analysis after mesh and time step refinements have been done based on those error estimators as discussed in section 4.1.

We denote by $\mathcal{T}_{nh}(\Omega)$ the actual simplicial triangulation of Ω and by $\mathcal{T}_{nH}(\Omega)$ a coarsened mesh. The spaces V^0_{nh} and V^0_{nH} stand for the associated finite element spaces. We refer to

$$I^n_H : C(\overline{\Omega}) \to V^0_{nH}$$

as the standard finite element interpolant.

Then, the coarsening error estimator is given by

$$\eta^{coarse}_n := (\Delta t_n)^{-1} \|u^n_h - I^n_H u^n_h\|^2_{0,\Omega} + |u^n_h - I^n_H u^n_h|^2_V. \qquad (5.11)$$

We further introduce three tolerances $TOL_{time} > 0$, $TOL_{space} > 0$ and $TOL_{coarse} > 0$, which are upper bounds for the time error estimator η_n, the space error estimator η^n_{space}, and the coarsening error estimator η^{coarse}_n, respectively. The following algorithm describes the adaptation of the adaptive strategy from [20] to progressive time stepping and adaptive refinement/coarsening for the fully discretized Black-Scholes equation:

Step 1: Once the approximate solution at a fixed time τ_n is known, compute the time error estimator. If it is larger than the tolerance $TOL_{time}/(2T)$, go back to the previous time and perform a computation with a reduced time step (e.g., half the size of the previously used one) until the time error estimator becomes smaller than the tolerance $TOL_{time}/(2T)$.

Step 2: Perform mesh adaptivity in the standard way by computing the space error estimators and refining the mesh where they are larger than the tolerance TOL_{space}/T. Compute the approximate solution with respect to the refined mesh and check the time error estimator. Perform a time step reduction, if necessary.

Step 3: Coarsen the actual mesh $\mathcal{T}_{nh}(\Omega)$ according to $\eta^{coarse}_n < TOL_{coarse}/T$, compute the approximate solution with respect to the coarsened mesh $\mathcal{T}_{nH}(\Omega)$ and compute the associated time error estimator.

Step 4: If the time error estimator is smaller than the tolerance $\Theta_{time}TOL_{time}/(2T)$ for some $\Theta_{time} \in (0,1)$, enlarge the time step (e.g., double it) and proceed to the next time step. Otherwise, go back to Step 1.

Chapter 6

Numerical Results

We present the results of numerical computations based on the adaptive refinement strategy described in subsection 5.2.1 for two examples (constant and variable volatility) where the data have been taken from [1].

Example 1: Constant Volatility. We choose the following data:

- $K = 100$ (strike)
- $T = 1$ (maturity 1 year)
- $\sigma = 0.2$ (volatility)
- $r = 0.04$ (interest rate)

and compute the price of a vanilla European put in the rectangle $[0, 200] \times [0, 1]$. The initial mesh is a uniform mesh with 20 nodes in t and 80 nodes in S.

Progressive Mesh Refinement

Using the strategy we described at the beginning of this chapter we get the following plots:

Figure 6.1 shows a visualization of the discrete solution after 5 refinement steps.

Figure 6.2 displays the error between the prices computed by the Black-Scholes formula and by the finite element method realized in this thesis. The four plots obtained after 0, 5, 10, and 19 refinements, respectively, clearly show a decrease in the error with progressive refinement. We also note that the errors are large around $S = 100$. Hence, we expect strong mesh refinement in this region.

Figure 6.3 shows the refined meshes after $0, 5, 10$ and 19 refinement steps of the adaptive algorithm. We observe a pronounced refinement around $S = 100$ as the process goes on.

Figure 6.4 displays the time error indicator η_n and the global S error indicator

$$(\frac{\Delta t_n}{\sigma_{min}^2} \sum_{T \in T_{nh}(\Omega)} \eta_{nT}^2)^{\frac{1}{2}}$$

versus time after $0, 5, 10$ and 19 refinement steps, respectively.

Figure 6.5 shows us the space error indicator η_{mT}. Obviously, this local indicator is also large around $S = 100$.

Figure 6.6 observes the changes in the error

$$\sigma \|u - u_{h,\Delta t}\|_{L^2((0,T);V)}$$

and in the estimated error

$$(\sum_m (\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2} \sum_{T \in T_{nh}(\Omega)} \eta_{n,T}^2))^{\frac{1}{2}}$$

as the refinement level (left) and the total number of nodes (right) change.

Table 6.1 contains the convergence history of the adaptive refinement process.

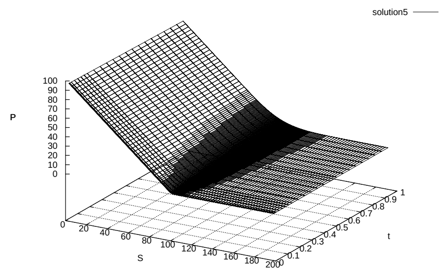


Figure 6.1: The discrete solution obtained after 5 steps.

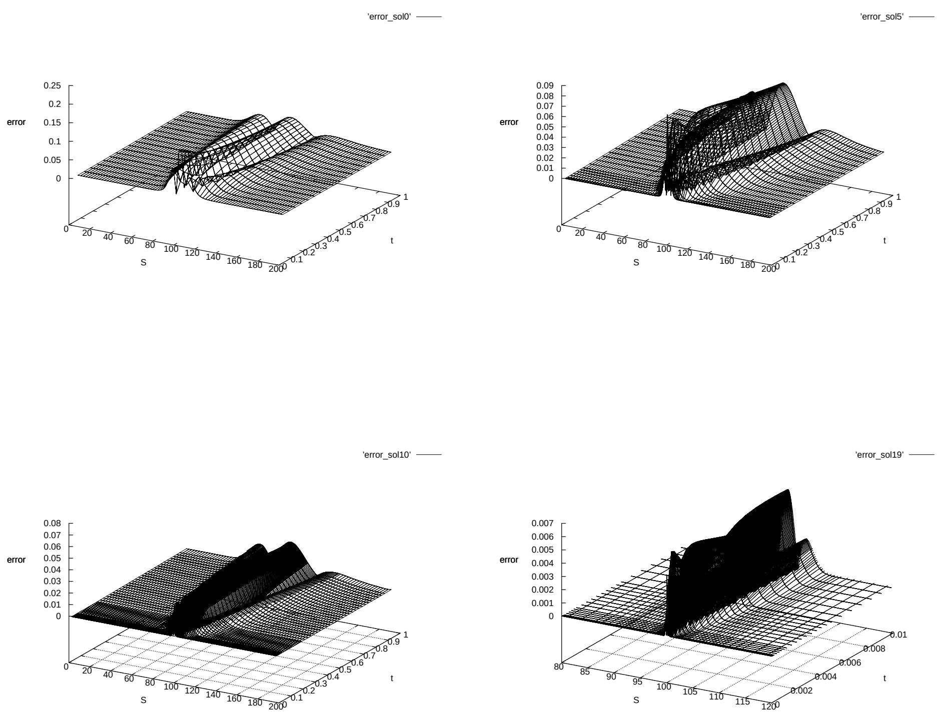


Figure 6.2: Errors obtained after $0, 5, 10$ and 19 refinement steps. The bottom right figure is a zoom.

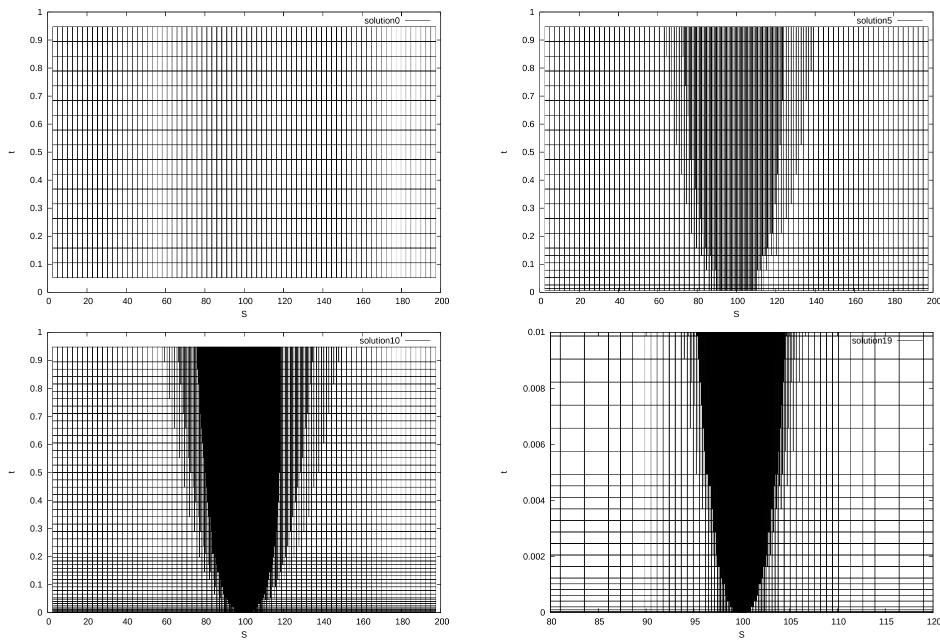


Figure 6.3: Refined mesh obtained after $0, 5, 10$ and 19 refinement steps. The bottom right figure is a zoom.

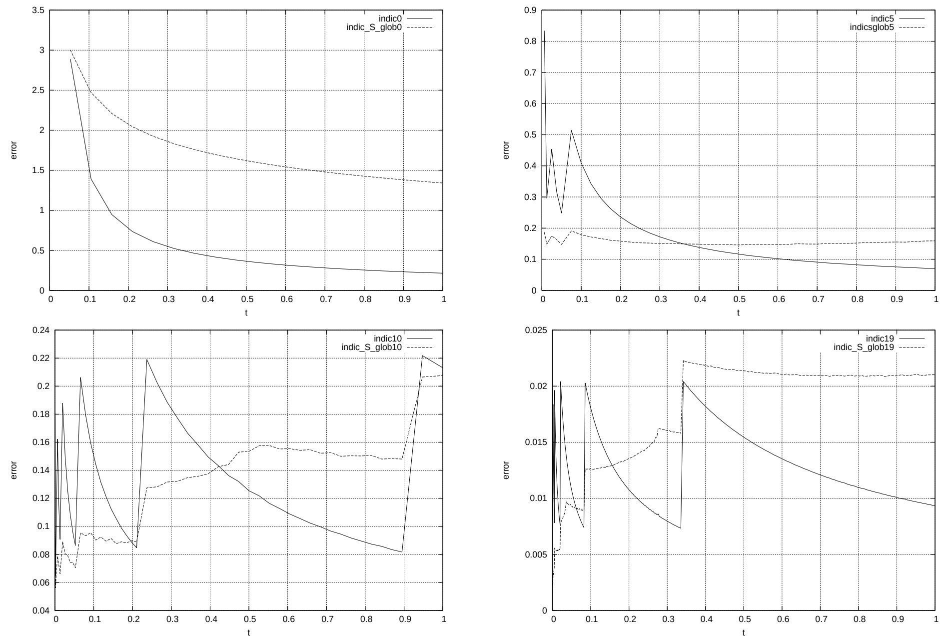
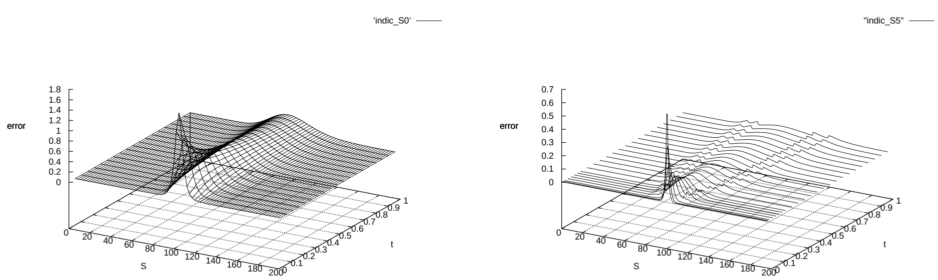


Figure 6.4: Time error indicator η_n and the global S error indicator $\left(\frac{\Delta t_n}{\sigma_{min}^2} \sum_{T \in T_{nh}(\Omega)} \eta_{n,T}^2\right)^{\frac{1}{2}}$ obtained after 0, 5, 10 and 19 refinement steps



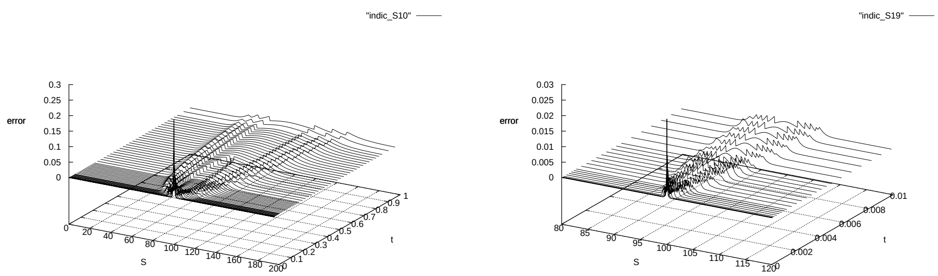


Figure 6.5: Space error $\eta_{n,T}$ obtained after $0, 5, 10$ and 19 refinement steps. The bottom right figure is a zoom.

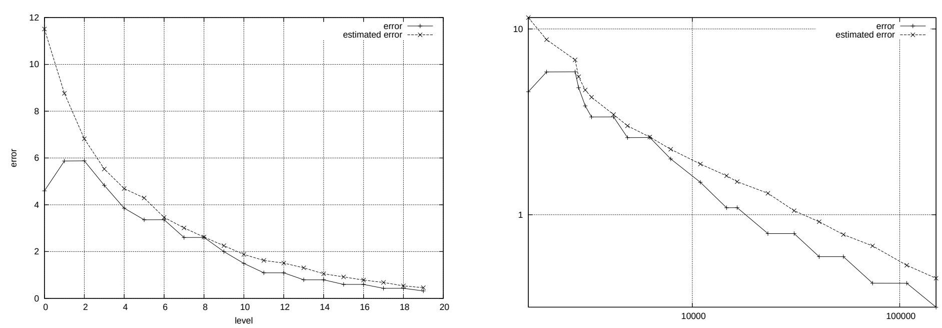


Figure 6.6: Error $\sigma \| u - u_{h,\Delta t} \|_{L^2((0,T);V)}$ and estimated error $(\sum_n (\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2} \sum_{T \in T_{nh}(\Omega)} \eta_{n,T}^2))^{\frac{1}{2}}$ as a function of the refinement level (left) and of the degrees of freedom (dof) on different levels (right).

level	dof	error	estimated error	refined variable
0	1680	4.60	11.51	S
1	2046	5.87	8.76	S
2	2848	5.88	6.82	t
3	2954	4.83	5.53	t
4	3166	3.86	4.69	t
5	3385	3.36	4.29	S
6	4303	3.36	3.46	t
7	4994	2.60	3.01	S
8	6437	2.60	2.62	t
9	8065	2.00	2.25	t
10	11102	1.49	1.88	t
11	15003	1.09	1.62	S
12	16773	1.09	1.51	t
13	23399	0.79	1.05	S
14	31533	0.79	1.05	t
15	42157	0.59	0.92	S
16	54819	0.59	0.78	t
17	75047	0.43	0.68	S
18	109915	0.43	0.53	t
19	153993	0.32	0.45	S

Table 6.1: Level, degrees of freedom (dof), error $\sigma \|u - u_{h,\Delta t}\|_{L^2((0,T);V)}$, estimated error $(\sum_n (\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2} \sum_{T \in T_{nh}(\Omega)} \eta_{n,T}^2))^{\frac{1}{2}}$, and type of the refined variable.

Aggressive Mesh Refinement

We can also use more aggressive mesh refinement strategies, but sometimes we may get a mesh which is too fine in some regions. The following figures display the results of splitting the elements into up to eight subelements (depending on the error indicators). The plots of the discrete solutions, the pointwise errors, the meshes and the time and space error indicators are similar to those obtained from the progressive refinement strategy. Therefore, we will only document the history of the refinement process.

Figure 6.7 shows the changes of the error $\sigma\|u - u_{h,\Delta t}\|_{L^2((0,T);V)}$ and the estimated error $(\sum_n(\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2}\sum_{T\in T_{nh}(\Omega)}\eta_{n,T}^2))^{\frac{1}{2}}$ as the refinement level (left) and the degrees of freedom (dof) (right) change.

Table 6.2 lists the values of the degrees of freedom (dof), the errors, the estimated errors, and indicates the type of the refined variable at each refinement step.

Figure 6.8 compares the errors $\sigma\|u - u_{h,\Delta}\|_{L^2((0,T);V)}$ and the estimated errors $(\sum_n(\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2}\sum_{T\in T_{nh}(\Omega)}\eta_{n,T}^2))^{\frac{1}{2}}$ obtained by the progressive and the aggressive refinement strategy.

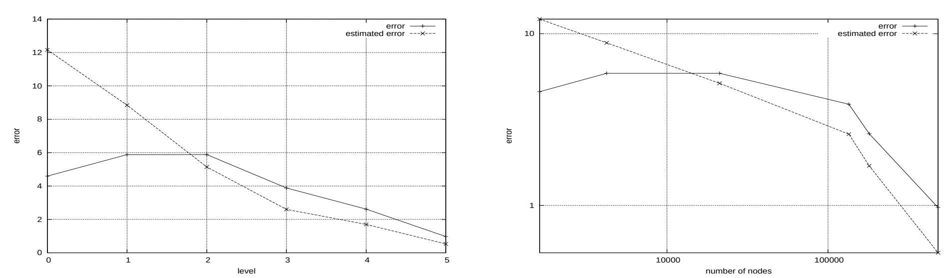


Figure 6.7: Aggressive mesh refinement: error $\sigma\|u - u_{h,\Delta}\|_{L^2((0,T);V)}$ and estimated error $(\sum_n(\eta_n^2+\frac{\Delta t_n}{\sigma_{min}^2}\sum_{T\in T_{nh}(\Omega)}\eta_{n,T}^2))^{\frac{1}{2}}$ as a function of the refinement level (left) and the degrees of freedom (dof) on different levels (right) of the refinement process.

level	dof	error	estimated error	refined variable
0	1680	4.60	12.16	S
1	4227	5.89	8.85	S
2	21229	5.88	5.14	S
3	134777	3.27	2.59	t
4	180237	2.62	1.73	t
5	482002	0.97	0.53	S

Table 6.2: Level, degrees of freedom (dof), error $\sigma\|u - u_{h,\Delta}\|_{L^2((0,T);V)}$, estimated error $(\sum_n(\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2} \sum_{T\in T_{nh}(\Omega)} \eta_{n,T}^2))^{\frac{1}{2}}$, and type of the refined variable.

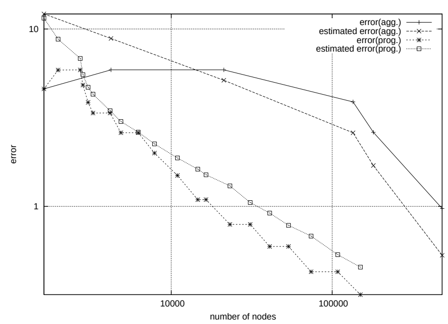


Figure 6.8: Progressive mesh refinement vs. aggressive mesh refinement: error $\sigma\|u - u_{h,\Delta}\|_{L^2((0,T);V)}$ and estimated error $(\sum_n(\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2} \sum_{T\in T_{nh}(\Omega)} \eta_{n,T}^2))^{\frac{1}{2}}$ as functions of the degrees of freedom (dof) on different levels.

Example 2. Local Volatility. We consider the data

- $K = 100$ (strike price)
- $T = 1$ (maturity 1 year)
- $\sigma = 0.05 + 0.25\mathbf{1}_{\frac{|S-100|^2}{400}+\frac{|t-0.5|^2}{0.01}\leq 1} + 0.25\mathbf{1}_{\frac{|S-100|^2}{400}+\frac{|t-0.9|^2}{0.01}\leq 1}$ (volatility)
- $r = 0.04$ (interest rate)

and compute the price of a vanilla European put in the rectangle $[0, 200] \times [0, 1]$. The initial mesh is a uniform mesh with 40 nodes in t and 100 nodes in S.

Progressive Mesh Refinement

Using the strategy as described at the beginning of this chapter we get the following results:

Figure 6.9 shows a visualization of the discrete solution after 2 refinement steps.

Figure 6.10 contains the refined meshes after $0, 5, 10$ and 19 refinement steps. We clearly observe a pronounced refinement in local regions where the volatility exhibits jumps.

Figure 6.11 displays the time error indicator η_n and the global S error indicator $(\frac{\Delta t_n}{\sigma^2_{min}} \sum_{T\in T_{nh}(\Omega)} \eta^2_{n,T})^{\frac{1}{2}}$ versus times after $0, 5, 10$ and 19 refinement steps. We observe that they are both large when the volatility jumps.

Figure 6.12 shows the space error indicator $\eta_{n,T}$ which is also large when the volatility jumps.

Figure 6.13 displays the changes of the estimated error $(\sum_n(\eta^2_n + \frac{\Delta t_n}{\sigma^2_{min}} \sum_{w\in T_{nh}} \eta^2_{n,T}))^{\frac{1}{2}}$ as the refinement level (left) and the degrees of freedom (dof) (right) change.

Table 6.3 contains the history of the refinement process in terms of the degrees of freedom (dof), the estimated error, and the type of the refined variable at each level.

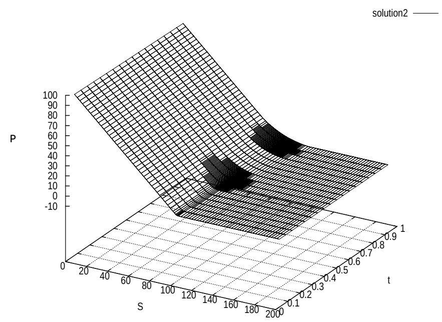
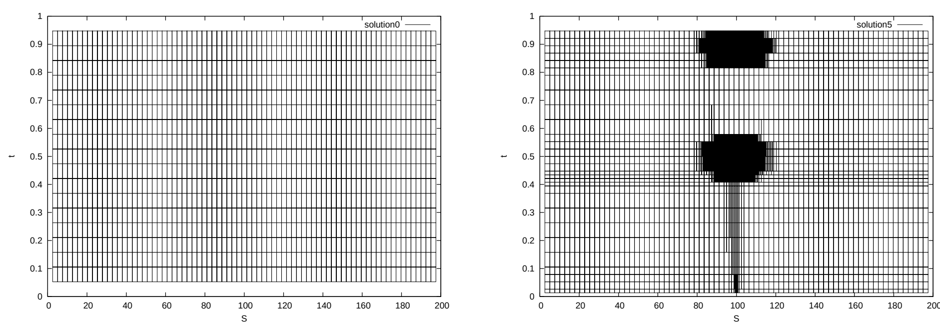


Figure 6.9: The discrete solution after 2 refinement steps.



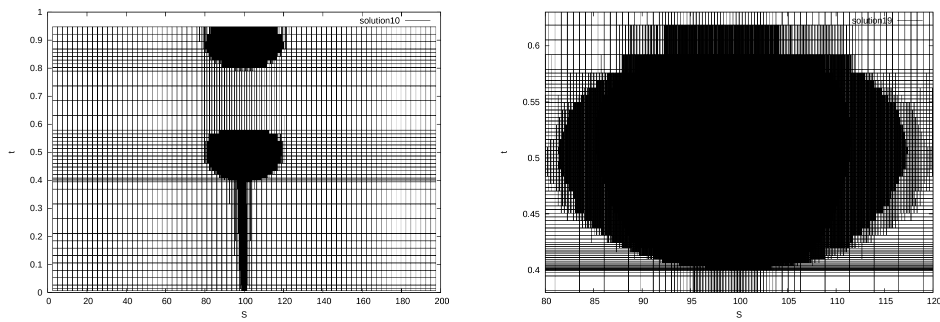


Figure 6.10: Refined meshes obtained after $0, 5, 10$ and 19 refinement steps. The bottom right figure is a zoom.

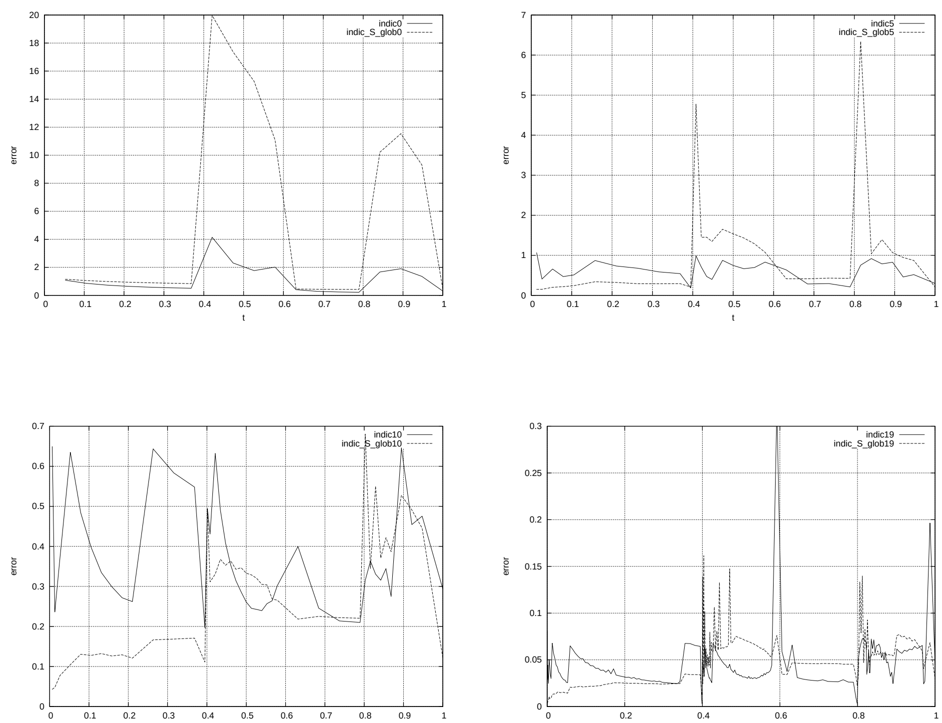


Figure 6.11: The time error indicator η_n and the global S error indicator $\left(\frac{\Delta t_n}{\sigma_{min}^2}\sum_{T\in T_{nh}(\Omega)}\eta_{n,T}^2\right)^{\frac{1}{2}}$ obtained after $0, 5, 10$ and 19 refinement steps.

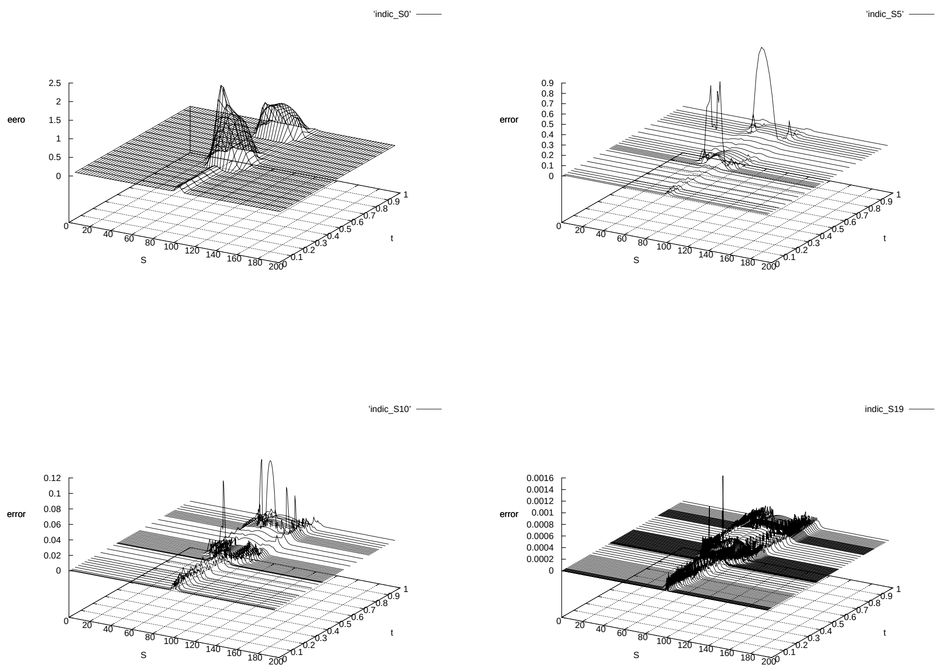


Figure 6.12: Space error $\eta_{n,T}$ obtained after $0, 5, 10$ and 19 refinement steps. The bottom right figure is a zoom.

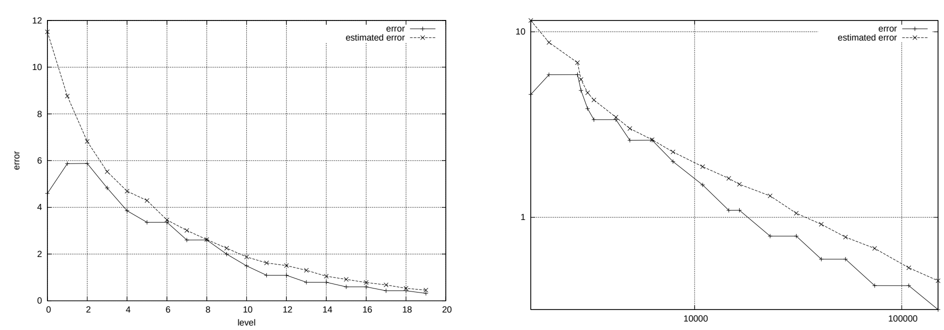


Figure 6.13: The error $\sigma \|u - u_{h,\Delta}\|_{L^2((0,T);V)}$ and the estimated error $(\sum_n (\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2} \sum_{T \in T_{nh}(\Omega)} \eta_{n,T}^2))^{\frac{1}{2}}$ as a function of the refinement level (left) and the degrees of freedom (dof) (right).

level	dof	estimated error	refined variable
0	4040	617.47	S
1	4288	317.90	S
2	4812	162.90	S
3	5860	85.13	S
4	8107	45.00	S
5	12688	24.61	S
6	21910	14.33	S
7	40374	9.17	S
8	77309	6.58	t
9	80486	6.58	S
10	80547	6.03	t
11	117803	6.20	S
12	118691	5.17	t
13	145762	5.21	S
14	146617	4.67	S
15	148631	4.45	S
16	179365	3.84	t
17	276324	3.69	S
18	276732	3.45	S
19	278093	3.36	S

Table 6.3: Level, degrees of freedom (dof), estimated error $(\sum_n(\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2}\sum_{T\in T_{nh}(\Omega)}\eta_{n,T}^2))^{\frac{1}{2}}$, and the type of the refined variable.

Aggressive Mesh Refinement

We present the results when using the aggressive mesh refinement strategy which splits the elements marked for refinement into up to eight subelements (depending on the error indicators).

Figure 6.14 shows the estimated error $(\sum_n(\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2}\sum_{T\in T_{nh}(\Omega)}\eta_{n,T}^2))^{\frac{1}{2}}$ as a function of the refinement level (aggressive strategy) (left) and of the degrees of freedom (dof) (right) for aggressive and progressive refinement.

Table 6.4 lists the degrees of freedom (dof), the values of the estimated errors, and the type of the refined variable at each level of the refinement process.

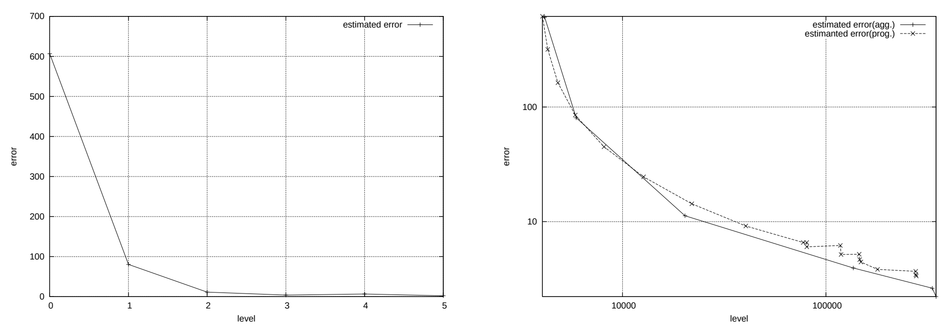


Figure 6.14: Aggressive mesh refinement: Estimated error $(\sum_n(\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2}\sum_{T\in T_{nh}(\Omega)}\eta_{n,T}^2))^{\frac{1}{2}}$ as a function of the refinement level (left) and the degrees of freedom (dof) at different levels (right).

level	dof	estimated error	refined variable
0	4141	606.64	S
1	5921	80.50	S
2	20278	11.27	S
3	136463	3.95	t
4	333592	2.63	S
5	347960	2.22	S

Table 6.4: Level, degrees of freedom (dof), estimated error $(\sum_n(\eta_n^2 + \frac{\Delta t_n}{\sigma_{min}^2}\sum_{T\in T_{nh}(\Omega)}\eta_{n,T}^2))^{\frac{1}{2}}$, and the type of the refined variable.

Chapter 7

Conclusions

The a posteriori error analysis presented in this thesis is restricted to the Black-Scholes equation for European options on a single equity share. It can be extended to European basket options, i.e., options on more than one stock. Likewise, the generalization to exotic options such as single- or double-barrier European basket options should be feasible. However, there is a natural limitation on the number of stocks in a basket, since with an increasing number of stocks one is faced with the curse of dimensionality. The numerical solution of such high-dimensional problems by appropriate reduction techniques is currently in the focus of active research.

As far as American options are concerned, the situation is more difficult, since they give rise to parabolic variational inequalities of obstacle type. Although adaptive finite elements have been studied recently for variational inequalities associated with elliptic boundary value problems [14, 15], the parabolic case is still widely unknown territory.

Bibliography

[1] Y. Achdou and O. Prionneau; *Computational Methods for Option Pricing.* SIAM, Philadelphia, 2005.

[2] M. Ainsworth, J.T. Oden; *A Posteriori Error Estimation in Finite Element Analysis.* Wiley, Chichester, 2000.

[3] Black, Fischer and Myron Scholes;*The Pricing of Options and Corporate Liabilities.* Journal of Political Economy. **81**, 637-654, 1973.

[4] I. Babuska, T. Strouboulis; *The Finite Element Method and its Reliability.* Clarendon Press, Oxford, 2001.

[5] W. Bangerth, R. Rannacher; *Adaptive Finite Element Methods for Differential Equations.* Lectures in Mathematics. ETH-Zürich. Birkhäuser, Basel, 2003.

[6] A. Bergam, C. Bernardi, and Z. Mghazli; *A posteriori analysis of the finite element discretization of a nonlinear parabolic equation.* Math. Comp. **74**, 1097–1116, 2005.

[7] P. Binev, W. Dahmen, R. DeVore. *Adaptive Finite Element Methods with Convergence Rates.* Numer. Math. **97**, 219–268, 2004.

[8] F. Black and M. Scholes; *The pricing of options and corporate liabilities.* J. Pol. Econ. **81**, 637–659, 1973.

[9] T. Bollerslev, R.F. Engle, and D.B. Nelson; *ARCH models.* In: Handbook of Econometrics (R.F. Engle and D. McFadden (eds.), pp. 2959–3038, Elsevier, Amsterdam, 1986.

[10] F. Bornemann; *An adaptive multilevel approach to parabolic equations I. General theory and 1D-implementation.* IMPACT Comput. Sci. Engrg. **2**, 279–317, 1990.

[11] F. Bornemann; *An adaptive multilevel approach to parabolic equations II. Variable order time discretization based an a multiplicative error correction.* IMPACT Comput. Sci. Engrg. **3**, 93–122, 1991.

[12] F. Bornemann; *An adaptive multilevel approach to parabolic equations III. 2D error estimation and multilevel preconditioning.* IMPACT Comput. Sci. Engrg. **4**, 1–45, 1992.

[13] D. Braess; *Finite Elements. 3rd Edition.* Cambridge University Press, Cambridge, 2007.

[14] D. Braess, C. Carstensen, and R.H.W. Hoppe; *Error reduction in adaptive finite element approximations of elliptic obstacle problems.* J. Comput. Math. **27**, 170–183, 2009.

[15] D. Braess, R.H.W. Hoppe, and J. Schöberl; *A posteriori estimators for obstacle problems by the hypercircle method.* Comput. Visual. Sci. **11**, 351–362, 2008.

[16] S.C. Brenner and L.R. Scott; *The Mathematical Theory of Finite Element Methods. 2nd Edition.* Springer, Berlin-Heidelberg-New York, 2002.

[17] C. Carstensen and R.H.W. Hoppe; *Convergence analysis of an adaptive nonconforming finite element method.* Numer. Math. **103**, 251-266, 2006.

[18] C. Carstensen and R.H.W. Hoppe; *Error reduction and convergence for an adaptive mixed finite element method.* Math. Comp. **75**, 1033-1042, 2006.

[19] J.M. Cascon, C. Kreuzer, R.H. Nochetto, and K.G. Siebert; *Quasi-Optimal Convergence Rate for an Adaptive Finite Element Method.* SIAM J. Numer. Anal. **46**, 2524–2550, 2008.

[20] Z. Chen, F. Jia; *An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems.* Math. Comp. **73**, 1167-1193, 2004.

[21] P.G. Ciarlet; *The Finite Element Method for Elliptic Problems. Rev. Edition.* SIAM, Philadelphia, 2002.

[22] P.G. Ciarlet, P.G., J.-L. Lions (eds.); *Handbook of Numerical Analysis. Volume II. Finite Element Methods. Part 1.* Elsevier, Amsterdam, 1990.

[23] R. Cont and P. Tankov; *Financial Modelling with Jump Processes.* Chapman and Hall, Boca Raton, 2003.

[24] W. Dörfler. *A convergent adaptive algorithm for Poisson's equation.* SIAM J. Numer. Anal., **33**, 1106–1124, 1996.

[25] B. Dupire; *Pricing with a smile.* Risk **7**, 18–20, 1994.

[26] K. Eriksson, D. Estep, P. Hansbo, C. Johnson; *Computational Differential Equations.* Cambridge University Press, Cambridge, 1996.

[27] L.C. Evans; *Partial Differential Equations.* American Mathematical Society, Providence, 1983.

[28] R.H.W.Hoppe; *Vorlesungsskript SS 2007: Numerische Mathematik für WiMAs II.* Universität Augsburg, Institut für Mathematik, Augsburg, 2007.

[29] R.H.W. Hoppe and J. Schöberl; *Convergence of adaptive edge element methods for the 3D eddy currents equations.* J. Comput. Math. **27**, 657–676, 2009.

[30] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1991. Second Edition.

[31] D. Lamberton and B. Lapeyre, *Introduction au Calcul Stochastique Appliqué á la Finance*, Ellipses, Paris, 1997.

[32] R.C. Merton; *Theory of rational option pricing.* Bell J. Econom. Management Sci. **4**, 141–183, 1973.

[33] P. Morin, R.H. Nochetto and K.G. Siebert. *Data oscillation and convergence of adaptive FEM.* SIAM J. Numer. Anal., **38**, 2, 466–488, 2000.

[34] P. Neittaanmäki, S. Repin; *Reliable methods for mathematical modelling. Error control and a posteriori estimates.* Elsevier, New York, 2004.

[35] M. Renardy and R.C. Rogers; *An Introduction to Partial Differential Equations.* Springer, Berlin-Heidelberg-New York, 1993.

[36] A. Revuz and M. Yor, *Continuous Martingale Calculus*, Springer- Verlag, New York, 1990.

[37] R. Seydel; *Tools for Computational Finance. 2nd Edition.* Springer, Berlin-Heidelberg-New York, 2004.

[38] R. Stevenson; *Optimality of a standard adaptive finite element method.* Found. Comput. Math. **7**, 245–269, 2007.

[39] V. Thomée; *Galerkin Finite Element Methods for Parabolic Problems.* Springer, Berlin-Heidelberg-New York, 1997.

[40] R. Verfürth; *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Wiley-Teubner, New York, Stuttgart, 1996.

[41] R. Verfürth,*A posteriori error estimates for nonlinear problems: $L^r(0,T;L^p(\Omega))$-error estimates for finite element discretizations of parabolic equations*, Math. Comput. , 67: 1335-1360, 1998.