

MASTER

Inverse problems in Black-Scholes option price forecasting with locally reconstructed volatility

Wolf, J.

Award date:
2015

[Link to publication](#)

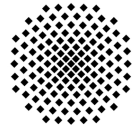
Disclaimer

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain



University of Stuttgart
Germany

Department of Mathematics and Computer Science
Centre for Analysis, Scientific computing and Applications

Inverse Problems in Black-Scholes: Option Price Forecasting with Locally Reconstructed Volatility

Master's Thesis

Jan-Philipp Wolf

Supervisors:

Prof. dr. Sorin Pop, TU Eindhoven (NL), University of Bergen (NO)

Dr. Michiel Hochstenbach, TU Eindhoven (NL)

Dr.ir. Jan ten Thije Boonkamp, TU Eindhoven (NL)

Prof. dr. Christian Rohde, Universität Stuttgart (DE)

SimTech No. 6

Eindhoven, 30th November 2015

Abstract

The Black-Scholes equation for option pricing is a famous model in financial mathematics. Given a strike price and expiration date it tries to determine at what price an option should trade today. Lots of research has been carried out in determining the validity of its suppositions, implementation details and extensions to accommodate for more realistic assumptions. Noting the equivalence of the Black-Scholes equation to the heat transfer equation, reversing the time in this equation leads to an *ill-posed* problem.

First, we present a way to treat this problem mathematically and numerically to forecast option prices via the solution of a time inverse parabolic problem based on the work of Klivanov [29]. This model introduces new boundary and initial conditions and leaves behind traditional notions like the strike price and expiration date of an option contract. We try to identify the most efficient approach for the numerical solution of the problem.

Secondly, we aim to improve prediction quality by incorporating more realistic market volatility data. In fact, obtaining volatility implied by today's option price requires the solution of another nonlinear inverse problem. Depending on the assumptions we impose on the dependence of volatility on maturity and strike price the complexity of problems and methods used to solve them varies greatly. Due to its computational simplicity we follow a linearization approach presented by Bouchaev & Isakov [11, 21]. The reconstructed volatility surface resulting from this method could then be used in the forecasting model and results be compared to the volatility function used in Klivanov [29]. However, note that this is merely an experiment because convergence of the approximate solution to the exact has not been proved yet in this case.

Next, we discuss a trading strategy based on the result of the time inverse model and back test this strategy with real market data. We assert that trading only makes sense if the bid ask spread of the option's underlying is large enough and see this conjecture confirmed when back testing. Finally, we are also interested in the prediction quality for different boundary conditions and the impact of decision criteria on the strategy's performance.

Preface

”A gift is pure when it is given from the heart to the right person at the right time and at the right place, and when we expect nothing in return.”

The Bhagavad Gita

After five years of study across multiple disciplines and at two universities there are quite a few people I feel indebted to. Let me start off by thanking my parents and my sister for their relentless and subtle support all the way through my academic life. I appreciate that they allowed me to follow my interests and never interfered with any of my plans.

Secondly, I would like to thank prof. Rohde in Stuttgart and prof. Pop in Eindhoven for supervising a self initiated project that did not fully match their field of expertise. However, prof. Pop’s oversight allowed me to work with just the right specialists at TU/e. I want to express my thankfulness to dr. Michiel Hochstenbach for his help concerning inverse problems and to dr. Jan ten Thije Boonkamp for his insights on numerical aspects. Furthermore, I express my gratitude to Mikhail Klivanov from the University of North Carolina at Charlotte for answering my questions regarding his paper.

Thirdly, I have to express my special gratitude to all my fellow students in Stuttgart who I remember for their kind helpfulness and willingness to try to understand a problem by its root. I largely enjoyed studying in groups and I will cherish these hours for the years to come. During my stay in Eindhoven I have met some of the most amazing people from all around the world. Thank you all for making this period of my life one of the most memorable experiences imaginable and broadening my horizon manifold.

Finally, thanks to Joos Buijs and Thijs Nugteren for providing the L^AT_EXtemplate for this thesis.




































Jan-Philipp Wolf
Eindhoven, October 2015

Contents

Contents	vii
List of Figures	ix
List of Tables	xi
1 Introduction	1
2 Preliminaries	5
2.1 Black-Scholes model (BSM)	5
2.2 On implied volatility	7
2.3 Model shortcomings and criticism	8
3 Time Inverse Parabolic Problem: Option Price Forecasting	9
3.1 In analogy to the heat equation	9
3.2 Option price forecasting: Time inverse BSM	11
3.3 Solution approaches	12
3.3.1 Error estimation	12
3.3.2 Tikhonov regularization	13
3.4 Convergence results: Carleman estimates	13
3.4.1 Carleman estimate	14
3.4.2 Stability estimates and Regularization	15
3.4.3 Discretization	17
3.4.4 LSQR	21
3.4.5 Minimization via Conjugate gradient method	22
3.4.6 Multifrontal Sparse QR factorization	22
3.5 Numerical Experiments	23
3.5.1 The backward-time problem: Crank-Nicolson	23
3.5.2 Verification test: time inverse heat solution	23
3.5.3 Assessment: How to include boundary conditions?	24
3.5.4 Bid-Ask spread requirements	24
3.5.5 Choice of regularization parameter	26
3.5.6 Numerical instability of gradient method and lsqr algorithm	28
3.5.7 Alternative Boundary conditions	28
4 Parameter Identification: Implied Volatility as an Inverse Problem	31
4.1 Ill-posedness of the inverse problem and literature overview	32
4.2 Linearization approach by Isakov	32
4.2.1 Continuous time formula: Dupire	32
4.2.2 Linearization at constant volatility	33
4.2.3 Uniqueness and Stability	34
4.2.4 Numerical solution of the forward problem	35
4.2.5 Numerical tests: reconstruction of given volatility surfaces	36

4.2.6	Towards a real world scenario: Smooth interpolation of market prices and remarks	37
5	Putting everything into action: A large-scale real market data study	39
5.1	Option exchanges & BATS	39
5.2	Data gathering	39
5.3	Data filtering	39
5.4	Trading strategy	40
5.5	Backtesting	40
5.5.1	Restriction on options with significant bid-ask spread: A Comparison . . .	41
5.5.2	Modification: Negativity of extrapolated ask price	41
5.5.3	Impact of boundary conditions and decision criterion	42
5.5.4	Influence of volatility coefficient	42
5.5.5	Test on most actively traded options	42
5.5.6	Alternative one-day trading strategy	42
5.5.7	Forecasting with locally reconstructed implied volatility: Future work . . .	44
6	Conclusions	45
	Bibliography	47

List of Figures

1.1	First known printed forward contract for a share in the Dutch West India Company worth 500 Flemish pounds at 178.5% with a maturity of one year. This print dates back to 1629.	2
1.2	Payoff  and profit  for a call option on the left respectively put option on the right with strike price $K = 1$	3
1.3	Payoff  and profit  for a short call option on the left respectively short put option on the right with strike price $K = 1$	4
3.1	Heat equation solved forward  and backward 	10
3.2	Price of an option with strike price $K = 50$ €, $T = 5/12$ years, in dependence of spot price S . Verification of analytical solution as calculated by Matlab for call  and put  vs. Crank-Nicolson scheme for call  and put 	24
3.3	Heat equation solved backward with noise level $\delta = 10^{-3}$  , $\delta = 10^{-2}$  and $\delta = 10^{-1}$  and true initial distribution $\psi(x)$  for different terminal times T_e	25
3.4	Solution of price forecasting for set boundaries  vs. free boundaries  compared to forward solution of the problem as obtained by Crank-Nicolson method 	25
3.5	L-Curve in double logarithmic scale for options with different bid ask spreads of the underlying: large spread $\Delta = 2.6\% \frac{s_a + s_b}{2}$  vs. tight spread $\Delta = 0.12\% \frac{s_a + s_b}{2}$  for regularization parameter $\lambda \in (10^{-8}, 1)$	26
3.6	GCV function $G(\lambda)$ for the full range of regularization parameters $\lambda \in (10^{-8}, 1)$ on the left and zoomed in around the minimum on the right.	27
3.7	Solution as obtained using different methods: reference solution  , Matlab's backslash (Multifrontal QR)  , Matlab's LSQR  and backslash solution of the gradient system  . Regularization parameter was set to $\lambda = 0.1$	29
3.8	Price using quadratic polynomial  , +5%,  , -5%, regression  , +5%,  , -5%, mean  , +5%,  , -5%, weighted average  , +5%,  , -5%.	29
4.1	Implied volatility surface for MSFT options. Data retrieved on 05.07.2015.	31
4.2	Reconstructed volatility surface  compared to prescribed volatility surface 	37
4.3	Price of put option for different strikes  on MSFT stock with maturity date 17.06.2016 as observed on 14.09.2015 with smoothing spline 	38

List of Tables

5.1	Comparison between samples from all options versus samples from options with significant bid-ask spread.	42
5.2	Effects of incorporating the modified trading rule mentioned in Section 5.5.2. . . .	42
5.3	Impact of different combinations of boundary conditions and decision criteria as presented in Section 3.4.3.	43
5.4	Experiment: Linear regression function as volatility.	43
5.5	Experiment: Most actively traded options.	43
5.6	Experiment: One-day trading strategy.	43

Chapter 1

Introduction

Prediction is very difficult, especially if it is about the future.

Markus M. Ronner, 1918

Correct attribution is hard, especially for the past.

Doug Arnold, 2010

Since this thesis deals with the prediction of option prices we give a short overview over the classes of market prediction techniques and their relation to the term market efficiency. Additionally, we outline the history of option trading to better understand the usefulness of this financial device to the world economy. Finally, modern day terminology and technical details are introduced and instructive examples serve to understand the mechanics of options trading, the leverage effect as well as some of the risks involved.

First of all, the question whether the prediction of market movement is possible at all hinges on the question of market efficiency. A market is deemed efficient if all relevant information is always incorporated into the current share prices. The efficient-market hypothesis then states that it is consequently impossible for anybody to outwit the market and stock prices behave according to the closely related random walk hypothesis which states that stock prices are governed by a random walk and are therefore unpredictable. However, the joint hypothesis problem refers to the fact that testing markets for efficiency is impossible because any such test would have to involve an equilibrium asset pricing model which by definition is not able to capture anomalous behavior. Any evidence for anomalous market behavior may therefore be the result of market inefficiency or an insufficient model. In short, there is no way to tell where anomalies originate from if equilibrium asset pricing models are used and therefore it is impossible to prove market efficiency. In fact, there are now three forms of market efficiency in existence. Namely weak-form, semi-strong-form and strong-form market efficiency. They vary in the degree to which information is incorporated into current share prices and the speed in which they are incorporated. For more information on these variations of market efficiency see e.g. Jones [23]. Behavioral finance or economics tries to explain market inefficiencies by analyzing prevalently human factors like irrational decision making, mental emotional filters and the like.

Stock market prediction has existed as long as modern day trading exists. The most important accounts are made by Joseph de la Vega [47] who describes aspects of technical analysis in Dutch markets as early as in the 17th century. There are three broad categories of prediction methods: Fundamental analysis, technical analysis and alternative methods. Whereas fundamental tries to make predictions based on an analysis of the stock issuing company technical analysis is concerned with finding patterns in past prices using e.g. time series analysis. Statistical techniques as various moving average methods and the use of candlesticks are widely applied as well. The third and most recent development concerns machine learning i.e. making predictions using feed forward neural networks.

This project is a challenging mathematical problem with a link to large real world data sets and application to but not limited to finance. For future works there is a connection of the mathematical field of inverse problems to machine learning to be explored. However, this project is based on a recent publication dealing with option price forecasting in connection with a classic inverse problem by Klibanov [29]. The beauty of mathematics is that the methods used in this topic are directly transferable to other inverse problems ranging from inverse scattering Au [4], electrocardiography Sarikaya [41], precision measurement in chip production Soulan [44].

The previous section dealt with stock market prediction whereas this thesis is about option price prediction. Let us clarify the terminology and connection between stocks and options now.

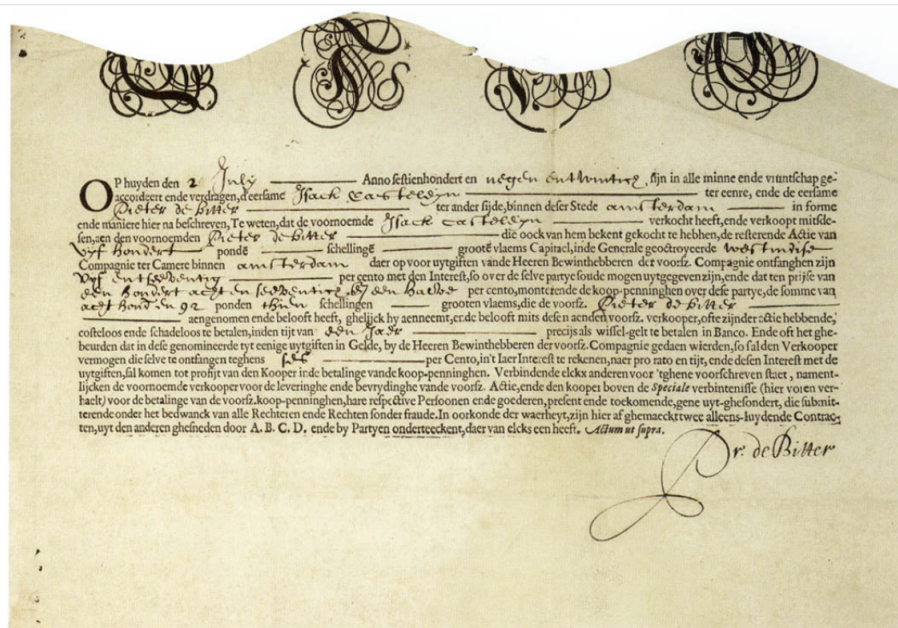


Figure 1.1: First known printed forward contract for a share in the Dutch West India Company worth 500 Flemish pounds at 178.5% with a maturity of one year. This print dates back to 1629.

An option is the *right* but not the *obligation* to buy or sell an agreed upon amount of stock at a certain point of time in the future at a predetermined price. The price paid for this right is called the option *premium*. When one obtains the option to buy stock this is referred to as a *call* option whereas when one obtains the right to sell stock this is referred to as a *put* option. Since the value of an option depends on the underlying, stock options are *derivatives* i.e. they derive their value from another financial product. Options are not restricted to stocks but are also available on commodities, interest rates, foreign currencies and bonds.

To better understand the (original) purpose of this special instruments let us have a look at the history of forward contracts and options. Forward contracts are closely linked to options with the only difference being that the buying respectively selling is not optional but mandatory. Forward trading of commodities as grain or salt had already appeared in medieval times. Indeed one of the first option buyers was the Greek philosopher Thales of Miletus who bought the right to rent olive presses in anticipation of a larger than usual olive harvest. When this harvest turned out to be larger than usual indeed he rented out the spaces for a higher price, thus making him one of the first option exercisers Saunders [42]. Even though forward contracts have also been found on Mesopotamian clay tablets written in cuneiform scripture dating back as far as 1750 BC, there are few records on how this form of financial instrument was traded among business men and how prices were set. Insightful records are only available from the 17th century with Joseph Penso de la Vega's *Confusion de confusiones* [47] from 1688 being one of the most important resources especially for the situation in the Netherlands. Around the year 1500 is the time when with the rise of a system of periodical fairs trading in options became common for people who were not directly involved in trading the underlying commodities. Venice became a city of regular fairs. The first permanent markets were to be found in Antwerp, Belgium with foreign merchants settling in the Scheldt port after 1500 Goetzmann [16]. Another milestone are the Dutch voyages to Asia starting around 1590 which resulted in the emergence of companies like VOC, Dutch East and West Company as well as in a spike in futures trading. One example of how such a contract looked like at that time can be seen in Fig. 1.1. Furthermore, the public opinion regarding derivatives had changed from a negative view influenced by the Roman laws under which it was regarded as taking chances and therefore equated to betting to a more liberal one where regulation was

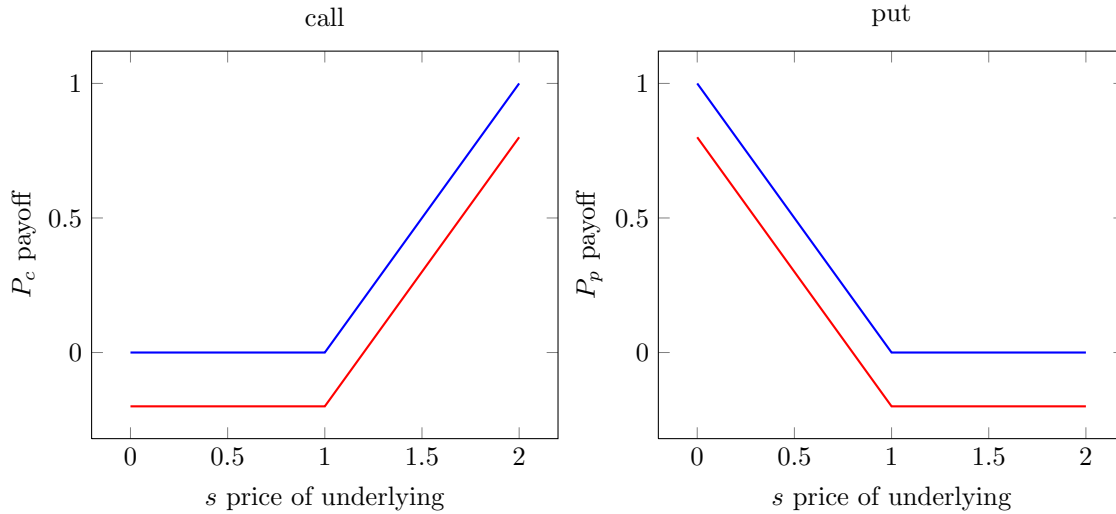


Figure 1.2: Payoff — and profit — for a call option on the left respectively put option on the right with strike price $K = 1$.

demanded rather than outright bans. For more information on the history of derivative trading we refer to Goetzmann [16] from which lots of inputs for this section have been taken. As stated in Haug [20] there were active option exchanges in London, Paris and New York as well as other European markets in the late 1800 and early 1900. Furthermore there was a Foreign Exchange Option Market from 1917 to 1921 as reported by Mixon [35]. There is proof that traders used sophisticated techniques to price options in the American equity option in 1870 (see Kairys [24]) or on the Johannesburg Stock Exchange in the early 20th century (see Moore [36]). With the installation of the first transatlantic telegraph cable in 1858 by Cyrus West Field and his Atlantic Telegraph Company arbitrage trading emerged. From Haug [20]: "Although American securities had been purchased in considerable volume abroad after 1800, the lack of quick communication placed a definite limit on the amount of active trading in securities between London and New York markets, (see Weinstein, 1931). Furthermore, one extant source, Nelson (1904), speaks volumes: an option trader and arbitrageur, S.A. Nelson published a book *The ABC of Options and Arbitrage* based on his observations around the turn of the twentieth century. The author states that up to 500 messages per hour and typically 2000 - 3000 messages per day were sent between the London and the New York market through the cable companies. Each message was transmitted over the wire system in less than a minute. In a heuristic method that was repeated in *Dynamic Hedging* by one of the authors, (Taleb, 1997), Nelson, describe in a theory-free way many rigorously clinical aspects of his arbitrage business: the cost of shipping shares, the cost of insuring shares, interest expenses, the possibilities to switch shares directly between someone being long securities in New York and short in London and in this way saving shipping and insurance costs, as well as many more similar tricks."

Nowadays, options on stocks, commodities, foreign currencies, bonds and interest rates are traded on exchanges all around the world. Along with options on different kinds of underlying there are also different styles of options. The most common option is the *European* style option. European options can only be executed *on* the day of maturity. On the contrary, *American* options can be executed *on or before* the day of maturity. These two most common types of options are also called *vanilla* options. Furthermore, there are *binary* options where the payoff is either a fixed amount or nothing and all kinds of *exotic* options which could involve foreign exchange rates as in *quanto* options, depend on multiple indexes as with *basket* options or depend on the value of the underlying not only at maturity but also on multiple dates before maturity (*Asian, lookback, barrier, digital, spread options*, etc). Note that these options are usually traded over

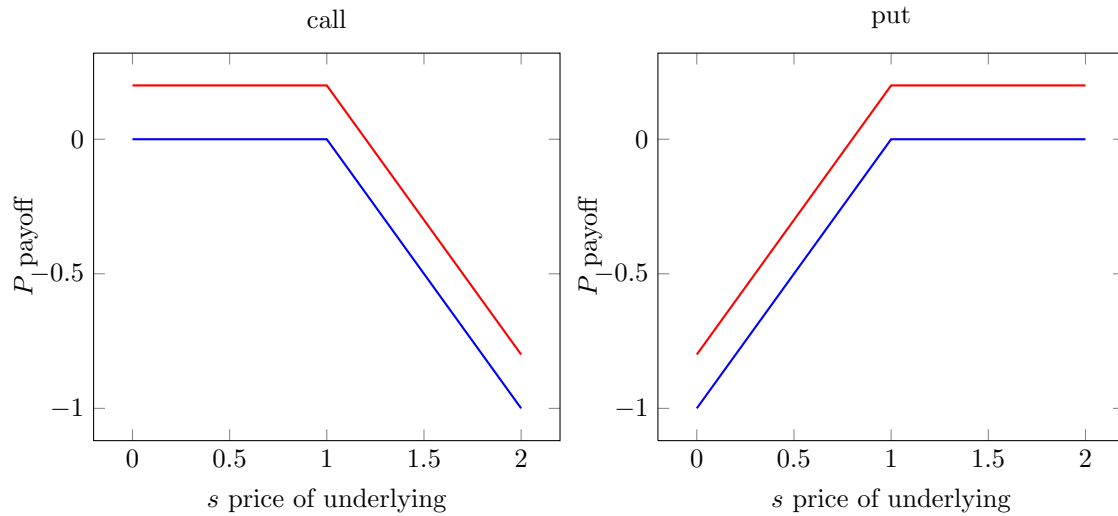


Figure 1.3: Payoff — and profit — for a short call option on the left respectively short put option on the right with strike price $K = 1$.

the counter (OTC). To get a feel for the significance of this kind of trading note that in 2013 the Chicago Board Options Exchange (CBOE) which is the biggest American option exchange reported a trading volume of 561.503.190.262\$. To better understand the link between stock price and option price and to understand how options work take a look at the payoff diagrams for a call and put option with strike price $K = 1\$$ in Fig. 1.2. The payoff — in the left plot of a call option is $P_c = \max(s - K, 0) = (s - K)^+$. The difference between payoff and profit is the option premium, i.e. the amount the buyer pays to obtain the right to exercise the option. Conversely, there is a payoff from a put option (plot on the right in Fig. 1.2) once the underlying price s at maturity is lower than the strike price and one profits from this option once this differences is greater than the premium paid to the put option seller. The profit from a put option is therefore $P_p = \max(K - s, 0) = (K - s)^+$. From these quantities another vital relation can be derived. The equality $P_c - P_p = (s - K)^+ - (K - s)^+ = s - K$ is called the put-call parity providing a direct link between the price of a put and the price of the corresponding call option. Furthermore, this means that only either the call *or* the put has to be priced. Next consider the case of writing call and put options. In the case of a short call this means that one sells the right to buy shares respective in the case of a short put the right to sell shares to a second party. The payoff and profit diagrams for these kinds of options are presented in Fig. 1.3. Note that short call options are especially risky since there is no limit on possible losses. Finally, it is worth mentioning that leverage can be created through options. Assume you have \$ 1000 to invest. Instead of buying 10 shares of company ABC at \$ 100 you could buy five option contracts at e.g. 200 a piece allowing you to control 500 shares instead of 10. Note that one option contract is usually written on 100 shares. This example also reveals the risk associated with option trading.

Chapter 2

Preliminaries

The underlying model for all the considerations to follow is the option pricing model of Fischer Black and Myron Scholes [8] from 1973. Therefore, we will explain the model's assumptions and simplifications, derive the model equation and a closed form solution for European style options ¹. Furthermore the model's limitations will be analyzed and the reader will be made aware of possible modifications improving the accuracy of results. In fact, one of the most delicate quantities in this model is the volatility coefficient σ . Historical data reveals that constant volatility as assumed in the standard Black-Scholes model can not reproduce real world market data. For this reason, strategies to better model volatility will be presented. Finally, we deal with the criticism BSM has received and we outline the similarities with older pricing options which have been used since the 19th century (see e.g. Haug & Taleb [20]).

2.1 Black-Scholes model (BSM)

A short derivation of the Black-Scholes equation is given. The aim is to calculate a fair price of an European option given the underlying does not pay dividends. A European call option is the right to buy an asset at a specified *strike price* K in time T from today. In this case the option writer is obliged to sell the asset. The price paid for the right to execute the deal by the buyer respectively the money received by the option issuer is called the *premium*. In the case of a put option the buyer of the option has the right to sell the asset at price K in time T whereas the option issuer is obliged to buy the asset. In 1973 Fischer Black and Merton Scholes [8] proposed their famous model on how to price European options with riskless interest rate r , where the underlying does not pay dividends. The basic idea is to buy and short sell a certain amount of options at the same time to eliminate risk. This is a list of assumptions on both assets and the market:

1. The risk-free rate of return is constant in time.
2. The underlying (i.e. the stock) does not pay dividends.
3. The logarithmic returns of the stock follow an infinitesimal geometric Brownian motion with constant drift and volatility.
4. The market allows for no arbitrage.
5. Fractional buy and sell positions on the stock are possible at all times.
6. Fractional lending and borrowing at the riskless rate r is possible at all times.
7. Frictionless market: there are no transaction fees or costs.

We assume the price of an option's underlying s follows a geometric Brownian motion:

$$\frac{ds}{s} = \mu dt + \sigma dW. \quad (2.1)$$

This signifies that the uncertainty is coming from the term Brownian motion $W(t)$. In a geometric Brownian motion the logarithm of the quantity (in this case s) follows a Brownian motion. Here, μ is the expected return of the stock and σ is related to the standard deviation and is called

¹European style: the option can only be exercised *on* the expiration date.

volatility of the log returns in percent per year. Since $dW(t)$ is a Wiener process the following properties hold:

$$E[dW] = 0, \quad E[dW^2] = dt. \quad (2.2)$$

To understand the relation between the volatility σ and the variance of ds consider

$$E[ds] = E[s\mu dt + \sigma sdW] = E[s\mu dt] + E[\sigma sdW] = s\mu dt + \sigma sE[dW] = s\mu dt, \quad (2.3)$$

$$\text{Var}[ds] = E[ds^2] - E[ds]^2 = E[(s\mu dt + \sigma sdW)^2] - (s\mu dt)^2 = \sigma^2 s^2 E[dW^2]. \quad (2.4)$$

Since the standard deviation is the square-root of the variance, the volatility σ is proportional to $\frac{\sqrt{\text{Var}[ds]}}{s\sqrt{dt}}$. We now want to find a fair price $u(s, t)$ for an option in dependence on the time t and value of the underlying s , where we know the price at maturity $u(s, T)$. Since s follows a stochastic process we introduce Ito's lemma for two variables.

Lemma 1 Assume X_t is an Ito drift diffusion process satisfying

$$dX_t = \mu_t dt + \sigma_t dB_t, \quad (2.5)$$

where B_t is a Brownian motion. Also assume that $f \in C^2$. Then, in the limit $dt \rightarrow 0$ we have

$$df = \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{\sigma_t^2}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + \frac{\partial f}{\partial x} dB_t. \quad (2.6)$$

For a proof see e.g. Karatzas [25].

We derive and apply Ito's lemma to derive the Black-Scholes equation. The Taylor expansion of $u(t, s)$ is

$$du = \frac{\partial u}{\partial t} dt + \frac{\partial u}{\partial s} ds + \frac{1}{2} \frac{\partial^2 u}{\partial s^2} ds^2 + O(dt^2, ds^3). \quad (2.7)$$

Recalling equation (2.1) we substitute $s(\mu dt + \sigma dW)$ for ds :

$$du = \left(\frac{\partial u}{\partial t} + \frac{\partial u}{\partial s} \right) (s\mu dt + \sigma s dW) + \frac{1}{2} \frac{\partial^2 u}{\partial s^2} (s^2 (\mu^2 dt^2 + 2\mu dt \sigma dW + \sigma^2 dW^2)) + O(dt^3, dW^3). \quad (2.8)$$

Forming the limit $dt \rightarrow 0$ and considering that dB^2 is $O(dt)$ we obtain

$$du = \left(\frac{\partial u}{\partial t} + s\mu \frac{\partial u}{\partial s} + \frac{1}{2} \sigma^2 s^2 \frac{\partial^2 u}{\partial s^2} \right) dt + \sigma s \frac{\partial u}{\partial s} dW. \quad (2.9)$$

We now consider a portfolio containing one short option and $\frac{\partial u}{\partial s}$ shares of the option's underlying. Let the value of this so-called delta hedge portfolio at time t be denoted by $D(t) = -u + \frac{\partial u}{\partial s} s$. Let $\Delta t > 0$ be a time discretization and consider the portfolio's change in value over the time interval $[t, t + \Delta t]$:

$$\Delta D = -\Delta u + \frac{\partial u}{\partial s} \delta s. \quad (2.10)$$

Discretizing equations (2.1) and (2.9) and inserting them into (2.10) gives

$$\Delta D = \left(-\frac{\partial u}{\partial t} - \frac{\sigma^2}{2} s^2 \frac{\partial^2 u}{\partial s^2} \right) \Delta t, \quad (2.11)$$

where the Brownian motion dW has vanished. Taking the riskless interest rate r into account, we have

$$\frac{\Delta D}{\Delta t} = rD. \quad (2.12)$$

Inserting (2.11) and $D = -u + \frac{\partial u}{\partial s} s$ into this equation and transforming the discretized operators into continuous ones we conclude with the famous Black-Scholes PDE:

$$\frac{\partial u}{\partial t} + \frac{\sigma^2}{2} s^2 \frac{\partial^2 u}{\partial s^2} + r s \frac{\partial u}{\partial s} - r u = 0. \quad (2.13)$$

The model is completed by boundary conditions

$$u(s = 0, t) = 0 \quad \forall t \quad u(s, t) \rightarrow s \text{ for } s \rightarrow \infty \quad (2.14)$$

and terminal condition

$$u(s, T) = \max\{s - K, 0\}. \quad (2.15)$$

In the case of an European option (an option that can only be exercised *at* the maturity date) Black and Scholes derived an analytical solution to the PDE. This can be done by transforming the equation into a heat diffusion equation.

Lemma 2 *The respective explicit formulas for the value of a call $C(t, s)$ and put $P(t, s)$ are*

$$C(t, s) = sN(d_1) - K \exp(-r(T - t))N(d_2) \quad (2.16)$$

$$P(t, s) = K \exp(-r(T - t))N(-d_2) - SN(-d_1), \quad (2.17)$$

where

$$d_1 = \frac{\ln s/K + (r + \frac{\sigma^2}{2})(T - t)}{\sigma\sqrt{T - t}} \quad (2.18)$$

$$d_2 = d_1 - \sigma\sqrt{T - t} \quad (2.19)$$

with the cumulative standard distribution function

$$N(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz. \quad (2.20)$$

This means that for every pair (K, T) there is exactly one fair option price u .

2.2 On implied volatility

The most delicate parameter in the Black-Scholes model is the volatility coefficient σ . In financial terminology volatility is defined as the standard deviation of the instrument's yearly logarithmic returns. To understand the concept of logarithmic returns let us first define the simple annual return or simple annual rate of return of an investment with value V_0 at the start of the year and V_1 at the end of the year as:

$$r = \frac{V_1 - V_0}{V_0}. \quad (2.21)$$

Furthermore, the relation between the return R over a period of time t can be converted to an annualized rate or return r by the following relationship:

$$1 + R = (1 + r)^t. \quad (2.22)$$

The continuously compounded return or logarithmic return is then given by

$$R_t = \ln \frac{V_1}{V_0}, \quad (2.23)$$

with the logarithmic rate of return over a period t given by

$$r_t = \frac{R_t}{t}. \quad (2.24)$$

One of the reasons for the usage of logarithmic returns in financial modeling is related to the definition of the cumulative returns r_i ($i \in \{0, 1, 2, \dots, k\}$) over $k + 1$ periods of time i.e.

$$R_c = (1 + r_0) \cdot (1 + r_1) \dots (1 + r_k) - 1, \quad (2.25)$$

which can be represented as a simple sum if the returns r_i are logarithmic:

$$R_c = \sum_{i=0}^k r_i. \quad (2.26)$$

Therefore, volatility is a measure for the magnitude of uncertainty regarding an instrument's (logarithmic) return. One distinguishes between two types of volatility namely actual and implied volatility. Actual volatility is the volatility inferred from historical data with the most recent data being today's data. One also speaks of actual historical volatility if the last available data dates from earlier than today and actual future volatility for the volatility of an instrument ending at a point of time in the future. On the other hand there is an implied volatility analogon for each of these types. Implied volatility is the volatility for which a financial model (such as the Black-Scholes model) returns the actual observed market price. The volatility is therefore "implied" by the model.

Recall that the volatility coefficient in BSM is a constant. In reality one observes a different scenario though: depending on which market is considered different strike prices will result in different volatility values. Usually options with very low or very high strike prices will imply higher volatilities than *near the money* options. This phenomenon is called the smile effect. Additionally, volatility does not seem to only depend on strike price but also on term structure i.e. maturity dates. Taking the dependences on strike K and maturity T into account one will end up with an implied volatility surface serving as a representation of the mapping

$$\sigma(K, T) : \mathbb{R}^2 \rightarrow \mathbb{R}. \quad (2.27)$$

A vast body of research is available on how to retrieve this volatility information from e.g. the (modified) Black-Scholes model. Some of the methods will be presented, others will be mentioned and references will be given in Chapter 4.

2.3 Model shortcomings and criticism

Haug & Taleb [20] also criticize that Black & Scholes use an academic and theoretical argument to derive a formula that has been in practitioner's hands for a long time. More specifically they criticize that the model is based on dynamic hedging meaning buying and selling fractional amounts of securities at arbitrary times. In reality this is not possible due to technical restrictions. Furthermore they argue that traders do not do valuation based on unknown probability distributions but price options such that they are compatible with put-call parity and other instruments in the market. As such the degrees of freedom are heavily restricted and pricing models like Black-Scholes are unnecessary. Further points of critique concern some of the main assumptions of the model. For example, the fact that the increments of the price of the underlying are modeled as normally distributed does not make the model robust to large unexpected changes in the market regime. Ultimately, that is what caused the decline of Scholes' and Robert C. Merton's LTCM hedge fund which relied on dynamic hedging strategies. Losses totaled \$4.6 billion within less than 4 months in 1998 and resulted in a bailout initiated by the United States Federal Reserve, see Edwards [15]. Research also exists on modifying the Black-Scholes framework to accommodate for and study the effects of transaction costs Amster [3], dividend payments and on incorporating non-constant volatility.

Since we are not interested in pricing options directly but in obtaining a trend into which direction an option might be trading in the future we think that it is still important to investigate whether Black-Scholes' model can be used for such purposes.

Chapter 3

Time Inverse Parabolic Problem: Option Price Forecasting

The Black-Scholes model is a way of deriving a fair price of an option today given a future expiration date T . The main idea of Klivanov [29] is to reverse the time t and calculate i.e. forecast a future option price given today's market data. First, we show why this inverse problem is ill-posed in the sense of Hadamard using the analogy to the heat equation as an accessible example. Next, the full model including new boundary and initial conditions is presented. Thereafter, we present the most common numerical solution approaches and reproduce convergence results which are based on Carleman estimates for one such class of procedures from Klivanov [29, 27]. The regularization technique treating this ill-posed problem is the famous Tikhonov regularization approach for the solution of ill-posed inverse problems. Some remarks on the role of data errors and the inevitability of having estimates on the input data error for the successful construction of regularization algorithms are made. Finally, some numerical experiments are carried out to determine optimal regularization parameter, influence of different boundary criteria and preferable numerical solution techniques.

3.1 In analogy to the heat equation

Recalling that the Black-Scholes is of the parabolic type let us consider the heat equation for a moment. In its simplest form the heat equation can be written as

$$u_t - u_{xx} = 0 \quad \text{in } \Omega \subset \mathbb{R} \times (0, \infty), \quad (3.1)$$

$$u(x, 0) = f(x) \quad \text{for } x \in \Omega, \quad (3.2)$$

$$u(\cdot, t) = 0 \quad \text{on } \partial\Omega. \quad (3.3)$$

For simplicity, let us consider the one-dimensional case and let the domain be bounded such that $x \in \Omega = (0, \pi)$ and let the terminal time $T > 0$ be chosen such that $t \in (0, T) \subset (0, \infty)$. It is well known that there exists a unique solution to this initial boundary value problem which can be represented by Fourier series.

Lemma 3 *The solution of the one-dimensional heat equation (3.1) with initial and boundary conditions (3.2)-(3.3) is given by the Fourier series*

$$u(x, t) = \sum_{k=1}^{\infty} f_k \sin(kx) \exp(-k^2 t), \quad (3.4)$$

with the Fourier coefficients f_k , $k = 1, 2, \dots, \infty$ given by

$$f_k = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(kx) dx. \quad (3.5)$$

Now, consider the time inverse problem

$$u_t + u_{xx} = 0 \quad \text{in } \Omega \subset \mathbb{R} \times (0, \infty), \quad (3.6)$$

$$u(x, 0) = f(x) \quad \text{for } x \in \Omega, \quad (3.7)$$

$$u(\cdot, t) = 0 \quad \text{on } \partial\Omega. \quad (3.8)$$

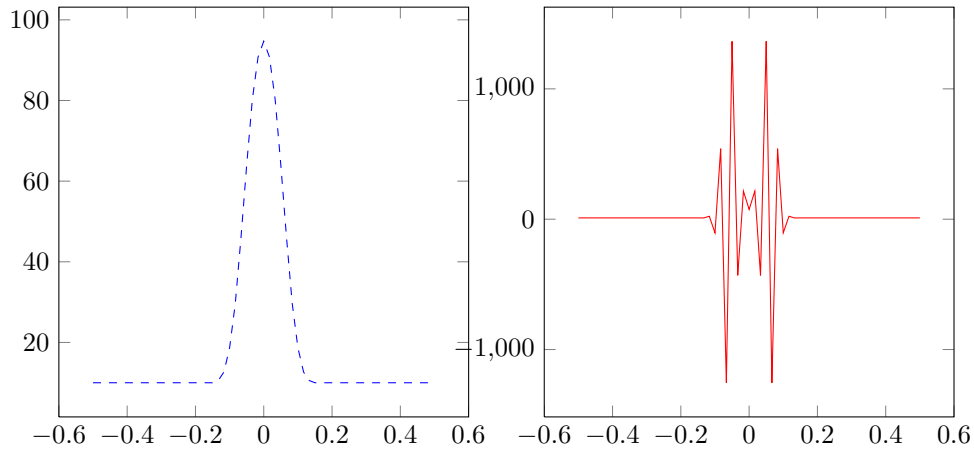


Figure 3.1: Heat equation solved forward --- and backward —.

Again, consider the one-dimensional case with the bounded domain $\Omega = (0, \pi)$ and $t \in (0, T) \subset (0, \infty)$.

Lemma 4 *The solution of the time inverse heat equation (3.6) with initial and boundary conditions (3.7)-(3.8) restricted to one space dimension is given by*

$$u(x, t) = \sum_{k=1}^{\infty} f_k \sin(kx) \exp(k^2 t). \quad (3.9)$$

This representation is instructive in understanding why the inverse time problem is ill-posed. Observe that for the solution (3.9) to exist the squares of the Fourier coefficients f_k with respect to k have to decay exponentially. Another typical problem in the context of inverse problems is the instability of the solution to small fluctuations in e.g. initial data. Consider the norm of the $u(x, t)$

$$\|u(x, t)\|_{L_2}^2 = \sum_{k=1}^{\infty} f_k^2 \exp(2k^2 t) \quad (3.10)$$

and remember that the Fourier coefficients are obtained via the convolution with the initial data i.e. $f(x)$ in our case. This means that small deviations in f_k can result in large variations of the solution $u(t, x)$. From (3.10) we can also deduce that the problem becomes more unstable as time t proceeds. To confirm these propositions and to get an impression for what this means in terms of numerical aspects we solve the heat equation forwards and backwards using a naive explicit solution scheme and plot the results in Fig. 3.1. For the purpose of demonstration we let $\Omega = \{(x, t) | x \in (-0.5, 0.5), t \in (0, 1)\}$ and let

$$f(x) = \begin{cases} 100 & |x| < \frac{1}{10} \\ 10 & \text{else.} \end{cases} \quad (3.11)$$

The result after $t = 10$ is plotted in Fig. 3.1 on the left. Applying the same scheme with time reversed leads to oscillations and a non realistic result as seen in Fig. 3.1 on the right after only 10 time steps. The discretization parameters dt, dx used fulfill the CFL condition: $\frac{dt}{dx^2} = \frac{10^{-4}}{(\frac{1}{60})^2} = 0.36 < \frac{1}{2}$. All numerical examples to appear from hereon have been executed with Matlab R2015a.

Thus this problem belongs to a special class of inverse problems, namely ill-posed inverse problems. According to Jacques Salomon Hadamard [18] a problem is *well-posed* if

1. A solution exists

2. The solution is unique
3. The solutions is depends continuously on the initial data.

A problem is *ill-posed* if any of these three stipulations does not hold.

The time reverse heat equation is a well-known and classic example for an inverse ill-posed Cauchy problem. For an overview of methods applied to tackle such problems see e.g. Seidman [43]. Useful information on data errors, error estimation and the impossibility to construct solution algorithms without any knowledge on the error level of the input data can be extracted from Yagola [48]. Alternative techniques as in e.g. Marie [34] try to solve well-posed approximations to the ill-posed problem rather than to solve the ill-posed problem.

3.2 Option price forecasting: Time inverse BSM

*And now remains that we find
out the cause of this effect.*

William Shakespeare, Hamlet

We use the model presented in Klibanov [29] which entails the Black-Scholes equation along with special initial and boundary conditions as follows.

$$Lu := \frac{\partial u}{\partial t} + \frac{\sigma(t)^2}{2} s^2 \frac{\partial^2 u}{\partial s^2} = 0 \text{ on } Q_T, \quad (3.12)$$

$$u(s, 0) = f(s) \text{ for } s \in (s_a, s_b), \quad (3.13)$$

$$u(s_b, t) = u_b(t), u(s_a, t) = u_a(t) \text{ for } t \in (0, T), \quad (3.14)$$

where the domain is given as

$$Q_T = \{(s, t) | s \in (s_b, s_a), t \in (0, T)\} \quad (3.15)$$

and with

$$f(s) = \frac{u_b - u_a}{s_b - s_a} s + \frac{u_a s_b - u_b s_a}{s_b - s_a}. \quad (3.16)$$

Note that $f(s)$ is simply a linear interpolation between the option's bid price u_b and ask price u_a . In the general context it is not necessary for f to be linear though. As we will see later forecasting is done for a rather short time horizon. This is why the boundary conditions in (3.14) namely $u_b(t)$ and $u_a(t)$ are the extension of the quadratic interpolation polynomial of preceding bid and ask prices in the time interval $(-T, 0)$. The same goes for the volatility coefficients $\sigma(t)$, which is the extrapolation of the quadratic interpolation polynomial for the volatility in $(-T, 0)$. Let the Sobolev space $H^k(Q_T)$ be defined as follows.

Definition Let $1 \leq p < \infty$ and $k \in \mathbb{Z}_+$. Suppose $u \in L_p(Q_T)$ and the weak derivatives $\partial^\alpha u$ exist for any multi-index α with $|\alpha| \leq k$ such that

$$\partial^\alpha u \in L_p(Q_T), \quad |\alpha| \leq k. \quad (3.17)$$

We then say that $u \in W^{k,p}$ and $W^{k,p}$ is called Sobolev space. For $p = 2$, we define $H^k = W^{k,2}$.

For our case let $H^{2,1}$ be similarly defined as the space of functions $u \in L_2(Q_T)$ with $\partial_s u, \partial_{ss} u, \partial_t u \in L_2(Q_T)$.

Deviations from the function

$$F(s, t) = \frac{u_b(t) - u_a(t)}{s_b - s_a} s + \frac{u_a(t)s_b - u_b(t)s_a}{s_b - s_a} \in H^2(Q_T) \quad (3.18)$$

will be penalized and these relations hold

$$F(s, 0) = f(s), \quad F(s_b, t) = u_b(t), \quad F(s_a, t) = u_a(t). \quad (3.19)$$

The fact that deviations of the solution u from the function F are penalized is already a regularization technique. In the spirit of Tikhonov regularization a solution that solves the original problem as well as possible but at the same time fulfills certain restrictions is searched for. Remember for example the noncontinuous dependence of the norm of $u(x, t)$ in (3.10) on the Fourier coefficients. In that case one could for example ask for an approximation to $u(x, t)$ with bounded norm. Something similar is done here: we want that equation (3.12) is fulfilled in the integral sense while at the same time we demand that u is close to an interpolation of option prices F . This gives rise to the Tikhonov functional

$$J_\lambda(u) = \int_{s_b}^{s_a} \int_0^T (Lu)^2 dt dx + \lambda \|u - F\|_{H^2(Q_T)}, \quad (3.20)$$

where the regularization parameter $\lambda \in \mathbb{R}$ determines how much deviations of u from F are penalized as well as how smooth the solution is.

In the following section the discretization of this model, the choice of the regularization parameter λ and alternative regularization techniques are discussed. Furthermore, the convergence proof for this Tikhonov functional is replicated from Klibanov [29] and some remarks on the iterative solver for the resulting system of linear equations are made. Another important problem is the way in which the model error can be estimated and the question how this error level relates to the choice of the regularization parameter λ . The chapter is concluded with a handful of numerical experiments in which the need for higher order Tikhonov regularization is debated and concrete recommendations for the mesh size and regularization parameter λ for this problem are made.

3.3 Solution approaches

There are several techniques that are regularly applied to these types of inverse problems. We will have a closer look at a method called Tikhonov regularization which will be used for this problem. However, the general problem is sketched first and thereafter we define what a solution of an ill-posed problem is in connection with perturbations in the input data.

3.3.1 Error estimation

Since we ultimately want to make statements on the convergence properties of regularized solutions of ill-posed problems to the exact solution of the problem, an interesting question to address is whether or not it is possible to estimate errors and if so, what a priori information is necessary to do so? A precise and general survey on data errors and error estimation is given by Yagola [48]. In general, a "stable regularization method does not guarantee the possibility of an error estimation or a comparison of convergence rates" [48]. Also, it is essential to include all a priori information in the mathematical problem formulation before trying to solve it. Let us start by answering the question whether it is possible to solve an ill-posed problem without knowledge of the data error? Consider the operator equation

$$Ax = \bar{u}, \quad x \in X, \bar{u} \in U. \quad (3.21)$$

Here the unknown is x , X, U are Hilbert spaces and the operator $A : X \rightarrow U$ is bounded and linear. Recalling Hadamard's requirements for well-posed problems we assert that stability means that if instead of the exact data (A, \bar{u}) we are given admissible data (A_h, u_δ) with $\|A - A_h\| \leq h$, $\|\bar{u} - u_\delta\| \leq \delta$, then the approximate solution converges to the exact one as h and δ approach zero. The problems of existence and uniqueness of the problem are usually dealt with by taking a generalized solution of the problem. In the case of non-uniqueness one usually takes the solution which has the smaller norm or smaller distance from a fixed element that arises from a priori knowledge. The main problem, however, is that the search for such a generalized solution can still

be unstable with respect to errors of A and \bar{u} . This brings us to the definition of solution of an ill-posed problem. Tikhonov [45], in his seminal paper on incorrectly formulated problems and his regularization method, states that to solve an ill-posed problem there has to be a mapping $R(h, \delta, A_h, u_\delta)$ which

- maps a regularized solution $x_{h,\delta} = R(h, \delta, A_h, u_\delta)$ to the data of the problem with $h \geq 0$, $\delta \geq 0$, A_h bounded and linear map, $u_\delta \in U$
- and guarantees convergence of $x_{h,\delta} \rightarrow x$ as $h, \delta \rightarrow 0$.

Note that not all ill-posed problems allow for the construction of such a regularizing algorithm and are therefore termed "nonregularizable ill-posed problems". One interesting result is the fact that it is impossible to construct stable regularizing algorithms without the knowledge of the error level (h, δ) ; see Yagola [48] or Bakushinskii [7] for a proof.

3.3.2 Tikhonov regularization

Apart from giving this definition of the solution to an ill-posed problem Tikhonov also proposed a regularization algorithm. The method searches for a minimum $x_\lambda = x_\lambda(h, \delta) \in X$ of the Tikhonov functional

$$T_\lambda(x) = \|A_h x - u_\delta\|^2 + \lambda \|x\|^2, \quad (3.22)$$

selecting the regularization parameter λ such that convergence of the approximate solution $x_{h,\delta}$ to \bar{x} is ensured. We do not give any more general information on Tikhonov functionals here because convergence properties of the Tikhonov-like functional for the case of our inverse problem (3.20) will be proved in Section 3.4. For more details on the general case see e.g. Tikhonov [45, 46].

3.4 Convergence results: Carleman estimates

He was a genius! My older friends in Uppsala used to tell me about the wonderful years they had had when Carleman was there. He was the most active speaker in the Uppsala Mathematical Society and a well-trained gymnast. When people left the seminar crossing the Fyris River, he walked on his hands on the railing of the bridge. Bo Kjellberg, 1995

To prove that the Tikhonov approach guarantees the existence of at most one solution to the inverse problem and to establish convergence rates for the numerical minimization of the Tikhonov functional Carleman estimates play a key role. Specifically, the Carleman estimate allows one to obtain a logarithmic stability estimate (Theorem 1) which in turn guarantees uniqueness of the solution to the parabolic problem and is needed to establish convergence rates of the Tikhonov minimizer to a solution of this problem. The following theorems from Klibanov [27] are presented for a general parabolic problem of the form

$$u_t + L_{par} u = f \text{ in } Q_T = \Omega \times (0, T), u \in H^2(Q_T), \quad (3.23)$$

$$u(s, 0) = g(s), \quad (3.24)$$

$$u(s, t) = p(s, t) \text{ on } \partial\Omega \times (0, T) \quad (3.25)$$

with

$$L_{par} u = \sum_{i,j=1}^n a^{ij}(s, t) u_{ij} + \sum_{j=1}^n b^j(s, t) u_j + b_0(s, t) u, \quad (3.26)$$

$$L_{0,par} u = \sum_{i,j=1}^n a^{ij}(s, t) u_{ij}, \quad A = \partial_t + L_{par}, \quad (3.27)$$

where $u_j = \frac{\partial u}{\partial s_j}$, $a^{ij} \in C^1(\bar{Q}_T)$, $b_j, b_0 \in B(\bar{Q}_T)$. Here $B(\bar{Q}_T)$ is the space of bounded functions and $L_{0,par}$ is the principal part of the parabolic operator for which the following ellipticity condition holds: there exist two constants $\mu_2 \geq \mu_1 > 0$ such that

$$\mu_1 |\eta|^2 \leq \sum_{i,j=1}^n a_{ij}(s,t) \eta_i \eta_j \leq \mu_2 |\eta|^2 \quad \forall s \in \bar{Q}_T, \forall \eta \in \mathbb{R}^n. \quad (3.28)$$

Even though it sounds more natural to assume $u \in H^{2,1}(Q_T)$ in (3.23), the extra smoothness is needed to prove stability and convergence results later on and is specifically mentioned in e.g. (3.37). This way problem (3.12) becomes a special case in which $f = 0$, $g(s) = f(s)$ with $f(s)$ as defined in (3.13) and $p(s,t) = u_b(t)$ respectively $p(s,t) = u_a(t)$ as in (3.14). Let the lower order terms of (3.26) be bounded by a constant $M > 0$:

$$\|b_j\|_{C(\bar{Q}_T)} \leq M, \quad j = 0, \dots, n. \quad (3.29)$$

3.4.1 Carleman estimate

Since the following estimate is the key to proving convergence of the Tikhonov functional to a unique solution of (3.23)-(3.25) we introduce the concept of Carleman estimates. For an introductory overview of the application of Carleman estimates for coefficient inverse problems see Timonov [31] and Klibanov [27] for the application to ill-posed Cauchy problems. Let the constant $k > 0$ be a number which will be chosen later. For $\lambda > 1$ the Carleman Weight Function $\varphi_\lambda(t)$ in our case is defined as

$$\varphi_\lambda(t) = (k+t)^{-\lambda}, \quad t > 0. \quad (3.30)$$

The combination of the following two lemmata which originate from Klibanov [30] will yield the Carleman estimate which will be used in the following proof.

Lemma 5 *There exists a number $\lambda_0 > 1$ and constant $C > 0$ both depending on μ_1, μ_2 and $\max_{i,j} \|a_{ij}\|_{C^1(\bar{Q}_T)}$ such that for all $\lambda \geq \lambda_0$ and for all $u \in C^{2,1}(\bar{Q}_T)$ the following estimate holds for all $(s,t) \in Q_T$*

$$(-u_t - L_{0,par}u)u\varphi_\lambda^2 \geq \mu_1 |\nabla u|^2 \varphi_\lambda^2 - \lambda u^2 \varphi_\lambda^2 + \operatorname{div} U_1 + \frac{\partial}{\partial t} \left(-\frac{u^2}{2} \varphi_\lambda^2 \right), \quad (3.31)$$

$$|U_1| \leq C|u| |\nabla u| \varphi_\lambda^2. \quad (3.32)$$

Lemma 6 *Additionally for λ_0, C from Lemma 5, for all $\lambda \geq \lambda_0$ and for all $u \in C^{2,1}(\bar{Q}_T)$ the following estimate holds for all $(s,t) \in Q_T$*

$$(u_t + L_{0,par}u)^2 \varphi_\lambda^2 \geq -C |\nabla u|^2 \varphi_\lambda^2 + \lambda (k+t)^{-2} u^2 \varphi_\lambda^2 + \quad (3.33)$$

$$\frac{\partial}{\partial t} \left(\lambda (k+t)^{-1} u^2 \varphi_\lambda^2 - \varphi_\lambda^2 \sum_{i,j=1}^n a_{ij} u_i u_j \right) + \operatorname{div} U_2,$$

$$|U_2| \leq C|u_t| |\nabla u| \varphi_\lambda^2. \quad (3.34)$$

The combination of the preceding two lemmata yields the Carleman estimate we are looking for.

Lemma 7 (Carleman estimate) *There exists a number $a_0 \in (0,1)$ depending on μ_1, μ_2 and $\max_{i,j} \|a_{ij}\|_{C^1(\bar{Q}_T)}$, M such that if $k+T < a_0$, there exists a number λ_1 depending on μ_1, μ_2 , constants $C, C_1 > 0$ and θ depending on μ_1, μ_2 and $\max_{i,j} \|a_{ij}\|_{C^1(\bar{Q}_T)}$, M such that for all $\lambda > \lambda_1$*

and for all $u \in C^{2,1}(\bar{Q}_T)$ the following estimate holds for all $(s, t) \in Q_T$

$$(u_t + L_{0,par}u)^2 \varphi_\lambda^2 \geq \frac{2}{3} \theta \mu_1 |\nabla u|^2 \varphi_\lambda^2 + C \lambda u^2 \varphi_\lambda^2 + \operatorname{div} U \quad (3.35)$$

$$+ \frac{\partial}{\partial t} \left(C \lambda (k+t)^{-1} (1 - C_1(k+t)) u^2 \varphi_\lambda^2 - C \varphi_\lambda^2 \sum_{i,j=1}^n a_{ij} u_i u_j \right),$$

$$|U| \leq C(|u_t| + |u|) |\nabla u| \varphi_\lambda^2. \quad (3.36)$$

Proof The proof for this lemma is only outlined here.

3.4.2 Stability estimates and Regularization

Consider the family of functions u_δ with $\delta \in (0, 1)$ satisfying

$$\int_{Q_T} (\partial_t u_\delta + L_{par} u_\delta)^2 ds dt \leq N \delta^2, \quad u_\delta \in H^2(Q_T), \quad (3.37)$$

$$\|u_\delta(\cdot, 0)\|_{L_2(\Omega)} \leq \sqrt{N} \delta, \quad (3.38)$$

$$u_\delta = 0, \quad \text{on } \partial\Omega \times (0, T), \quad (3.39)$$

where $N > 0$ is a constant independent of δ .

Theorem 1 *Let u_δ satisfy conditions (3.37)-(3.39). Let T fulfill $T > \frac{a_0}{2}$ with a_0 from Lemma 7. Then there exists a number $\delta_0 = \delta_0(L_{par}) \in (0, 1)$ and a number $C_{10} > 0$ such that the logarithmic stability estimate holds for all $u_\delta(s, T)$*

$$\|u_\delta(\cdot, T)\|_{L_2(\Omega)} \leq \frac{C_{10}}{\sqrt{\ln(\delta^{-1})}} (1 + \|u_\delta\|_{H^2(Q_T)}). \quad (3.40)$$

For every $\varepsilon \in (0, T/2)$ let $\beta = \beta(\varepsilon)$ be defined as

$$\beta = \beta(\varepsilon) = -\frac{\ln(1 - \varepsilon/T)}{2 \ln(1 - \varepsilon/(2T))} \in (0, \frac{1}{2}). \quad (3.41)$$

Then also this Hölder estimate holds

$$\|\nabla u_\delta\|_{L_2(Q_{T-\varepsilon})} + \|u_\delta\|_{L_2(Q_{T-\varepsilon})} \leq C_{10} (1 + \|u_\delta\|_{H^2(Q_T)}) \delta^\beta. \quad (3.42)$$

Theorem 2 *Uniqueness of the solution to problem (3.23)-(3.25) is a direct consequence of Theorem 1.*

With these estimates at hand the final step is to prove the existence of a solution to the minimization problem arising with the Tikhonov regularization approach and to find an estimate for the convergence rate in such a case. Assume that there exists a function $F \in H^2(Q_T)$ satisfying

$$F(s, 0) = g(s), \quad F = p(s, t) \quad \text{on } \partial\Omega \times (0, T). \quad (3.43)$$

The Tikhonov functional corresponding to the (inverse) problem (3.23) is

$$J_{\lambda, L_{par}}(u) = \|u_t + L_{par}u\|_{L_2(Q_T)}^2 + \lambda \|u - F\|_{H^2(Q_T)}^2, \quad u \in H^2(Q_T) \quad (3.44)$$

$$\text{subject to initial and boundary conditions (3.24) - (3.25)} \quad (3.45)$$

Since we are dealing with convergence analysis assume that there exists an exact solution $u^* \in H^2(Q_T)$ of (3.23)-(3.25).

$$\|g - g^*\|_{L_2(\Omega)} \leq \delta, \quad \|f - f^*\|_{L_2(Q_T)}, \quad \|F - F^*\|_{H^2(Q_T)} \leq \delta \quad (3.46)$$

The number δ can be interpreted as the error level in the input data. In the case of the forecasting problem the intervals s and t are defined on are small and it is assumed that approximating the exact F^* linearly in s and t by $F(s, t)$ as defined in (3.19) results in a statement of the form (3.46) with $\delta > 0$ sufficiently small.

Theorem 3 (Existence) *If there exists a function $F \in H^2(Q_T)$ satisfying (3.43), then for every $\lambda > 0$ there exists a unique minimizer $u_\lambda \in H^2(Q_T)$ of the functional (3.44) and the following estimate holds*

$$\|u_\lambda\|_{H^2(Q_T)} \leq \frac{C_{11}}{\sqrt{\lambda}} (\|F\|_{H^2(Q_T)} + \|f\|_{L_2(Q_T)}). \quad (3.47)$$

Proof Define $H_0^2(Q_T) = \{v \in H^2(Q_T) : v|_\Gamma = \partial_n v = 0\}$. Letting $v = u - F$, one can argue that the Tikhonov functional is a bounded linear functional on the space $H_0^2(Q_T)$. The aim is to minimize the following functional subject to boundary conditions (3.24) and (3.25)

$$\hat{J}_\lambda(v) = \|Av + (AF - f)\|_{L_2(Q_T)}^2 + \lambda \|v\|_{H^2(Q_T)}, \quad v \in H_0^2(Q_T). \quad (3.48)$$

Assume the minimizer is given by $v_\lambda = u_\lambda - F \in H_0^2(Q_T)$ and note that due to the variational principle it should satisfy

$$(Av_\lambda, Ah) + \lambda[v_\lambda, h] = (Ah, f - AF), \quad \forall h \in H_0^2(Q_T), \quad (3.49)$$

where (\cdot, \cdot) denotes the scalar product in $L_2(Q_T)$ and $[\cdot, \cdot]$ denotes the scalar product in $H^2(Q_T)$. Define

$$\{v, h\}_\lambda = (Av, Ah) + \lambda[v, h], \quad \forall v, h \in H_0^2(Q_T). \quad (3.50)$$

The expression $\{v, h\}$ is a new scalar product on $H_0^2(Q_T)$ and its corresponding norm fulfills

$$\sqrt{\lambda} \|v\|_{H^2(Q_T)} \leq \{v\}_\lambda \leq C_{11} \|v\|_{H^2(Q_T)}. \quad (3.51)$$

The norms $\{v\}_{H^2(Q_T)}$ and $\|v\|_{H^2(Q_T)}$ are equivalent and we can rewrite

$$\{v_\lambda, h\}_\lambda = (Ah, f - AF), \quad \forall h \in H_0^2(Q_T). \quad (3.52)$$

This yields

$$|(Ah, f - AF)| \leq C(\|F\|_{H^2(Q_T)} + \|f\|_{L_2(Q_T)}) \{h\}_\lambda. \quad (3.53)$$

This shows that the right hand side of (3.52) is a bounded linear functional on $H_0^2(Q_T)$. Riesz' theorem then implies a unique solution of the minimization problem. The Carleman estimate is not used in this proof.

Theorem 4 (Convergence rate) *Assume (3.46) holds and let $\lambda = \lambda(\delta) = \delta^{2\alpha}$ with $\alpha \in (0, 1]$. Also let $T < a_0/2$ with $a_0 \in (0, 1)$, let ε, β and let the Carleman estimate be defined as before. Then there exists a sufficiently small number $\delta_0 = \delta_0(L_{par}, a_0, Q_T) \in (0, 1)$ and constant $C_{12} > 0$ such that for all $\delta \in (0, \delta_0^{1/\alpha})$ the following convergence rates hold true*

$$\|u_{\lambda(\delta)}(\cdot, T) - u^*(\cdot, T)\|_{L_2(\Omega)} \leq \frac{C_{12}}{\sqrt{\ln(\delta^{-1})}} (1 + \|u^*\|_{H^2(Q_T)}), \quad (3.54)$$

$$\|\nabla u_{\lambda(\delta)} - \nabla u^*\|_{L_2(Q_{T-\varepsilon})} + \|u_{\lambda(\delta)} - u^*\|_{L_2(Q_{T-\varepsilon})} \leq C_{12} (1 + \|u^*\|_{H^2(Q_T)}) \delta^{\alpha\beta}. \quad (3.55)$$

Note that the stronger Hölder estimate (3.55) only holds for times close to T whereas the logarithmic estimate (3.54) is valid on the whole of Q_T .

Proof Let $v^* = u^* - F^*$, so $v^* \in H_0^2(Q_T)$ and $Av^* = f^* - AF^*$. Consider

$$(Av^*, Ah) + \gamma[v^*, h] = (Ah, f^* - AF^*) + \lambda[v^*, h], \quad \forall h \in H_0^2(Q_T). \quad (3.56)$$

Subtraction of 3.49 and the definitions $\tilde{v}_\lambda = v^* - v_\lambda$, $\tilde{f} = f^* - f$, $\tilde{F} = F^* - F$ yield

$$(A\tilde{v}_\lambda, Ah) + \lambda[\tilde{v}_\lambda, h] = (Ah, \tilde{f} - A\tilde{F}) + \lambda[v^*, h], \forall h \in H_0^2(Q_T). \quad (3.57)$$

The test function is set to $h \equiv \tilde{v}_\lambda$ and thus

$$\|A\tilde{v}_\lambda\|_{L_2(Q_T)}^2 + \lambda\|\tilde{v}_\lambda\|_{H^2(Q_T)} = (A\tilde{v}_\lambda, \tilde{f} - A\tilde{F}) + \lambda[v^*, \tilde{v}_\lambda]. \quad (3.58)$$

Application of Cauchy-Schwarz inequality gives

$$\|A\tilde{v}_\lambda\|_{L_2(Q_T)}^2 + \lambda\|\tilde{v}_\lambda\|_{H^2(Q_T)}^2 \leq \frac{1}{2}\|A\tilde{v}_\lambda\|_{L_2(Q_T)}^2 + \frac{1}{2}\|\tilde{f} - A\tilde{F}\|_{L_2(Q_T)}^2 + \frac{\lambda}{2}(\|v^*\|_{H^2(Q_T)} + \|\tilde{v}_\lambda\|_{H^2(Q_T)}). \quad (3.59)$$

Using the a priori information (3.46) this means

$$\|A\tilde{v}_\lambda\|_{L_2(Q_T)}^2 + \lambda\|\tilde{v}_\lambda\|_{H^2(Q_T)}^2 \leq C_{12}\delta^2 + \lambda\|v^*\|_{H^2(Q_T)}^2. \quad (3.60)$$

Noting that $\lambda = \delta^{2\alpha}$ with $\alpha \in (0, 1]$, we deduce $\delta^2 \leq \lambda$ and this is why (3.60) implies

$$\|\tilde{v}_\lambda\|_{H^2(Q_T)} \leq C_{12}(1 + \|v^*\|_{H^2(Q_T)}) \quad (3.61)$$

$$\|A\tilde{v}_\lambda\|_{H^2(Q_T)}^2 \leq C_{12}(1 + \|v^*\|_{H^2(Q_T)}^2)\delta^{2\alpha} \quad (3.62)$$

$$(3.63)$$

For convenience denote $w_\lambda = \tilde{v}_\lambda(1 + \|v^*\|_{H^2(Q_T)})$. Combining (3.61) and the Hölder estimate from Theorem 1 imply $\|w_\lambda\|_{H^2(Q_T)} \leq C_{12}\delta^{\alpha\beta}$ and

$$\tilde{v}_\lambda\|_{H^1(Q_{T+\varepsilon})} \leq C_{12}(1 + \|v^*\|_{H^2(Q_T)})\delta^{\alpha\beta}, \quad \forall \delta \in (0, \delta_0). \quad (3.64)$$

Applying the triangle inequality onto $\tilde{v}_\lambda = (u^* - u) + (F^* - F)$ and using (3.46) once again, yields

$$\|\tilde{v}_\lambda\|_{H^1(Q_{T+\varepsilon})} \geq \|u_\lambda - u^*\|_{H^1(Q_{T+\varepsilon})} - \|F^* - F\|_{H^1(Q_{T+\varepsilon})} \geq \|u_\lambda - u^*\|_{H^1(Q_{T+\varepsilon})} - \delta \quad (3.65)$$

Putting the last two inequalities together and arguing that $\delta^{\alpha\beta} > \delta$ since $\beta, \delta \in (0, 1)$ and $\alpha \in (0, 1]$ we get the final result presented in the theorem above.

3.4.3 Discretization

The domain is given by $Q_T = (s_b, s_a) \times (0, T)$ with $T = 2\tau$, where $\tau \in (0, \frac{1}{4})$ is time in years. The limitation on τ is due to the restriction on T in Theorem 1. Furthermore, $s_b < s_a$ always holds. In our case $\tau = 1/255$ represents one day. The domain is discretized in tuples of the discretizations of s and t . The stock prices $s \in (s_b, s_a)$ are discretized into M equidistant points $s_i = s_b + i \cdot ds$, $i = 0, \dots, M$. Respectively time $t \in (0, T)$ is split into N equidistant points such that $t_i = i \cdot dt$, $i = 0, \dots, N$. For all computations $N = M = 100$. Let $u_{n,m}$ denote the option price corresponding to a stock price s_m at time t_n , i.e. $u_{n,m} \approx u(s_m, t_n)$. The values $u_{n,m}$ are formatted into a vector of this form:

$$u = [u_{11}u_{12} \dots u_{1M-1}u_{1M}u_{21}u_{22} \dots u_{2M-1}u_{2M} \dots u_{N-11}u_{N-12} \dots u_{N-1M-1}u_{N-1M} \\ u_{N1}u_{N2} \dots u_{NM-1}u_{NM}]^T \in \mathbb{R}^{N \times M} \quad (3.66)$$

The discretized version of the Tikhonov functional (3.20) can then be written as:

$$\hat{J}_\alpha(u) = dt ds \sum_{j=M+1}^{NM-M} \left(\frac{u_{j+M} - u_{j-M}}{2dt} + \frac{\sigma(t(\lfloor \frac{j-1}{M} \rfloor + 1))^2}{2} s(j-1 \pmod{M} + 1)^2 \frac{u_{j+1} - 2u_j + u_{j-1}}{ds^2} \right)^2 + \\ \alpha ds dt \sum_{j=M+1}^{NM-M} (u_j - F_j)^2 + \left(\frac{1}{2dt} (u_{j+M} - u_{j-M} - F_{j+M} + F_{j-M}) \right)^2 + \\ \left(\frac{1}{2ds} (u_{j+1} - u_{j-1} - F_{j+1} + F_{j-1}) \right)^2 + \left(\frac{1}{dt^2} (u_{j+M} - 2u_j + u_{j-M} - F_{j+M} + 2F_j - F_{j-M}) \right)^2 + \\ \left(\frac{1}{ds^2} (u_{j+1} - 2u_j + u_{j-1} - F_{j+1} + 2F_j - F_{j-1}) \right)^2. \quad (3.67)$$

For the discretization of the Sobolev norm second order finite differences are used to approximate the derivatives. For the first derivatives a central differencing scheme is used such that the time derivative at the m -th price s_m is given by

$$\nabla_t u_{:,m} = Dt^* u_m + O(dt) \quad (3.68)$$

with

$$Dt^* = \frac{1}{dt} \begin{bmatrix} -1 & 1 & 0 & \dots & \dots & \dots & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & \ddots & & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \dots & \dots & \dots & 0 & -1 & 1 \end{bmatrix} \in \mathbb{R}^{N \times N}. \quad (3.69)$$

To assure correct approximations of the derivative on the boundaries, first order accurate forward differences are used on the boundaries. Consequently the full set of derivatives is given by

$$\nabla_t u_{n,m} = Dt u_{n,m} + O(dt) \quad (3.70)$$

with

$$Dt = Dt^* \otimes I^M \quad (3.71)$$

and is only first-order accurate in total.

Definition The Kronecker product \otimes is defined as follows. Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$, then the Kronecker product $A \otimes B \in \mathbb{R}^{mp \times nq}$ is the block matrix defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix}. \quad (3.72)$$

The procedure for the derivative with respect to s is analogous:

$$Ds^* = \frac{1}{ds} \begin{bmatrix} -1 & 1 & 0 & \dots & \dots & \dots & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & \ddots & & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \dots & \dots & \dots & 0 & -1 & 1 \end{bmatrix} \in \mathbb{R}^{M \times M} \quad (3.73)$$

$$Ds = I^N \otimes Ds^* \quad (3.74)$$

For the second derivatives the following representation is obtained.

$$\Delta_{tt} u_{:,m} = Dtt^* u_{:,m} + O(dt) \quad (3.75)$$

with

$$Dtt^* = \frac{1}{dt^2} \begin{bmatrix} \frac{1}{2} & -1 & \frac{1}{2} & 0 & \dots & \dots & 0 \\ 1 & -2 & 1 & 0 & & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 & -2 & 1 \\ 0 & \dots & \dots & 0 & \frac{1}{2} & -1 & \frac{1}{2} \end{bmatrix} \in \mathbb{R}^{N \times N}. \quad (3.76)$$

Again note that in order to obtain reasonable approximations of the derivative on the boundary, the first and last row of Dtt^* are modified when compared to standard Toeplitz finite difference matrix. The choice of the stencil $(\frac{1}{2}, -1, \frac{1}{2})$ is in accordance with a first order accurate approximation of the second derivative on the boundary. To see this, consider the following two Taylor series

$$u_2 = u_1 + dt \frac{\partial u_1}{\partial t} + dt^2 \frac{\partial^2 u_1}{\partial t^2} + O(dt^3) \quad (3.77)$$

$$u_3 = u_1 + 2dt \frac{\partial u_1}{\partial t} + 4dt^2 \frac{\partial^2 u_1}{\partial t^2} + O(dt^3) \quad (3.78)$$

Addition of (3.78)-2·(3.77) gives

$$\frac{\partial^2 u_1}{\partial t^2} = \frac{u_1 - 2u_2 + u_3}{2dt^2} + O(dt). \quad (3.79)$$

Consequently the full set of derivatives is given by

$$\Delta_{tt} u_{n,m} = Dtt u_{n,m} + O(dt) \quad (3.80)$$

with

$$Dtt = Dtt^* \otimes I^M \quad (3.81)$$

Similarly for the second derivative with respect to s one obtains:

$$\Delta_{ss} u_{n,:} = Dss^* u_{n,:} + O(ds) \quad (3.82)$$

with

$$Dss^* = \frac{1}{ds^2} \begin{bmatrix} \frac{1}{2} & -1 & \frac{1}{2} & 0 & \dots & \dots & 0 \\ 1 & -2 & 1 & 0 & & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 & -2 & 1 \\ 0 & \dots & \dots & 0 & \frac{1}{2} & -1 & \frac{1}{2} \end{bmatrix} \in \mathbb{R}^{M \times M}. \quad (3.83)$$

The derivation of the boundary conditions is completely analogous to the one for the Dtt^* matrix and the full set of derivatives is given by

$$\Delta_{ss} u_{n,m} = Dss u_{n,m} + O(ds) \quad (3.84)$$

with

$$Dss = Dss^* \otimes I^N. \quad (3.85)$$

Also, define the Sobolev matrix D

$$D = [I \ Dt \ Ds \ Dtt \ Dss]^T. \quad (3.86)$$

The resulting approximation \hat{J}_α of the Tikhonov functional J_α is then given by

$$\hat{J}_\alpha = dsdt \|Au\|_2^2 + \alpha dsdt \left(\|u - F\|_2^2 + \|D_t(u - F)\|_2^2 + \|D_s(u - F)\|_2^2 + \|D_{tt}(u - F)\|_2^2 + \|D_{ss}(u - F)\|_2^2 \right), \quad (3.87)$$

where $A = Dt + RDss$ with $R \in \mathbb{R}^{MN \times MN}$ being a diagonal matrix with elements corresponding to the factor $\frac{\sigma(t)^2}{2} s^2$.

Incorporation of initial and boundary conditions

The boundary conditions (3.14) and initial condition (3.13) have to be considered as well. Since the initial and boundary conditions are prescribed, they do not constitute degrees of freedom in the problem of minimizing the Tikhonov functional (3.67), but they still have to influence the solution in the right way. To achieve this the right hand side of $Au = 0$ is modified as follows. Let $u_{bd} \in \mathbb{R}^{MN}$ be the vector containing the discrete boundary values and initial values in the entries corresponding to the boundaries and zeros otherwise. In order to eliminate the prescribed nodes from system $Au = 0$, we subtract $-b = Au_{bd}$ which yields the new system $Au = -Au_{bd}$. The rows and columns corresponding to the free nodes can then easily be extracted using an index vector. The resulting matrices are named \bar{A} , \bar{D} , \bar{u} . Considering that there are $2N + M$ set values the minimization problem reduces slightly in dimension from NM to $NM - (2N + M)$ unknowns and from $5NM$ to $5NM - 5(2N + M)$ equations. This results in the following minimization problem:

$$\min \left\| \begin{bmatrix} \bar{A} \\ \sqrt{\alpha} \bar{D} \end{bmatrix} \bar{u} - \begin{bmatrix} \bar{b} \\ \sqrt{\alpha} \bar{D} \bar{F} \end{bmatrix} \right\|_2, \quad (3.88)$$

where $A = Dt + RDss$ is the realization of the discretization of the first term in (3.67) and the second set of equations in (3.88) corresponds to the discretization of the Sobolev norm in (3.67). There are different ways to solve this minimization problem, namely applying the LSQR algorithm for the solution of Sparse Linear Equations and Least Square Problems as published in Paige [37] or solving the linear system resulting from the first order optimality condition on the gradient of J or applying a sparse QR solver to the linear system (3.90). All methods are briefly described in the next sections and one of them is chosen based on performance thereafter.

Alternative method

Another way of incorporating the initial and boundary conditions is to see them as free variables and introduce a penalty term in (3.44) penalizing deviations from the prescribed initial and boundary values. This modified Tikhonov functional is given by

$$\bar{J}_\lambda = \|u_t + L_{par}u\|_{L_2(Q_T)}^2 + \lambda \|u - F\|_{H^2(Q_T)}^2 + \gamma \|u - u_{bd}\|_{L_2(Q_T)}^2. \quad (3.89)$$

However, note that there neither is proof for convergence of this functional towards the exact solution nor results on how to choose the parameters α, γ . We mention this to point out that there are certain slightly different formulations of the problem which might vary in the quality of solution and might possess more or less favourable numerical properties. Furthermore, it serves as a (numerical) test to determine the influence of the (incorporation of) boundary conditions on the overall characteristics of the solution. The minimization problem to be solved transforms into

$$\min \left\| \begin{bmatrix} A \\ \alpha D \\ \gamma I_F \end{bmatrix} u - \begin{bmatrix} 0 \\ \sqrt{\alpha} DF \\ \gamma u_{bd} \end{bmatrix} \right\|_2, \quad (3.90)$$

where $I_F \in \mathbb{R}^{NM}$ is a diagonal matrix containing ones in the entries corresponding to initial and boundary values.

3.4.4 LSQR

LSQR is an iterative method for the solution of sparse (asymmetric) linear equations and sparse least square problems i.e. problems of the form $Ax = b$ or $\min \|Ax - b\|_2$ with A being large and sparsely populated. Published in 1982 by Paige & Saunders [37] it is based on the bidiagonalization procedure of Golub and Kahan. The algorithm is analytically equivalent to the standard conjugate gradient method but achieves better numerical results.

We aim for a solution of the system $Bx = b$ with $B \in \mathbb{R}^{m \times m}$ symmetric and positive definite. Briefly, the method is an orthogonal projection onto the Krylov space

$$K_m(r_0, B) = \text{span}\{r_0, Br_0, B^2r_0, \dots, B^m r_0\}, \quad (3.91)$$

where $r_0 = Bx_0 - b$ is the initial residual. The Lanczos algorithm can be used to compute an orthonormal basis of the Krylov space K_m . In the process it generates a sequence of vectors v_i and scalars $\alpha_i, \beta_i, i > 0$ such that B is reduced to tridiagonal form.

Lanczos process

```

 $\beta_1 v_1 = b$ 
for  $i = 1, 2, \dots$  do
     $w_i = Bv_i - \beta_i v_{i-1}$ 
     $\alpha_i = v_i^T w_i$ 
     $\beta_{i+1} v_{i+1} = w_i - \alpha_i v_i$ 
end for,
    
```

where $v_0 = 0$ and all $\beta_i \geq 0$ are chosen such that $\|v_i\| = 1$. After k steps we have

$$BV_k = V_k T_k + \beta_{k+1} v_{k+1} e_k^T, \quad (3.92)$$

where $T_k = \text{tridiag}(\beta_i, \alpha_i, \beta_{i+1})$ and $V_k = [v_1, v_2, \dots, v_k]$. Multiplying equation (3.92) by an arbitrary vector y_k we obtain

$$BV_k y_k = V_k T_k y_k + \beta_{k+1} v_{k+1} \bar{y}, \quad (3.93)$$

where \bar{y} is the last entry of y_k . By definition $V_k(\beta_1 e_1) = b$ and we define

$$\begin{aligned} T_k y_k &= \beta_1 e_1 \\ x_k &= V_k y_k. \end{aligned} \quad (3.94)$$

Combining this with (3.92) we have $Bx_k = b + \bar{y}\beta_{k+1}v_{k+1}$ up to working precision. From here it is possible to construct iterative methods that solve $Bx = b$. The main difference is the way in which y_k is being eliminated from (3.94). For conjugate gradients one uses the Cholesky factorization $T_k = L_k D_k L_k^T$ for example.

Let us now turn our attention to the *damped least-squares* problem of the form

$$\min \left\| \begin{bmatrix} A \\ \lambda I \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|_2 \quad (3.95)$$

with $\lambda \in \mathbb{R}$. The solution to this system also satisfies

$$\begin{bmatrix} I & A \\ A^T & -\lambda^2 I \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix} \quad (3.96)$$

where $r = b - Ax$. Applying the Lanczos algorithm to this system is the same as using a bidiagonalization procedure from Golub & Kahan [17]. Furthermore, the relations (3.94) can after $2k + 1$ iterations be permuted to

$$\begin{bmatrix} I & B_k \\ B_k^T & -\lambda^2 I \end{bmatrix} \begin{bmatrix} t_{k+1} \\ y_k \end{bmatrix} = \begin{bmatrix} \beta_1 e_1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} r_k \\ \lambda I \end{bmatrix} = \begin{bmatrix} U_{k+1} & 0 \\ 0 & V_k \end{bmatrix} \begin{bmatrix} t_{k+1} \\ y_k \end{bmatrix}, \quad (3.97)$$

where $B_k \in \mathbb{R}^{k+1 \times k}$ is a lower bidiagonal matrix. One can see that y_k is the solution of another least-squares problem

$$\min \left\| \begin{bmatrix} B_k \\ \lambda I \end{bmatrix} y_k - \begin{bmatrix} \beta_1 e_1 \\ 0 \end{bmatrix} \right\|_2, \quad (3.98)$$

which in turn can efficiently be solved using orthogonal transformations.

3.4.5 Minimization via Conjugate gradient method

Another way of minimizing the functional (3.67) with respect to the values of u_λ at the grid points is to form the partial derivatives of the discrete version of the functional with respect to the entries of $u_\lambda \in \mathbb{R}^{MN}$. The discrete analogon of (3.44) is (3.67).

Using the relation

$$\frac{\partial u_j}{\partial u_{\bar{j}}} = u_j \delta_{j\bar{j}} \text{ and } \sum_{j=1}^{MN} u_j \delta_{j\bar{j}} = u_{\bar{j}} \quad (3.99)$$

one could obtain the gradient of the Tikhonov functional by forming all partial derivatives $\frac{\partial \hat{J}_\lambda}{\partial u_{\bar{j}}}$. This can be expressed in matrix form as follows

$$\nabla \hat{J}_\lambda = Gu - HF, \quad (3.100)$$

where $u, F \in \mathbb{R}^{MN}$ contains the entries of the gradient that are independent of u_λ and $G, H \in \mathbb{R}^{MN \times MN}$ is symmetric and positive definite (SPD). The first order optimality condition leads to the solution of the following linear system of equations

$$Gu - HF = 0. \quad (3.101)$$

One possible method for the solution of such linear systems with G being SPD is the conjugate gradient method (CG). It is also possible to apply direct sparse solvers to the system. Alternatively, an equivalent and much easier derivation of this linear system is obtained by considering the discrete version of the Tikhonov functional in the form

$$J_\lambda(u) = \|Au - b\|_2^2 + \lambda [\|Iu - IF\|_2^2 + \|D_t u - D_t F\|_2^2 + \|D_s u - D_s F\|_2^2 + \|D_{tt} u - D_{tt} F\|_2^2 + \|D_{ss} u - D_{ss} F\|_2^2]$$

and applying the gradient operator to obtain

$$\begin{aligned} \nabla J_\lambda(u) &= A^T Au - A^T b + \lambda I^T I[u - F] + \lambda D_t^T D_t[u - F] + \lambda D_s^T D_s[u - F] \\ &\quad + \lambda D_{tt}^T D_{tt}[u - F] + \lambda D_{ss}^T D_{ss}[u - F] \\ &= [A^T A + I^T I + D_t^T D_t + D_s^T D_s + D_{tt}^T D_{tt} + D_{ss}^T D_{ss}]u - A^T b \\ &\quad - \lambda [I^T I + D_t^T D_t + D_s^T D_s + D_{tt}^T D_{tt} + D_{ss}^T D_{ss}]F, \end{aligned} \quad (3.102)$$

where $A = D_t + RD_{ss}$ is the discrete differential operator as defined before.

3.4.6 Multifrontal Sparse QR factorization

An important way to efficiently solve large (sparse) linear system and least-squares problems directly is to factorize matrices into the form $A = QR$, where Q is orthogonal and R is upper triangular. Such a transformation is called QR factorization. The first QR factorization methods transformed A one row or column at a time and did not reach theoretical performance due to irregular access of memory. An efficient alternative is an approach which involves symbolic analysis, reordering to eliminate singletons and makes use of different forms of parallelism. The most relevant details for this multifrontal multithreaded QR factorization as implemented from Matlab R2009a upwards are extracted from Davis [13].

The procedure starts with the elimination of singletons, which are columns j with a single nonzero entry $a_{ij} > \tau$, where τ is a threshold value or no entry. Singletons are permuted to the left and

row i is permuted to the top. Column j and row i are then removed from the matrix. This is done for all singletons leading to a matrix

$$\begin{bmatrix} R_{11} & R_{12} \\ 0 & A_{22}, \end{bmatrix} \quad (3.103)$$

where R_{11} is upper triangular with diagonal entries larger than τ . The remaining matrix A_{22} is then ordered to reduce fill using Matlab's "COLAMD" method, which orders $A^T A$ without explicitly forming it. Next, "COLMOD" symbolically finds a LL^T factorization of the ordered matrix $A^T A$ by only analyzing A itself. For this purpose the elimination tree of $A^T A$ and the number of nonzero entries in each row of R are computed. The information contained in the elimination tree is important for the symbolic analysis as well as the numerical factorization. The next step is the identification of supernodes, which are groups of m adjacent rows with (nearly) identical nonzero pattern. Each supernode results in one frontal matrix which is a dense subset of A on which the actual QR factorization is performed. The information about supernodes is stored in the frontal matrix tree which gives rise to the frontal matrix after ordering, determination of the frontal matrix sizes and a corresponding assembly strategy. A Householder transformation is then applied to the frontal matrix. Potential ways of exploiting parallelisms are the parallel QR factorization of parts of the frontal matrix and the QR factorization of a single frontal matrix can be accelerated by using shared-memory multi-core capabilities of the "LAPACK" QR factorization package.

3.5 Numerical Experiments

In order to extract suitable model and discretization parameters and to validate the model some numerical tests are carried out. These tests are based on the performance of the regularization algorithm tested against the downward in time solution of the standard Black-Scholes equation with suitable final and boundary data. Thus the numerical solution of the Black-Scholes backwards in time is shortly outlined in the first section. Thereafter, the algorithm is tested for the inverse time heat problem. We determine the best way to solve the least-squares problem numerically and introduce potential alternative boundary conditions. The choice of the regularization parameter will also be dealt with.

3.5.1 The backward-time problem: Crank-Nicolson

The Black-Scholes equation is solved backwards in time, which is the forward problem to our inverse problem, to assess the quality of the regularization procedure described above. The Crank-Nicolson method will be applied to numerically solve (2.13). Since the interest rate is assumed to be zero we obtain the following discretization. For all reference forward solution of the problem we use $N = 400$ data points in each dimension. To verify that the implementation works, results for one example are compared to Matlab's implementation of the Black-Scholes price called "blsprice". In Fig. 3.2 one can see a plot of call \rightarrow and put \leftarrow as calculated using the Crank-Nicolson scheme for an option with expiration $T = 5/12$, strike $K = 50$ €, interest rate $r = 0$ and volatility $\sigma = 0.5$.

3.5.2 Verification test: time inverse heat solution

To validate the algorithm for a slightly different problem, the 1D heat equation is solved backward in time. Let the initial boundary value problem be given by

$$\frac{\partial T}{\partial t} + \alpha_d \frac{\partial^2 T}{\partial x^2} = 0 \quad \text{with } x \in (0, \pi), \quad t \in (0, T_e), \quad (3.104)$$

$$T(x, 0) = \psi(x), \quad x \in [0, \pi], \quad (3.105)$$

$$T(x, t) = 0, \quad x = \{0, \pi\}, \quad t \in [0, T_e], \quad (3.106)$$

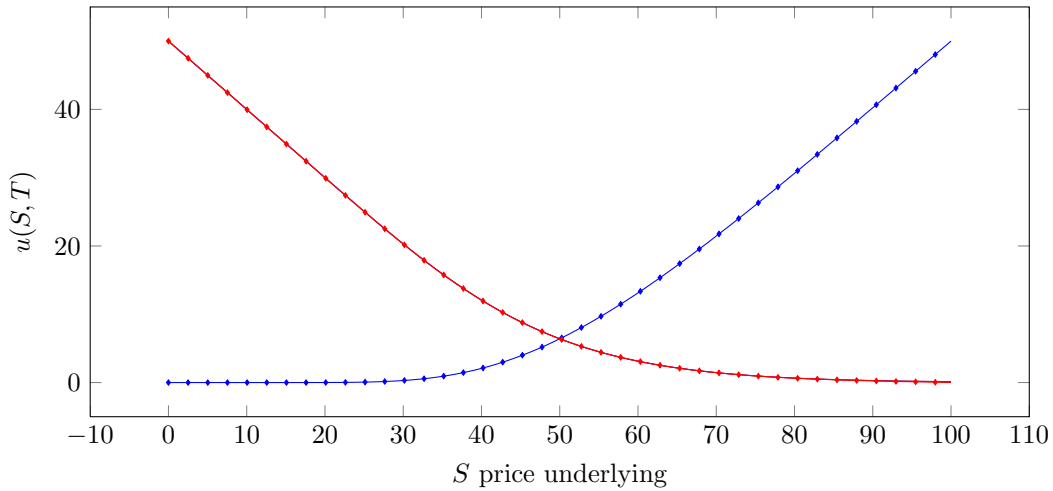





Figure 3.2: Price of an option with strike price $K = 50$ €, $T = 5/12$ years, in dependence of spot price S . Verification of analytical solution as calculated by Matlab for call — and put — vs. Crank-Nicolson scheme for call — and put —.

where α_d is the thermal diffusivity with units $[\frac{m^2}{s}]$ and is set to $\alpha_d \equiv 1$. The initial temperature distribution is prescribed by $\psi(x) = \sin(x)$. The equation is then first solved forwards in time using the Crank-Nicolson scheme as mentioned before to obtain the temperature distribution at the final times $T_e = \{0.05, 0.1, 0.2, 0.4\}$ s. These distributions are then contaminated with a random signal with noise level $\delta = \{10^{-1}, 10^{-2}, 10^{-3}\}$ in order to prevent committing inverse crime. The resulting distribution is used as the terminal condition for the algorithm solving the inverse problem up until time $t = 0$. Finally, the reconstructed initial distribution is compared to the true initial condition $\psi(x)$ in Fig. 3.3. For this test the regularization parameter λ is chosen heuristically and is set to $\lambda = 0.25$ for error level $\delta = 10^{-2}$. This way the constant α from (3.54) can be obtained. In our test case we obtain $\alpha = 0.15$ and we thus choose the regularization parameter according to $\lambda(\delta) = \delta^{0.3}$. In accordance with expectations one can see that the quality of the solution decreases with an increase in noise level and or for larger end times T_e . Note, that stable solutions are only guaranteed for $T_e \in (0, \frac{1}{2})$ because of the restriction $T_e < a_0/2$ in Theorem 3.54. Indeed, increasing T_e to $\frac{1}{2}$ in our numerical example leads to unusable results. The minimization problem (3.90) is solved using Matlab's sparse QR solver.

3.5.3 Assessment: How to include boundary conditions?

As discussed in Section 3.4.3 we consider two ways of implementing the initial (3.13) and boundary (3.14) conditions. In Fig. 3.4 the results are compared. One can observe that with a certain choice of parameters λ and γ the results are similar for both methods. However, the fixed boundary method  gives better results compared to the alternative boundary method  which makes sense since the forward problem  was solved using the same boundary conditions. In the following numerical computations the boundary conditions will be directly prescribed.

3.5.4 Bid-Ask spread requirements

When dealing with real market data, we assess that bid-ask spreads of an option's underlying are very tight. The bid-ask spread Δ is defined as $\Delta = s_a - s_b$. We observe a typical bid-ask spread of $\Delta = 0.05\% \frac{s_a + s_b}{2}$. The resulting problem is that the dynamics on the domain $Q_{2\tau}$ are not sufficient to extract any additional information when compared to the interpolation of option prices between $u_a(t)$ and $u_b(t)$ as defined in (3.14). Thus, we exclude options with spreads $\Delta < 0.5\% \frac{s_a + s_b}{2}$ from

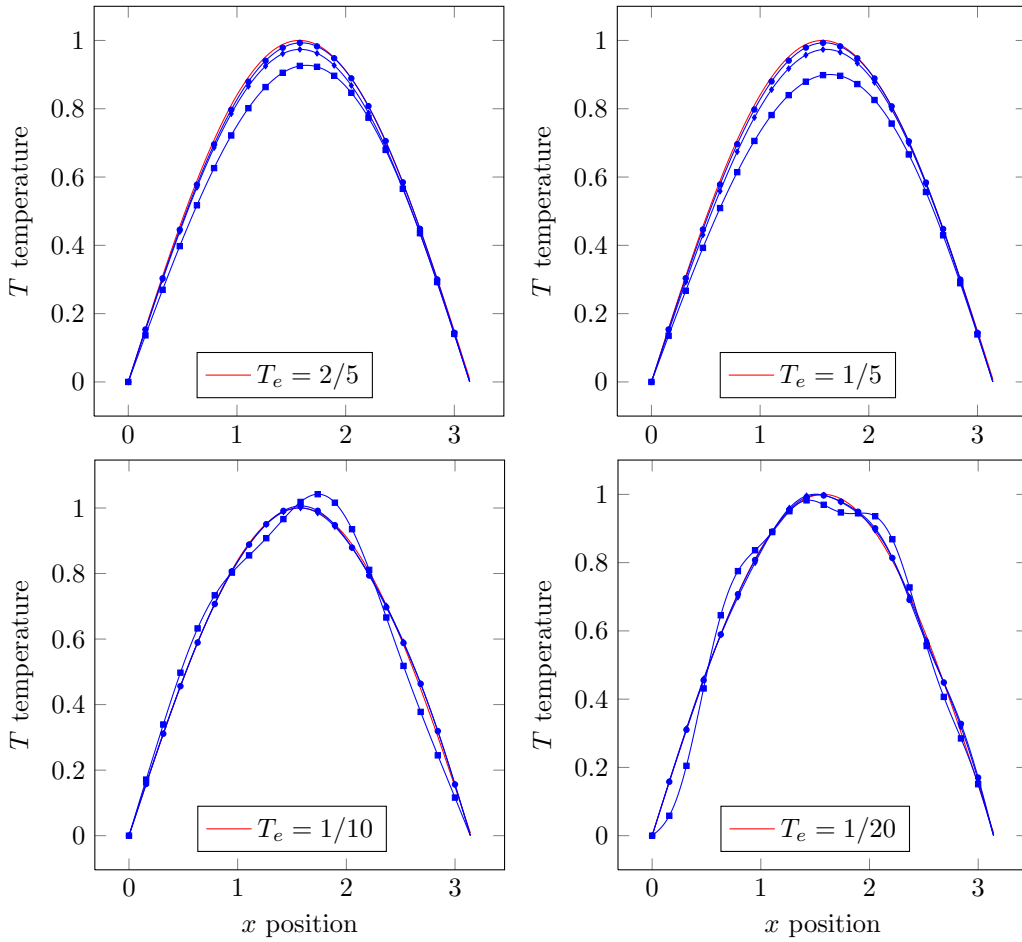


Figure 3.3: Heat equation solved backward with noise level $\delta = 10^{-3}$ \bullet , $\delta = 10^{-2}$ \bullet and $\delta = 10^{-1}$ \bullet and true initial distribution $\psi(x)$ --- for different terminal times T_e .

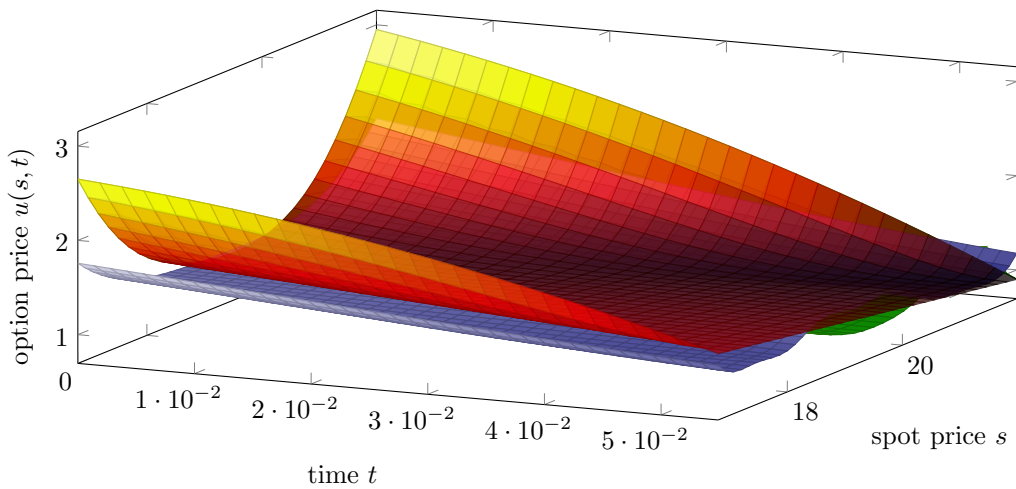


Figure 3.4: Solution of price forecasting for set boundaries \blacklozenge vs. free boundaries \blacklozenge compared to forward solution of the problem as obtained by Crank-Nicolson method \blacklozenge .

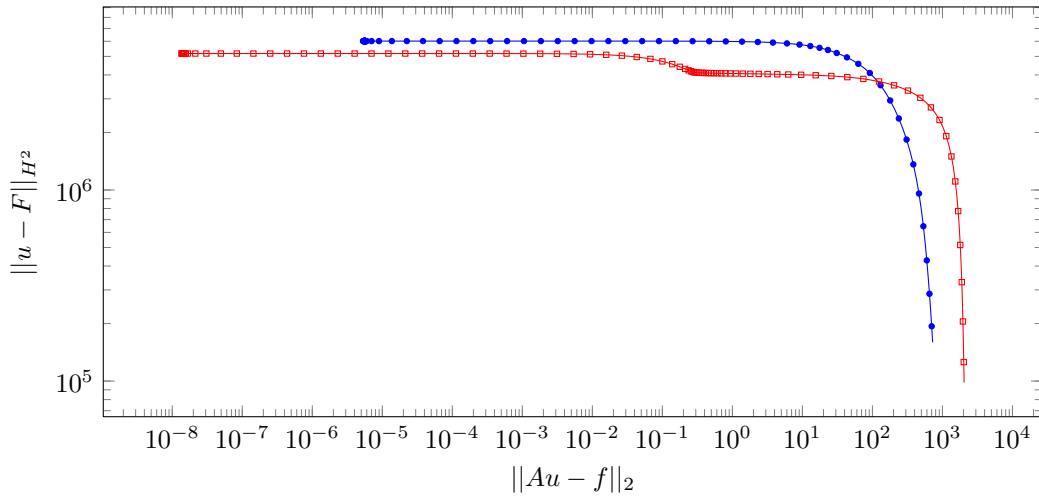


Figure 3.5: L-Curve in double logarithmic scale for options with different bid ask spreads of the underlying: large spread $\Delta = 2.6\% \frac{s_a + s_b}{2}$ —□— vs. tight spread $\Delta = 0.12\% \frac{s_a + s_b}{2}$ —●— for regularization parameter $\lambda \in (10^{-8}, 1)$.

the following considerations.

3.5.5 Choice of regularization parameter

The magnitude of the regularization parameter λ greatly influences the solution's behavior. There are two classes of parameter selection algorithms: methods that rely on information on the noise level in the input data and such methods that do not and are therefore called noise free parameter choice rules. However, in the deterministic case there is *always* a noise such that for any regularization method with error free parameter choice the regularized solution will not converge to the true solution. This result is also referred to as Bakushinskii's veto [7].

L-curve

The L-curve criterion is a popular error free parameter choice method. It is a graphical tool which for all valid regularization parameters plots the discrete smoothing norm (in our case the discrete analogon of $\|u - F\|_{H^2}$) versus the corresponding residual norm $\|Au - f\|_2$. In the case of Tikhonov regularization the L-curve is continuous because one can choose the regularization parameter from an interval of real number values. The L-curve criterion visualizes the tradeoff between minimization of the two norms mentioned. In a lot of cases the plot in double logarithmic scale of these two norms against each other has the shape of the letter L. One then aims to seek the point where the curvature is maximal, being the point corresponding to the corner of the letter L. Let us calculate the norms for regularization parameters $\lambda \in (10^{-8}, 1)$ for one randomly picked option and plot the results in double logarithmic scale. In Fig. 3.5 one can see that the resulting plot —●— does not resemble the letter L. However, increasing the difference between the bid and ask price of the underlying $-s_b$ and s_a respectively—one can observe that the resulting L-curve —□— actually resembles the letter L. Since the bid-ask spreads we are dealing with are typically smaller and since it is difficult to numerically determine this point in this case, we move on to the next possible parameter selection criterion.

Generalized Cross Validation (GCV)

Another error free parameter selection method is the Generalized Cross Validation method (GCV). The idea is to minimize the predictive mean-square error $\|Au_\lambda - f^*\|_2$, where f^* is the exact data

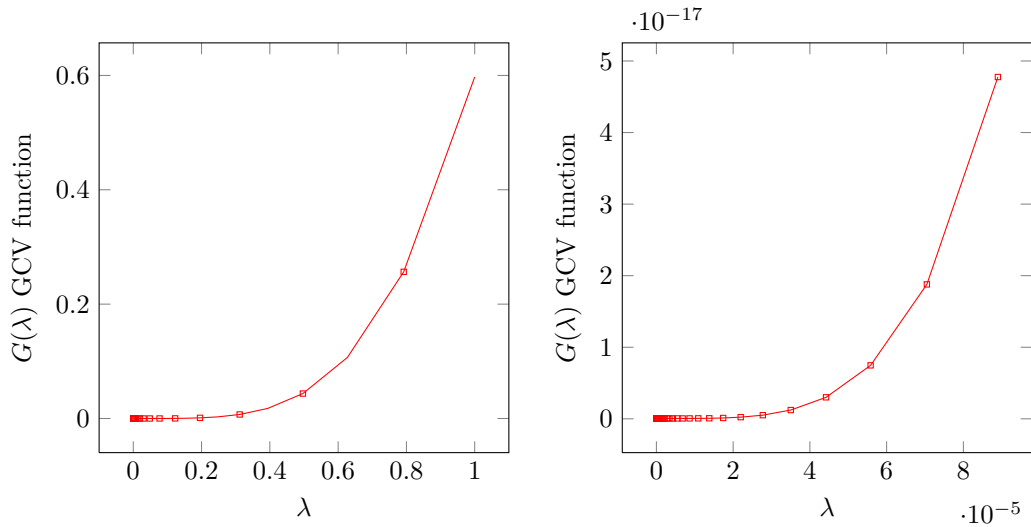


Figure 3.6: GCV function $G(\lambda)$ for the full range of regularization parameters $\lambda \in (10^{-8}, 1)$ on the left and zoomed in around the minimum on the right.

as defined in (3.46). The GCV function works with the given data f and is defined as

$$G(\lambda) = \frac{\|Au_\lambda - f\|_2^2}{(\text{Tr}(I - AA^\#))^2}, \quad (3.107)$$

where $A^\# = (A^T A + \lambda D^T D)^{-1} A^T$ is the regularized inverse. Using generalized singular value decomposition the matrices A and D can be decomposed as

$$A = UCX^{-1}, \quad D = V SX^{-1}, \quad (3.108)$$

where U, V are orthonormal matrices, X is nonsingular with $A^T A$ orthogonal columns and C, S are diagonal matrices with diagonal elements being ordered non-decreasingly. Then consider the relations

$$A(A^T A + \lambda D^T D)^{-1} A^T = UCX^{-1}(X^{-T}(C^2 + \lambda S^2)X^{-1})^{-1} X^{-T} C^T U^T \quad (3.109)$$

$$\text{Tr}(I - A(A^T A + \lambda D^T D)^{-1} A^T) = \text{tr}(I) - \text{tr}(C^2(C^2 + \lambda S^2)^{-1}), \quad (3.110)$$

where we used the fact that U is orthogonal and the relation $\text{Tr}(ABC) = \text{Tr}(CAB)$. Remarks on the general form Tikhonov form and parameter selection criteria in this case can be found in Chung [12]. The λ_0 value corresponding to the minimum of the GCV function $G(\lambda_0)$ is the optimal parameter then. Fig. 3.6 shows one such plot for a random option. Observe that the minimum found does not represent the best parameter when compared to the reference forward solution u_{CS} and is also hard to determine numerically since machine precision level is reached. A similar scenario is found on page 198 in Hansen [19], where the occasional failure of the method is mentioned and the GCV function around the minimum is very flat. Thus, we refrain from using the GCV method for our problem and continue with a heuristic estimation of the optimal parameter in the next section.

Heuristics: Parameter selection by forward solution fitting

Since it seems to be hard to find a regularization parameter in our case, we introduce a heuristic approach now. Let the numerical solution of the forward problem as obtained by Crank-Nicolson scheme be u_{CS} . The idea is to pick 100 different options and choose the average regularization parameter which minimizes the difference between the regularized solution and the forward solution. Let the mean-square error between these two solutions be defined as $e_2 = \|u_\lambda - u_{CS}\|_2$. The

resulting regularization parameter is $\lambda = 5 \cdot 10^{-3}$, which is in good agreement with the parameter used by Klibanov [29].

More Heuristics: Multiparameter Regularization

We observed that the reconstruction quality is sometimes higher and the range of admissible regularization parameters larger when the different terms in the (discretized) Sobolev norm $\|u - F\|_{H^2}$ are assigned different weights. More specifically, this means that there is not only one regularization parameter λ but five of them. One way to do this is to remove the natural scaling stemming from the finite difference discretization of the Sobolev operator, such that

$$\|u - F\|_{H^2} = \lambda \|u - F\|_2 + \lambda 2dt \|D_t(u - F)\|_2 + \lambda 2ds \|D_s(u - F)\|_2 + \lambda dt^2 \|D_{tt}(u - F)\|_2 + \lambda ds^2 \|D_{ss}(u - F)\|_2. \quad (3.111)$$

Note, that by doing so we essentially make the problem a *multiparameter Tikhonov* regularization problem and also leave the framework of the convergence proof presented before. For more information on multiparameter Tikhonov regularization and possible parameter selection criteria in this case see among few available papers on this topic for example Ito [22], Reichel [38] or Chung [12]. Ito [22] provides further relevant references in section 1 of his paper. Rerunning the analysis from the preceding section we retrieve as the regularization parameter $\lambda = 0.75$ for our problem. Since, this is only an experimental observation and since the convergence results from Section 3.4 do not hold for the case of multiple parameters, we will not actually apply this approach when forecasting prices in Chapter 5.

3.5.6 Numerical instability of gradient method and lsqr algorithm

Solving the minimization problem using Matlab's implementation of the LSQR algorithm from Section 3.4.4 and the gradient method discussed in Section 3.4.5 led to more unstable results when compared to the multifrontal QR solution from Section 3.4.6. The matrix product $A^T A$ is actually calculated in the gradient formulation leading to considerable round-off error because the condition number of $A^T A$ is high, which explain its unsatisfactory performance. In Fig. 3.7 the results using different solvers are plotted.

3.5.7 Alternative Boundary conditions

Since the boundary conditions chosen by Klibanov [29] are very sensitive to price changes and have a strong influence on the trading strategy (see Section 5.4). As an insightful thought experiment, consider for example an option trading at constant 0.95\$ for two days and then making a 5% jump to 1\$ on the third day, i.e. prices $p = [1, 1, 1.05]$ corresponding to the times $t = [-2\tau, -\tau, 0]$. Quadratic interpolation and evaluation of the resulting polynomial until time 2τ imply an anticipated price of $p_{2\tau} = 1.3$. However, compare this to the result for a 5% decrease at the third day, yielding $p_{2\tau} = 0.7$. This corresponding quadratic functions $\rightarrow\bullet$, $\leftarrow\bullet$ are plotted in Fig. 3.8. Therefore, alternative choices of the boundary values (3.14) are proposed. The aim is to still use the past observations for u_b, u_a at $\tau \in \{-2\tau, -\tau, 0\}$, but to reduce the strong dynamics implied by the quadratic polynomial in (3.14).

Linear Regression

As a first alternative, we propose to use the linear regression function as boundary values. As one can see in Fig. 3.8 the dynamics are damped compared to the quadratic polynomial while still exhibiting a tendency to increase $\rightarrow\bullet$, decrease $\leftarrow\bullet$.

Constant Average

One might also argue that a single jump in price does not justify modeling future prices with a lasting trend. In this case we suggest to use constant boundary conditions using the mean value

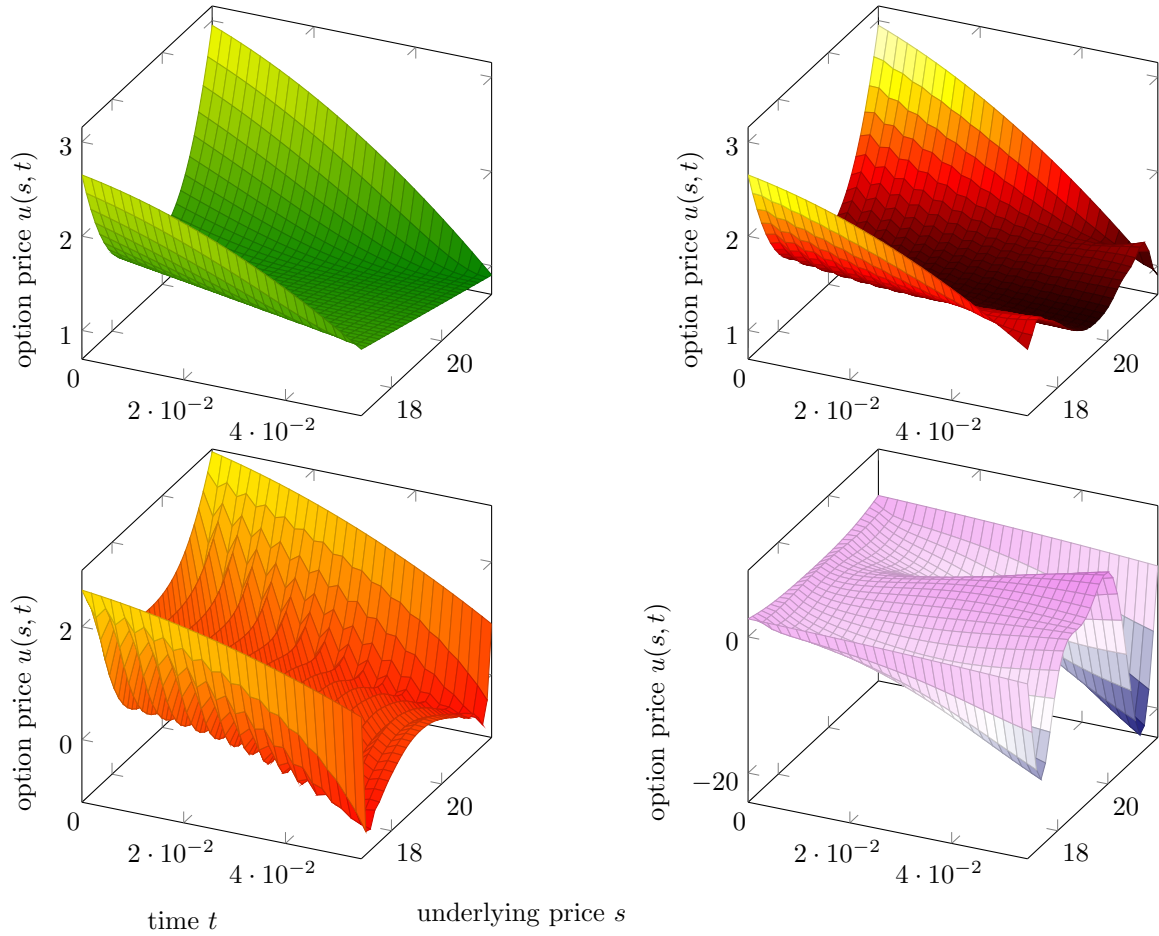


Figure 3.7: Solution as obtained using different methods: reference solution , Matlab's backslash (Multifrontal QR) , Matlab's LSQR and backslash solution of the gradient system . Regularization parameter was set to $\lambda = 0.1$.

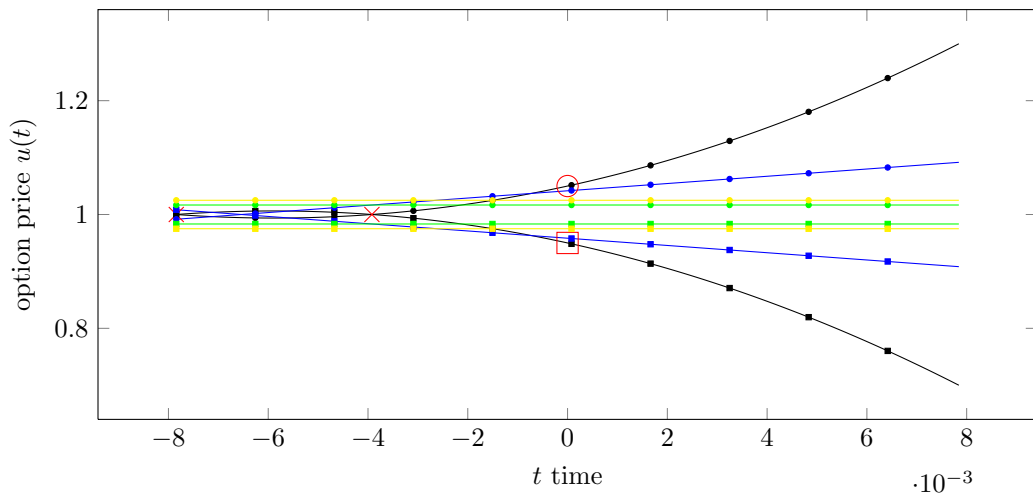
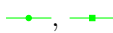


Figure 3.8: Price using quadratic polynomial \bullet , +5%, \blacksquare , -5%, regression \bullet , +5%, \blacksquare , -5%, mean \bullet , +5%, \blacksquare , -5%, weighted average \bullet , +5%, \blacksquare , -5% .

of the past observations as the constant boundary condition. The corresponding plots , are also there in Fig. 3.8.

Constant Weighted Average

Similarly, the more recent observations could be given more weight when calculating the mean value. One possibility is to assign the weights $\omega = [\frac{1}{6}, \frac{2}{6}, \frac{3}{6}]$ to the prices p , resulting in the weighted average

$$\bar{p}_w = \omega_1 p_1 + \omega_2 p_2 + \omega_3 p_3 = 1.025/0.975. \quad (3.112)$$

Chapter 4

Parameter Identification: Implied Volatility as an Inverse Problem

As already pointed out in the introduction the volatility implied by the Black-Scholes equation is observed to depend on both the underlying price s and the time t . In reality one can observe that options whose strike is either far out of or far into the money have a greater implied volatility while volatility tends to decrease the more the time to expiry. This dependence can be seen in so-called volatility surface plots in which implied volatility is plotted against the strike price K and time to expiry T . For example consider Fig. 4.1 in which the volatility data for 288 call options on Microsoft (MSFT) shares on 05.07.2015 were plotted. According to Rubinstein [40] this effect has only been observed in stock options after the financial breakdown of 1987 whereas it was already known for foreign currency options by then. As a result many practitioners and researchers tried to extend the Black-Scholes model to incorporate the volatility coefficients dependence on stock price and time. This class of models is therefore called *local volatility* models and the Black-Scholes in the extended form becomes

$$\frac{\partial u}{\partial t} + \frac{1}{2}s^2\sigma(s,t)^2\frac{\partial^2 u}{\partial s^2} + s\mu\frac{\partial u}{\partial s} - ru = 0. \quad (4.1)$$

The volatility coefficient is assumed to satisfy $0 < m < \sigma(s,t) < M < \infty$ and belong to the Hölder space $C^\lambda(\bar{\omega}^*)$, $0 < \lambda < 1$ for some interval $\omega^* \subset (0, \infty)$. Furthermore there is still a final condition imposed at the expiry date T :

$$u(s, T) = (s - K)^+ = \max(0, s - K), \quad 0 < s. \quad (4.2)$$

The inverse problem of option pricing can then be formulated as determining the volatility σ for given values

$$u(s^*, t^*; K, T) = u^*(K, T), \quad K \in \omega^*, \quad T \in I^* = \{T_1, T_2\}, \quad (4.3)$$

where s^* is the price of the underlying at time t^* and is assumed to be contained in ω^* , i.e. $s^* \in \omega^*$. The main problem with coefficient inverse problems (CIP) is that the cost functionals associated with the Tikhonov formulation of the problem are non-convex. This means that the cost functional

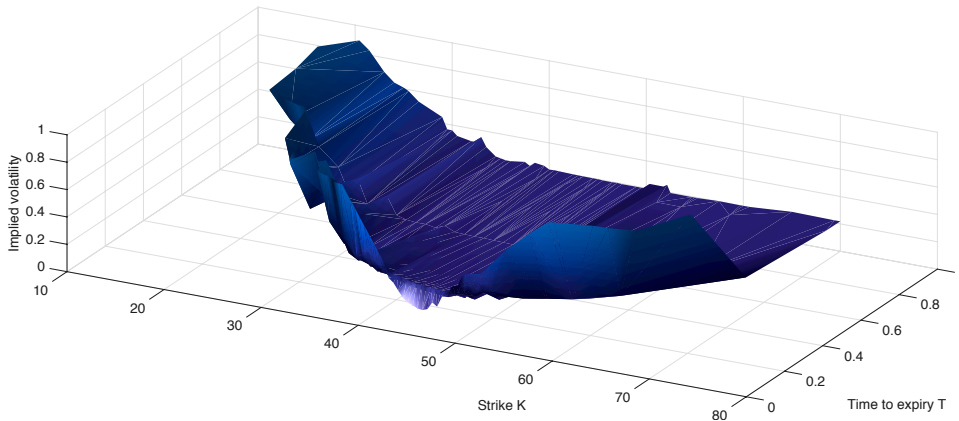


Figure 4.1: Implied volatility surface for MSFT options. Data retrieved on 05.07.2015.

potentially has many ravines and one should roughly know the location of the minimum to find a global minimum. One way to obtain globally convergent numerical algorithms is to construct strictly convex Tikhonov-like cost functionals using Carleman Weight Functions (CWF). For some theory on this topic see for example Klibanov [28], Timonov [31]. In the following, the non-linear coefficient inverse problem (4.1)-(4.3) will be linearized to obtain estimates of the volatility $\sigma(s, t)$. However, if one wants to use this s -dependent volatility coefficient in the forecasting problem from section 3.2 the convergence estimates of section 3.4 do not hold anymore. The challenge is to construct a modified Tikhonov-like functional such that convergence can be proved in this case. Inspiration for this task can be drawn from e.g. Klibanov [26, 28] or Timonov [31]. In any case, the algorithm from section 3.2 will be tested with the data obtained from the solution of the CIP in a numerical experiment.

4.1 Ill-posedness of the inverse problem and literature overview

In the past 35 years many approaches have been made to reconstruct the volatility surface given a discrete set of option prices. Avellaneda [5] applies dynamic programming to minimize a relative entropy functional, Bouchaev [9] approximates the option price as the sum of the Black-Scholes price and a correction for the non-constant volatility stemming from the parametrix expansion of the fundamental solution whereas Lagnado and Osher [32] minimize a Tikhonov like functional to solve the inverse problem. For more complete surveys on the topic see e.g. Zubelli [49] or Reisinger [39]. Since the solution of the nonlinear inverse option pricing model (nonlinear in the sense of the volatility depending nonlinearly on K and T) is unstable, therefore numerically challenging and regularization techniques hard to apply on sparse real market data we turn to a linearization approach suggested by Isakov [11, 21] which is numerically more robust to sparse data and better suited to the analysis of a large data set in terms of computational complexity.

4.2 Linearization approach by Isakov

Based on the aforementioned papers by Isakov [11, 21] the volatility is being linearized at two different times T_1, T_2 . The local volatility is assumed to be given by

$$\frac{1}{2}\sigma(s, t)^2 = \frac{1}{2}\sigma_0^2 + f_0^*(s) + tf_1^*(s), \quad (4.4)$$

where $f_0^*, f_1^* \in C^0(\omega)$ are small perturbations of σ_0 which in turn is the constant implied volatility defined as the solution of the Black-Scholes equation with constant volatility coefficient. Calculation of σ_0 is straightforward and can be done by applying a Newton fix point algorithm to the Black-Scholes equation. Alternative ways as e.g. rational approximation Li [33] also exist. Outside of the interval ω^* we set $f_0^* = f_1^* = 0$. This conjecture is acceptable if the volatility is not changing fast with respect to stock price and slow with respect to time.

Since the analysis of the problem is analytically and numerically easier in terms of the strike price K and maturity date T the Black-Scholes equation is formulated in terms of these variables yielding the Dupire equation. This dual equation is derived in the next section.

4.2.1 Continuous time formula: Dupire

Dupire [14] derived a dual equation to the Black-Scholes in 1993, which is not defined in terms of s and t but describes the option price in dependence of strike K and expiry T . The result stems from the question whether under certain conditions the Black-Scholes model can be modified to accommodate for non-constant volatility while keeping the model complete in the sense of being able to hedge the option with the underlying asset. The question was whether it is possible to

find a risk-neutral diffusion process of the form

$$\frac{dS}{S} = r(t)dt + \sigma(S, t)dW, \quad (4.5)$$

given arbitrage-free prices $u(K, T)$. The relation

$$u(K, T) = \int_0^\infty \max(S - K, 0)\phi_T(S)dS = \int_K^\infty (S - K)\phi_T(S)dS \quad (4.6)$$

results in

$$\phi_T(K) = \frac{\partial^2 u}{\partial K^2}(K, T) \quad (4.7)$$

after differentiating twice with respect to K , where ϕ_T is the risk-neutral density function of the spot at time T . The question is whether there exists a unique diffusion process $dx = a(x, t)dt + b(x, t)dW$ generating these densities ϕ_T ? Assuming for the moment that the interest rate is zero consider $dx = b(x, t)dW$. Using the Fokker-Planck equation and some calculation it is shown in Dupire [14] that the following relation holds

$$\frac{b^2(K, T)}{2} \frac{\partial^2 u}{\partial K^2} = \frac{\partial u}{\partial T} \quad (4.8)$$

for European options. From this b can be obtained because $\frac{\partial^2 u}{\partial K^2}$ and $\frac{\partial u}{\partial T}$ are known from observable market data. One only has to think of suitable interpolation or discrete differentiation procedures to obtain these quantities. Comparing this to the diffusion process (4.5) one identifies the volatility $\sigma = \frac{b(S, t)}{S}$. The derivation for non-zero interest rate is slightly more complicated and results in the Dupire equation

$$\frac{\partial u}{\partial T} - \frac{1}{2}K^2\sigma(K, T)^2 \frac{\partial^2 u}{\partial K^2} + \mu K \frac{\partial u}{\partial K} + (r - \mu)u = 0. \quad (4.9)$$

4.2.2 Linearization at constant volatility

Equation (4.9) and final condition (4.2) are transformed by

$$y = \ln\left(\frac{K}{S^*}\right), \quad \tau = T - t^* \quad (4.10)$$

$$a(y, \tau) = \sigma(s^* e^y, \tau + t^*), \quad U(y, \tau) = u(s^* e^y, \tau + t^*) \quad (4.11)$$

into

$$\frac{\partial u}{\partial \tau} - \frac{1}{2}a(y, \tau)^2 \frac{\partial^2 u}{\partial y^2} + \left(\frac{1}{2}a(y, \tau)^2 + \mu\right) \frac{\partial u}{\partial y} + (r - \mu)u = 0 \quad (4.12)$$

$$u(y, 0) = s^*(1 - e^y)^+, \quad y \in \mathbb{R}. \quad (4.13)$$

The additional market data become

$$u(y, \tau_j) = u_j(y), \quad y \in \omega, j = 1, 2, \quad (4.14)$$

where ω is the transformed interval ω^* and $u_j(y) = u^*(s^* e^y, T_j)$. Also note that

$$\frac{1}{2}a^2(y, \tau) = \frac{1}{2}\sigma_0^2 + f_0(y) + \tau f_1(y), \quad (4.15)$$

with

$$f_0(y) = f_0^*(s^* e^y) + t^* f_1^*(s^* e^y), \quad f_1(y) = f_1^*(s^* e^y). \quad (4.16)$$

The idea is to construct the solution u as $u = V_0 + V + v$, where v vanishes quadratically with respect to (f_0, f_1) and V_0 solves (4.12) with $a = \sigma_0$ and V is the solution to

$$\frac{\partial V}{\partial \tau} - \frac{1}{2}\sigma_0^2 \frac{\partial^2 V}{\partial y^2} + \left(\frac{1}{2}\sigma_0^2 + \mu\right) \frac{\partial V}{\partial y} + (r - \mu)V = \alpha_0(f_0(y) + \tau f_1(y)), \quad (4.17)$$

$$V(y, 0) = 0, \quad y \in \mathbb{R}, \quad (4.18)$$

with

$$\alpha_0(y, \tau) = \frac{s^*}{\sqrt{2\pi\tau\sigma_0}} e^{-\frac{y^2}{2\tau\sigma_0^2} + cy + d\tau}, \quad (4.19)$$

$$c = \frac{1}{2} + \frac{\mu}{\sigma_0^2}, \quad d = -\frac{1}{2\sigma_0^2} \left(\frac{\sigma_0^2}{2} + \mu\right)^2 + \mu - r \quad (4.20)$$

and

$$V(y, \tau_j) = V_j(y) = u_j(y) - V_0(y, \tau_j), \quad j = 1, 2, \quad y \in \omega. \quad (4.21)$$

This linearization is in accordance with the theory presented in e.g. Bouchaev [10] for the inverse option pricing model. Further substitution of $V = e^{cy+d\tau}W$ yields

$$\frac{\partial W}{\partial \tau} - \frac{1}{2}\sigma_0^2 \frac{\partial^2 W}{\partial y^2} = \alpha(f_0(y) + \tau f_1(y)), \quad 0 < \tau < \tau^*, \quad y \in \mathbb{R}, \quad (4.22)$$

$$W(y, 0) = 0, \quad y \in \mathbb{R}, \quad (4.23)$$

with

$$\alpha(y, \tau) = \frac{s^*}{\sqrt{2\pi\tau} e^{-\frac{y^2}{2\tau\sigma_0^2}}} \quad (4.24)$$

and the final data

$$W(y, \tau_j) = W_j(y) = e^{-cy-d\tau_j} V_j(y). \quad (4.25)$$

In one last simplification α is replaced by the y independent function

$$\alpha_1(\tau) = \frac{s^*}{\sqrt{2\pi\tau}\sigma_0} \quad (4.26)$$

and the domain \mathbb{R} is restricted to a finite interval $\Omega \supset \bar{\omega}$.

4.2.3 Uniqueness and Stability

Theorem 5 *Let $\omega = (-b, b)$ and $f_0 = f_1 = 0$ outside ω . Furthermore demand*

$$\frac{\tau_1^2 + \tau_2^2 + \sqrt{\tau_1\tau_2}(\tau_1 + \tau_2)3b^2}{\tau_1\tau_2(\tau_2 - \tau_1)2\sigma_0^2} < 1, \quad (4.27)$$

$$\frac{(\sqrt{\frac{\tau_1}{\tau_2}} + \sqrt{\frac{\tau_2}{\tau_1}} + 2)\frac{b^2}{\sigma_0^2} + 2\sqrt{2\pi}(\sqrt{\tau_1} + \sqrt{\tau_2})\frac{b}{\sigma_0}}{2(\tau_2 - \tau_1)} < 1. \quad (4.28)$$

This guarantees the uniqueness of the solution $(f_0, f_1) \in L^\infty(\omega) \times L^\infty(\omega)$ of the inverse option pricing problem (4.22)-(4.26).

Proof The proof will only be sketched here. For the complete proof see Isakov [21]. One starts by proving the uniqueness of a solution to the approximate inverse option pricing problem (4.22)-(4.26) by applying a Fourier transform to (4.22), solving the generated ordinary differential equation and using a Taylor expansion of the exponential function to prove boundedness of the system of integral equations. Using the definition of Sobolev norms via the Fourier transform the stability and uniqueness is proved.

Equivalence to the following system of Fredholm equations

$$\int_{\omega} K_0(x, y; \tau_1) f_0(y) + K_1(x, y; \tau_1) f_1(y) dy = W_1(x), x \in \omega \quad (4.29)$$

$$\int_{\omega} K_0(x, y; \tau_2) f_0(y) + K_1(x, y; \tau_2) f_1(y) dy = W_2(x), x \in \omega, \quad (4.30)$$

is then shown, where

$$K_0(x, y; \tau) = S_* \int_{\frac{|x-y|+|y|}{\sigma_0\sqrt{2\tau}}}^{\infty} e^{-\tau^2} d\tau, \quad (4.31)$$

$$K_1(x, y; \tau) = S_* \left(\frac{\sqrt{\tau}}{\sqrt{2}\sigma_0} (|y| - |x - y|) e^{-\frac{(|x-y|+|y|)^2}{2\tau\sigma_0^2}} + \left(\frac{x^2 - 2xy}{\sigma_0^2} + \tau \right) \int_{\frac{|x-y|+|y|}{\sigma_0\sqrt{2\tau}}}^{\infty} e^{-\tau^2} d\tau \right), \quad (4.32)$$

with $S_* = \frac{s^*}{\sigma_0^2\sqrt{\pi}}$ and $W_j(x) = e^{-cx-d\tau_j} V_j(x)$ with $V_j(x) = u_j(x) - V_0(x, \tau_j)$ as defined before. This means that it suffices to show the uniqueness and stability of the solution to this system of Fredholm equations to prove uniqueness of the linearized option pricing problem.

4.2.4 Numerical solution of the forward problem

This section deals with the numerical solution of the forward problem.

$$\frac{\partial u}{\partial \tau} - \frac{1}{2} a(y, \tau)^2 \frac{\partial^2 u}{\partial y^2} + \left(\frac{1}{2} a(y, \tau)^2 + \mu \right) \frac{\partial u}{\partial y} + (r - \mu)u = 0, \quad (4.33)$$

The forward problem is solved using an implicit respective combination of explicit and implicit methods. Consider the Dupire equation (4.12) which is being discretized using the so-called θ -method. For simplicity we set $r = \mu = 0$. Time derivatives are discretized using first order finite differences whereas derivatives with respect to the strike price are second order accurate finite differences. The θ -method requires a θ -weighted average of explicit and implicit discretization to be formed. Inserting these averages into (4.35) yields:

$$\begin{aligned} \frac{u_j^{n+1} - u_j^n}{\delta t} = & \frac{1}{2} a(y, \tau)^2 \left[\theta \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{\delta x^2} + (1 - \theta) \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{\delta x^2} \right] \\ & + \frac{1}{2} a(y, \tau)^2 \left[\theta \frac{u_{j+1}^{n+1} - u_{j-1}^{n+1}}{2\delta x} + (1 - \theta) \frac{u_{j+1}^n - u_{j-1}^n}{2\delta x} \right]. \end{aligned} \quad (4.34)$$

Defining $\sigma = \frac{a(y, \tau)^2 \delta t}{2\delta x^2}$ and $\rho = -\frac{a(y, \tau)^2 \delta t}{4\delta x}$ we rearrange equation (4.35)

$$\begin{aligned} u_j^{n+1} + 2\sigma\theta u_j^{n+1} - \sigma\theta u_{j+1}^{n+1} - \rho\theta u_{j+1}^{n+1} - \sigma\theta u_{j-1}^{n+1} + \rho\theta u_{j-1}^{n+1} = \\ u_j^n - (1 - \theta)\sigma 2u_j^n + (1 - \theta)\sigma u_{j+1}^n + (1 - \theta)\rho u_{j+1}^n + \sigma(1 - \theta)u_{j-1}^n - \rho(1 - \theta)u_{j-1}^n \\ j = 2, \dots, N - 1 \end{aligned} \quad (4.35)$$

We prescribe zero Dirichlet boundary conditions such that $U_1^n = U_N^n = 0$ which yields

$$j = 1 : -\sigma\theta u_2^{n+1} - \rho\theta u_2^{n+1} = (1 - \theta)\sigma u_2^n + (1 - \theta)\rho u_2^n \quad (4.36)$$

$$j = N : -\sigma\theta u_{N-1}^{n+1} + \rho\theta u_{N-1}^{n+1} = (1 - \theta)\sigma u_{N-1}^n - (1 - \theta)\rho u_{N-1}^n \quad (4.37)$$

In matrix notation we have

$$A = \begin{bmatrix} 0 & -\sigma\theta - \rho\theta & 0 & 0 & \dots & 0 \\ -\sigma\theta + \rho\theta & 1 + 2\sigma\theta & -\sigma\theta - \rho\theta & 0 & \dots & 0 \\ 0 & -\sigma\theta + \rho\theta & 1 + 2\sigma\theta & -\sigma\theta - \rho\theta & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\sigma\theta + \rho\theta & 1 + 2\sigma\theta & -\sigma\theta - \rho\theta \\ 0 & \dots & \dots & 0 & -\sigma\theta + \rho\theta & 0 \end{bmatrix} \quad (4.38)$$

$$(4.39) \quad B = \begin{bmatrix} 0 & (1-\theta)(\sigma+\rho) & 0 & 0 & \cdots & 0 \\ (1-\theta)(\sigma-\rho) & 1-2\sigma(1-\theta) & (1-\theta)(\sigma+\rho) & 0 & \cdots & 0 \\ 0 & (1-\theta)(\sigma-\rho) & 1-2\sigma(1-\theta) & (1-\theta)(\sigma+\rho) & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (1-\theta)(\sigma-\rho) & 1-2\sigma(1-\theta) & (1-\theta)(\sigma+\rho) \\ 0 & \cdots & \cdots & 0 & (1-\theta)(\sigma-\rho) & 0 \end{bmatrix}$$

$$Au^{n+1} = Bu^n \quad (4.40)$$

For all computations we chose $\theta = 1$.

4.2.5 Numerical tests: reconstruction of given volatility surfaces

In order to test the proposed algorithm, the Dupire equation (4.9) is solved with a prescribed volatility function $a(y, \tau) = \frac{\sigma_0^2}{2} + f(y)$ to obtain the values $u_j(y)$. Then the system of Fredholm equations (4.29), (4.30) is solved and the reconstructed volatility perturbation $\bar{a}(y, \tau) = \frac{\sigma_0^2}{2} + f_0(y) + \tau f_1(y)$ is compared to the given volatility $a(y, \tau)$. We make use of the fact that

$$K_0(x, y; \tau) = S_* \int_{\frac{|x-y|+|y|}{\sigma_0\sqrt{2\tau}}}^{\infty} e^{-\tau^2} d\tau = \frac{\sqrt{\pi}}{2} S_* \operatorname{erfc} \left(\frac{|x-y|+|y|}{\sigma_0\sqrt{2\tau}} \right), \quad (4.41)$$

where erfc is the complementary error function. The integrals are calculated using a composite midpoint rule. For more information on the numerics of (ill-posed) Fredholm integral equations of the first kind see e.g. Baker [6]. We discretize the domain $\omega = (-b, b)$ with N grid points. This way (4.29), (4.30) becomes

$$\sum_{k=1}^N dy K_0(x, y_k, \tau_1) f_0(y_k) + K_1(x, y_k, \tau_1) f_1(y_k) = W_1(x), \quad x \in \omega \quad (4.42)$$

$$\sum_{k=1}^N dy K_0(x, y_k, \tau_2) f_0(y_k) + K_1(x, y_k, \tau_2) f_1(y_k) = W_2(x), \quad x \in \omega. \quad (4.43)$$

If we let the collection of y_k , $k = 1, 2, \dots, N$ coincide with the points x_j , $j = 1, 2, \dots, N$ this results in a square block matrix $K \in \mathbb{R}^{2N \times 2N}$ consisting of the square matrices $K_{10}, K_{11}, K_{20}, K_{21}$ with entries

$$K_{10,jk} = dy K_0(x_j, y_k, \tau_1), \quad K_{11,jk} = dy K_1(x_j, y_k, \tau_1) \quad (4.44)$$

$$K_{21,jk} = dy K_0(x_j, y_k, \tau_2), \quad K_{22,jk} = dy K_1(x_j, y_k, \tau_2), \quad (4.45)$$

such that

$$K = \begin{bmatrix} K_{10} & K_{11} \\ K_{20} & K_{21} \end{bmatrix}. \quad (4.46)$$

This way the system of equations (4.29), (4.30) can be written

$$K\bar{f} = \bar{W}. \quad (4.47)$$

The unknown volatility perturbation $f(y)$ is then obtained by solving the linear system (4.47) as

$$\bar{f} = K^{-1}\bar{W}, \quad (4.48)$$

where $\bar{f} = [\bar{f}_0 \bar{f}_1]^T$, $\bar{W} = [\bar{W}_1 \bar{W}_2]^T \in \mathbb{R}^{2N}$ and where the $\bar{\cdot}$ quantities are the approximations to their respective smooth counterparts. For example the entries of $\bar{f}_0 \in \mathbb{R}^N$ are $\bar{f}_{0,j} \approx f_0(x_j)$, $j = 1, \dots, N$.

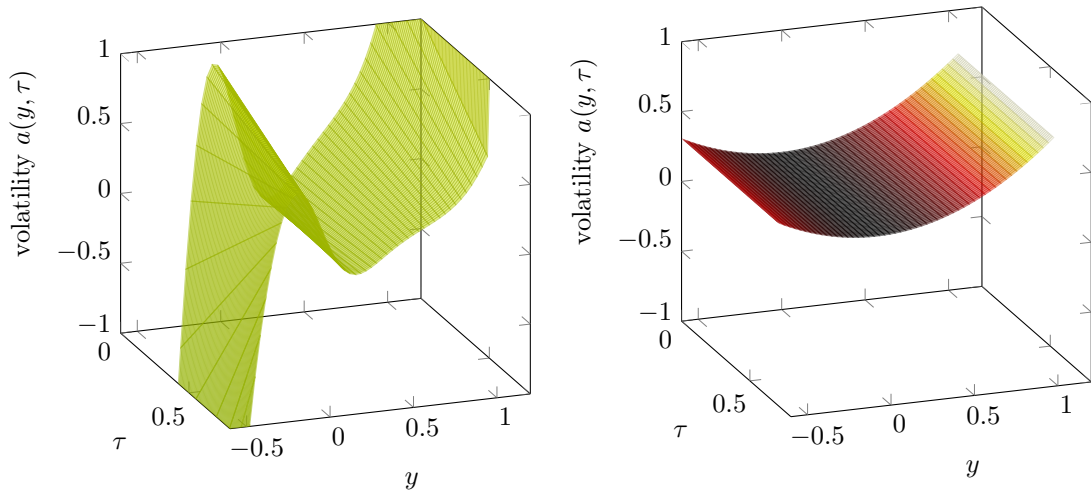


Figure 4.2: Reconstructed volatility surface  compared to prescribed volatility surface .

Numerical example

We try to recover a given volatility surface function $a(y, \tau) = \frac{1}{2}\sigma_0^2 + \frac{1}{2}y^2$. Let $s = 20$, $\sigma_0 = \frac{1}{2}$, $y \in (-1, 1.5)$. The equation (4.12) is solved backwards with volatility σ_0 and $a(y, \tau)$ to obtain $V_j(y)$ and thereafter the Fredholm system (4.47) is solved. The results are presented in Fig. 4.2. The quality of the reconstruction is not satisfactory and the reason for this behaviour are not clear at the moment.

4.2.6 Towards a real world scenario: Smooth interpolation of market prices and remarks

Since $u_j, j = 1, 2$ has to be at least twice differentiable on ω in order for f to be continuous the real market prices $u(\tau_j, y), j = 1, 2$ which are only sparsely available, i.e. for certain combinations of maturity T and strike price K , have to be smoothly approximated and or interpolated. For example one could use smoothing splines to achieve this. Given n data (x_i, y_i) with $x_1 < x_2 < x_3 < \dots < x_n$ such that $y_i = \mu(x_i)$. The smoothing spline approximation $\bar{\mu}$ of μ is the minimizer of

$$\sum_{i=1}^n (y_i - \bar{\mu}(x_i))^2 + \eta \int_{x_1}^{x_n} \left(\frac{d^2 \bar{\mu}(x)}{dx^2} \right)^2 dx. \quad (4.49)$$

Here, $\eta > 0$ is a smoothing parameter regulating the trade-off between a good data fit and a smooth approximation. In fact, for $\eta \rightarrow 0$, the solution will converge towards the interpolating spline, whereas for $\eta \rightarrow \infty$ the solution will get smoother and converge towards the linear least squares fit. A small test on some prices of put options on the MSFT stock suggests that this method is suitable for this purpose. The result which can be seen in Fig. 4.3 has been generated using Matlab's Curve Fitting Toolbox and the suggested smoothing parameter in this case is $\eta = 0.3738$.

In the context of the inverse problem of section 3.2 one could take the corresponding option's strike price and maturity date and reconstruct the volatility as follows. For the choice of $I^* = \{T_1, T_2\}$ there are two possibilities: either take the options maturity as T_1 and take the next biggest contract duration as T_2 or take T_2 as the option's maturity and T_1 as the next smallest maturity. Regarding the interval ω ; it should contain today's stock price s^* . One approach would be to take all strikes in a range of $\pm 30\%$ around the stock price s^* . This choice reflects the observed reconstruction quality of the volatility which deteriorates for options that are deep in or out of the money.

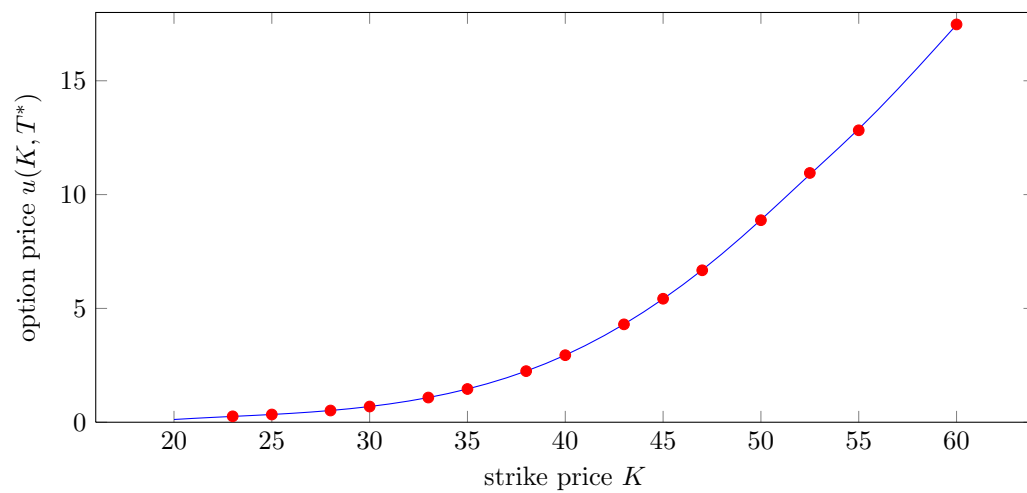


Figure 4.3: Price of put option for different strikes \bullet on MSFT stock with maturity date 17.06.2016 as observed on 14.09.2015 with smoothing spline --- .

Chapter 5

Putting everything into action: A large-scale real market data study

5.1 Option exchanges & BATS

The data was collected via Yahoo Finance <http://finance.yahoo.com/options> which retrieves its options data from the BATS exchange which had a market share in the American options market of 9.9% in March 2015 [1]. The biggest American options exchange is the Chicago Board of Options (CBOE) with a option trading volume of more than \$561 billion in 2013 [2]. Regular trading hours on the American exchanges are generally from 0930 ET to 1615 ET. There are also extended trading hours from 0030 ET until 0930 ET and from 1615 ET until 0915 ET +1d. Within the European market Deutsche Börse owned Eurex Exchange is the most important exchange.

5.2 Data gathering

For information on how to use the Yahoo Finance API visit: http://www.jarloo.com/yahoo_finance. To retrieve the option data we used the Remote Data Access functionality of the Python Data Analysis Library *pandas*: <http://www.pandas.pydata.org>. Quite some effort went into implementing this multi-threaded data grabber incorporating delays to synchronize real-time stock quotes with delayed option quotes and creating a MongoDB <http://www.mongodb.org> database to store the collected data. Data was collected in one hour intervals starting from 0930 ET during regular trading hours. This means that 7 data sets were collected each day starting from 1st July 2015 until 4th August 2015. The information is stored in CSV format and one such data set has a file size of around 40MB. A list of 3746 symbols for which option data was collected can be downloaded from http://s000.tinyupload.com/?file_id=08865665819161381532.

5.3 Data filtering

We collect data in the following format:

Strike	Expiry	Type	Symbol	Last
Bid	Ask	Chg	PctChg	Vol
Open_Int	IV	Root	IsNonstandard	Underlying
Underlying_Price	Quote_Time	Underlying_Symbol	Underlying_Bid	Underlying_Ask

In order to obtain meaningful results the raw data has to be filtered first. Only options with sufficient trade volume were taken into account. Next, for backtesting reasons option data that were not available for five consecutive days as well as zero entries are removed from the data set and a mapping of dates to consecutive trading days with results being stored in a vector called *TV* is appended. Option with identical last price or volume on two consecutive days were also filtered because this is interpreted as a sign of inactivity. Additionally, a vector containing the time to expiry in years named *TTE* is also appended. Eventually the datasets (per option) which are imported to Matlab contain the following data:

Strike	Bid	Ask	Implied volatility	Underlying_Bid	Underlying_Ask	<i>TTE</i>	<i>TV</i>
--------	-----	-----	--------------------	----------------	----------------	------------	-----------

Assume there are n_t days for which data is available for option *XYZ*. The dataset *XYZ* for option *XYZ* will then be of dimension $XYZ \in \mathbb{R}^{8 \times n_t}$. For the following computations only every tenth

entry of the database was considered to keep calculation times manageable on a consumer class computer.

5.4 Trading strategy

Remember that this model does not distinguish between call and put options since this information is usually incorporated by means of the boundary conditions. However, we solve the (inverse time) Black-Scholes equation on the new interval (s_b, s_a) . For trading decisions the calculated price of the option is evaluated at the midpoint between the ask and bid price of the underlying $s_m = \frac{s_b + s_a}{2}$. Since the bid-ask spread is narrow trades are only executed if our calculated price $u_\alpha(s_m, \tau)$ surmounts the so-called decision criterion $d(t)$ by an amount of $b = 0.02\$$ at least. Since we are calculating $u_\alpha(s_m, t)$ for $t \in (0, 2\tau)$, the trading strategy is a two time unit strategy. This means that depending on the prices $u_\alpha(s_m, \tau)$, $u_\alpha(s_m, 2\tau)$ we execute one of the strategies listed below. Remember that τ can be one minute, one hour, one day or every fraction of 255 days between 0 and 1/4.

1. If $u_\alpha(s_m, \tau) > d(\tau) + b$ and $u_\alpha(s_m, 2\tau) > d(2\tau) + b$, buy two contracts of the option today, i.e. at $t = 0$. Then, sell one contract at $t = \tau$ and the second one at $t = 2\tau$. Apply the strategy after two time units, i.e. at $t = 2\tau$ again.
2. If $u_\alpha(s_m, \tau) > d(\tau) + b$, but $u_\alpha(s_m, 2\tau) \leq d(2\tau) + b$, buy one contract today, i.e. at $t = 0$. Sell the contract after one time unit, i.e. at $t = \tau$ and apply the strategy on that day again.
3. If $u_\alpha(s_m, \tau) \leq d(\tau) + b$, but $u_\alpha(s_m, 2\tau) > d(2\tau) + b$, buy one contract today, i.e. at $t = 0$ and sell it after two time units, i.e. at $t = 2\tau$.
4. If $u_\alpha(s_m, \tau) \leq d(\tau) + b$ and $u_\alpha(s_m, 2\tau) \leq d(2\tau) + b$, do not buy anything and apply the strategy after time τ has passed.

For now we use the extrapolated ask price $u_a(t)$ as the decision criterion, i.e. $d(t) = u_a(t)$. We mention this because we will choose different combinations of boundary conditions and decision criteria in later sections.

5.5 Backtesting

- Options are usually traded by means of option contracts which is written on 100 shares of the underlying.
- Trading is possible irrespective of the account balance \rightarrow money can be borrowed at all times and without costs.
- Transaction costs are not considered, but one might argue that the buffer b accounts for them.
- Interest rates are not taken into account, i.e. $r \equiv 0$.
- Data is required to be available for 5 consecutive days in order to back test. In a real world application of the algorithm more decisions would be made.

Since our trading strategy only incorporates trades with holding periods of one and two day(s), we store profits respectively losses in the matrices X_1 and X_2 . For n_t days of observable data we have $n_t - 2$ trading opportunities for one day holding periods and $n_t - 2 - 1$ for two day holdings. This is due to the fact that we need two days of history data to apply our strategy. If n_o options are considered the matrices will have the following dimensions: $X_1 \in \mathbb{R}^{n_o \times n_t - 2}$ and $X_2 \in \mathbb{R}^{n_o \times n_t - 2 - 1}$. For implementation reasons we append a zero column vector $z = 0 \in \mathbb{R}^{n_o \times 1}$ to X_2 to make its dimension coincide with the dimension of X_1 . I.e. X_2 is redefined as $X_2 = [X_2 \ z]$.

To clarify matters, assume there is a list of options labeled $1, 2, \dots, j, \dots, n_o$. Now entry $X_1(i, j)$ contains the change in price of the i th option from the j th day to the $j + 1$ th day of observation i.e. $X_1(i, j) = (u_b(j+1) + u_a(j+1))/2 - u_a(j)$. Note that the ask price on the day of buying $u_a(j)$ is quoted whereas the midpoint between bid and ask price on the day of selling is quoted under the assumption that we are able to trade within the bid ask range on average. Next an indicator matrix $I \in \mathbb{R}^{n_o \times n_t - 2}$ is defined. Whenever the algorithm suggests a strategy for a certain option the corresponding number will be written into I at the corresponding position. The labeling is as follows: strategy 1 $\hat{=}$ 1, strategy 2 $\hat{=}$ 2, strategy 3 $\hat{=}$ 3, no trade $\hat{=}$ 0. In a next step this indicator matrix is split into three separate matrices of the same dimension, such that I_1 contains the non zero entries of I which equal one and I_2, I_3 the non zero entries which equal 2 and 3 respectively. These indicator matrices allow us to pick the daily/two day profit/loss from X_1 and X_2 for the days on which the algorithm suggested trading. This is a comfortable way to calculate profit per option or per strategy. Consider for example the profits (or losses) generated by strategy 1. Since this strategy includes two holding periods the total profit is a combination of entries in X_1 and X_2 . More precisely, let X_{1d1} and X_{2d1} be the entries of X_1 and X_2 respectively whose positions coincide with the non zero entries of I_1 . The subscript $1d1$ encodes a holding period of one day ($1d$) for strategy 1 (1). For example, one obtains the profit per option for strategy 1 by summing along the second dimension of X_1 . The account balance as plotted in the figures across the following sections is the sum of all options' profit/losses for all strategies combined, such that balance with initial balance b_0 on the i th trading day is

$$b(i) = b_0 + \sum_{a=1}^{n_o} \sum_{b=1}^{n_t-2} X_{1d1}^i(a, b) + X_{2d1}^i(a, b) + X_{1d2}^i(a, b) + X_{2d3}^i(a, b), \quad i = 1, 2, \dots, n_t - 2. \quad (5.1)$$

5.5.1 Restriction on options with significant bid-ask spread: A Comparison

As witnessed and explained in the previous section, the information content for options whose underlyings have very small bid-ask spreads is very small. To mitigate this we restrict our analysis to those options with significant bid-ask spreads and investigate possible potential for profit among this data subset. Let the bid-ask spread of the underlying be defined as $\Delta = s_a - s_b$. Such a data subset contains only options for which the relative spread $\frac{\Delta}{(s_a + s_b)/2} > \varepsilon$, where ε is the relative spread threshold. We demand a minimum relative spread of $\varepsilon = 0.6\%$ for an option to be considered "an option with significant bid ask spread". The testing procedure is as follows. The dataset containing all options is referred to as dataset 1. The dataset containing the options with significant bid ask spread is referred to as dataset 2. We take 40 random samples out of each of the datasets such that the number of decisions made by the algorithm is roughly the same. Then we simulate the trading strategy on each of the datasets' samples and calculate mean values of characteristic quantities q , which we denote by $\emptyset q$. The initial account balance is $b_0 = 100\text{\$}$ and it is assumed that money can be borrowed at no costs at any time. A comparison of back testing results can be seen in Table 5.1. The algorithm triggers fewer trades for dataset 2 than for dataset 1. This results in significantly higher end balance $\emptyset b(n_t - 2)$: 94.93\\$ vs. 88.68\\$. Note however, that the ratio between profitable and losing trades is still the same.

5.5.2 Modification: Negativity of extrapolated ask price

In case the extension of the quadratic polynomial $u_a(t)$ onto $(0, 2\tau)$ is negative at any point $t \in (0, 2\tau)$, no trade is recommended. The reason therefore is that the extrapolated ask price is a boundary condition for the inverse problem and even though the solution of such a problem could still contain information about the evolution of the price when compared to the decision criterion u_a , we do feel like it is not reasonable to allow for negative option prices. The results, which can be seen in Table 5.2 suggest improved results for both datasets. Especially, the increase in performance for dataset 1 is remarkable when compared with the data from Table 5.1 ($\Delta \emptyset b(n_t - 2) = +6.93\text{\$}$). Therefore, all the computations to follow respect this modified trading rule.

Measure	Dataset 1	Dataset 2
$\emptyset b(n_t - 2)$	88.68\$	94.93\$
$\emptyset \#$ decisions made	114	112
$\emptyset \#$ profitable trades	13	8
$\emptyset \#$ losing trades	40	27

Table 5.1: Comparison between samples from all options versus samples from options with significant bid-ask spread.

Measure	Dataset 1	Dataset 2
$\emptyset b(n_t - 2)$	95.61\$	96.65\$
$\emptyset \#$ decisions made	108	115
$\emptyset \#$ profitable trades	10	5
$\emptyset \#$ losing trades	30	21

Table 5.2: Effects of incorporating the modified trading rule mentioned in Section 5.5.2.

5.5.3 Impact of boundary conditions and decision criterion

From the boundary conditions from Section 3.4.3 we chose to test the regression boundary conditions and back test it. Furthermore, we also investigate the effects of using different combinations of boundary conditions and decision criteria. The abbreviations "bc" and "dc" stand for the boundary conditions and decision criterion used, respectively. In this context, "quad" refers to the quadratic polynomial and "reg" refers to the case where regression is applied on to the respective data for use as boundary condition or decision criterion. See Table 5.3 for the results. An important observation is that the regression boundary conditions and regression as the decision criterion lead to way lower number of trades. It might be interesting to further investigate the behavior of this strategy for longer observation horizon.

5.5.4 Influence of volatility coefficient

Another important quantity is the volatility coefficient $\sigma(t)$. The quadratic extrapolation mentioned in Section 3.2 is compared with a linear regression function obtained from the volatility observations at times $\{-2\tau, \tau, 0\}$. The impact can be seen in Table 5.4. Fewer trades are recommended when compared to the cases with the volatility which is quadratic in time (see Table 5.2).

5.5.5 Test on most actively traded options

One expects the prediction quality to be better for actively traded, i.e. liquid options. On the other hand tight bid ask spreads are associated with these options. In dataset 3 we collect the 56 most active options and run our strategy on this subset. The results are shown in Table 5.5.

5.5.6 Alternative one-day trading strategy

As prediction quality theoretically decreases for longer prediction horizons, we test a one-day trading strategy. The trading rule is as follows:

1. If $u_\alpha(s_m, \tau) > d(\tau) + b$, buy one contract today, i.e. at $t = 0$. Sell the contract after one time unit, i.e. at $t = \tau$ and apply the strategy on that day again.
2. If $u_\alpha(s_m, \tau) \leq d(\tau) + b$, do not buy anything and apply the strategy after time τ has passed.

We use $d(t) = u_a(t)$. In Table 5.6 the results of this strategy can be seen for all samples from all options (dataset 1), for options with significant bid-ask spread (dataset 2) and for most actively traded options.

Measure	Dataset 1	Dataset 2
bc: quad, dc: regression		
$\emptyset b(n_t - 2)$	89.73\$	93.96\$
$\emptyset \#$ decisions made	121	113
$\emptyset \#$ profitable trades	19	9
$\emptyset \#$ losing trades	52	32
bc: regression, dc: quad		
$\emptyset b(n_t - 2)$	91.18\$	92.85\$
$\emptyset \#$ decisions made	122	112
$\emptyset \#$ profitable trades	14	7
$\emptyset \#$ losing trades	54	35
bc: regression, dc: regression		
$\emptyset b(n_t - 2)$	98.47\$	98.99\$
$\emptyset \#$ decisions made	88	115
$\emptyset \#$ profitable trades	2	1
$\emptyset \#$ losing trades	4	1

Table 5.3: Impact of different combinations of boundary conditions and decision criteria as presented in Section 3.4.3.

Measure	Dataset 1	Dataset 2
bc: quad, dc: quad		
$\emptyset b(n_t - 2)$	91.46\$	96.74\$
$\emptyset \#$ decisions made	109	113
$\emptyset \#$ profitable trades	10	5
$\emptyset \#$ losing trades	30	21
bc: quad, dc: reg		
$\emptyset b(n_t - 2)$	93.22\$	93.96\$
$\emptyset \#$ decisions made	128	122
$\emptyset \#$ profitable trades	20	9
$\emptyset \#$ losing trades	55	36

Table 5.4: Experiment: Linear regression function as volatility.

Measure	Dataset 3
bc: quad, dc: quad	
$\emptyset b(n_t - 2)$	97.40\$
$\emptyset \#$ decisions made	100
$\emptyset \#$ profitable trades	7
$\emptyset \#$ losing trades	12

Table 5.5: Experiment: Most actively traded options.

Measure	Dataset 1	Dataset 2	Dataset 3
bc: quad, dc: quad			
$\emptyset b(n_t - 2)$	97.75\$	99.38\$	99.13\$
$\emptyset \#$ decisions made	98	115	100
$\emptyset \#$ profitable trades	5	3	3
$\emptyset \#$ losing trades	11	6	7

Table 5.6: Experiment: One-day trading strategy.

5.5.7 Forecasting with locally reconstructed implied volatility: Future work

Note that we are leaving the inverse Black-Scholes framework presented in Section 3 due to the fact that the volatility coefficient now depends both on time t and stock price s . Since the retrieval of volatility in Isakov [11] is not based on the same model as in Klivanov [29], in this case the tests presented here are of experimental nature only. However, the authors of [29] believe that based on the theory of globally convergent algorithms which was successfully applied in Klivanov [28, 26] it should be possible to construct functionals such that convergence can be proved. The key idea is to modify the Tikhonov-like functional to become globally strictly convex.

Since the results for the volatility reconstruction in Chapter 4 are not satisfactory, we refrain from providing further commentary here.

Chapter 6

Conclusions

What has been done

The study was set out to investigate the possibility of forecasting options prices via the solution of the time inverse Black-Scholes equation. To this end, fundamental terminology and basic concepts regarding the derivative form of options were introduced. Since the problem we are dealing with is ill-posed in the sense of Hadamard some basic theory on the treatment and analysis of such problems was presented in Section 3.3. The regularization technique applied here is Tikhonov regularization. We cited a convergence proof of the minimizer of the Tikhonov functional to the exact solution of the time inverse problem and analyzed different parameter selection criteria for their applicability to the inverse time option pricing problem. The analogy to the heat equation served as an accessible example to point out some of the particularities one encounters in the context of (ill-posed) inverse problems. A numerical example confirmed the theoretic result of solving the backward time heat equation for up to $T = 0.5s$, which is a remarkable result considering the instantaneous failure of standard Euler scheme like solvers for this problem.

One aspect of this thesis was to test the conjecture of profitability of the reconstruction algorithm when applied with a simple trading strategy to real market data. To do so, data was collected from Yahoo Finance and accumulated using Python's pandas library. The queried data was imported into a MongoDB database, filtered using the statistical computing software R and formatted for use with Matlab. Finally, backtesting was done using different boundary conditions and decision criteria and a two day strategy was compared to a one day strategy. The number of options considered (samples taken from a database of roughly 300.000 different options) was significantly higher than the 20 options tested in the original paper by Klibanov [29].

What the results mean

The reconstruction algorithm works when a reasonable regularization parameter is chosen. However, for the computational domain we are interested in, it is difficult to find automated parameter selection procedures. At the moment, good heuristic approaches seem to be the best way of choosing such parameters in this case. The results obtained in Sections 5.5.1-5.5.4 should not be seen in terms of absolute return but as insight into which changes the boundary conditions, trading rules and decision criteria have on the overall performance of the algorithm. Since the test sets only comprised one month's worth of data, prolonged observations are imperative for a final assessment of the strategies performance. The results support our conjecture of the prediction to be rather pointless for options whose underlying's bid-ask spreads are too tight. In that case the prediction generated by the solution of the time inverse problem essentially corresponds with the interpolation function $F(s, t)$. One basically tries to predict price by extrapolation of a polynomial in that case and the solution of the inverse problem is superfluous.

What can be done in future work

As already mentioned before, the algorithm should be tested for longer periods of time than one month, preferably more than one year to observe long term behavior of the algorithm. Additionally, the data filtering can probably improved based on heuristics and experience with option trading. Furthermore, in some cases better results were achieved when weighting the regularization terms in (3.67) differently. In this context it might be interesting to have a look at multi parameter Tikhonov regularization techniques. Note however that the convergence results from Section 3.4 are not valid in this case.

Next, the performance of the algorithm for shorter time intervals, for example hourly trading could be investigated. In theory, results should be better.

Additionally, improved volatility coefficients could further improve the reliability of the prediction. Unfortunately, numerical results for the linearization approach presented in Chapter 4 were not satisfactory and the volatility therefore not incorporated into the model.

Also, we observed that the largest jumps in option price usually occur on the day of a company's earning report. This might seem obvious, but one could try to map estimated earning call dates to stock symbols and pause the algorithm on these days. This is also in line with the fact that the Black-Scholes framework is per assumption not able to predict or account for sudden market events.

Another possibility might be to come up with more advanced trading strategies that make use of the information obtained by the solution of the inverse problem in a better way than our rather primitive strategy. In this context, one could introduce confidence levels for certain symbols or options based on experience, i.e. past decisions. Finally, machine learning techniques might be applied therefor, as well as for the selection of regularization parameters.

Bibliography

- [1] BATS Press Release April 2015. http://cdn.batstrading.com/resources/press_releases/BATSOOptions_April2015_Volume_FINAL.pdf. Accessed: 2015-08-11. 39
- [2] CBOE Market Statistics. <http://www.cboe.com/data/marketstats-2013.pdf>. Accessed: 2015-08-11. 39
- [3] Pablo Amster, Corina Averbuj, Maria Mariani, and Diego Rial. A Black–Scholes option pricing model with transaction costs. *Journal of Mathematical Analysis and Applications*, (2):688–695, 2005. 8
- [4] John Au. Chapter 3 Inverse Problems. *An Ab Initio Approach to the Inverse Problem-Based Design of Photonic Bandgap Devices*, pages 20–38, 2007. 1
- [5] Marco Avellaneda, Craig Friedman, Richard Holmes, and Dominick Samperi. Calibrating volatility surfaces via relative-entropy minimization. *Applied Mathematical Finance*, 4(1):37–64, 1997. 32
- [6] Christopher Baker, Leslie Fox, David Mayers, and K Wright. Numerical solution of Fredholm integral equations of first kind. *The Computer Journal*, 7(2):141–148, 1964. 36
- [7] Anatolii Bakushinskii. Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion. *USSR Computational Mathematics and Mathematical Physics*, 24(4):181–182, 1984. 13, 26
- [8] Fischer Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973. 5
- [9] Ilia Bouchouev. Derivatives valuation for general diffusion processes. In *The International Association of Financial Engineers (IAFE) Conferences*, volume 292, 1998. 32
- [10] Ilia Bouchouev and Victor Isakov. Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets. *Inverse Problems*, 15(3):R95–R116, 1999. 34
- [11] Ilia Bouchouev, Victor Isakov, and Nicolas Valdivia. Recovery of volatility coefficient by linearization. *Quantitative Finance*, 2:257–263, 2002. iii, 32, 44
- [12] Julianne Chung, Malena I Español, and Tuan Nguyen. Optimal Regularization Parameters for General-Form Tikhonov Regularization. pages 1–21, 2014. 27, 28
- [13] Timothy A Davis. Algorithm 915, SuiteSparseQR. *ACM Transactions on Mathematical Software*, 38(1):1–22, 2011. 22
- [14] Bruno Dupire et al. Pricing with a smile. *Risk*, 7(1):18–20, 1994. 32, 33
- [15] Franklin R Edwards. Hedge funds and the collapse of long-term capital management. *Journal of Economic Perspectives*, 13:189–210, 1999. 8

- [16] William N Goetzmann and Geert K Rouwenhorst. *The origins of value: The financial innovations that created modern capital markets*. Oxford University Press, 2005. 2, 3
- [17] Gene Golub and William Kahan. Calculating the singular values and pseudo-inverse of a matrix. *Journal of the Society for Industrial & Applied Mathematics, Series B: Numerical Analysis*, 2(2):205–224, 1965. 21
- [18] Jacques Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, 13(49-52):28, 1902. 10
- [19] Per C Hansen. *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*, volume 4. Siam, 1998. 27
- [20] Espen G Haug and Nassim N Taleb. Option traders use (very) sophisticated heuristics, never the Black-Scholes-Merton formula. *Journal of Economic Behavior and Organization*, 77(2):97–106, 2011. 3, 5, 8
- [21] Victor Isakov. Recovery of volatility coefficient by linearization, 2013. iii, 32, 34
- [22] Kazufumi Ito, Bangti Jin, and Tomoya Takeuchi. Multi-parameter Tikhonov regularization. *arXiv preprint arXiv:1102.1173*, 2011. 28
- [23] Steven L Jones and Jeffrey M Netter. Efficient capital markets. the concise encyclopedia of economics. 2008. library of economics and liberty, 2008. [Online; accessed 31-July-2015]. 1
- [24] Joseph P Kairys and Nicholas Valerio. The market for equity options in the 1870s. *The Journal of Finance*, 52(4):1707–1723, 1997. 3
- [25] Ioannis Karatzas. Brownian motion and stochastic calculus (graduate texts in mathematics). page 150, 2001. 6
- [26] Michael V Klivanov. Recovering of dielectric constants of explosives via a globally strictly convex cost functional. pages 1–22, 2014. 32, 44
- [27] Michael V Klivanov. Carleman estimates for the regularization of ill-posed Cauchy problems. *Applied Numerical Mathematics*, 94:46–74, 2015. 9, 13, 14
- [28] Michael V Klivanov and Vladimir G Kamburg. Globally strictly convex cost functional for an inverse parabolic problem. pages 1–10, 2010. 32, 44
- [29] Michael V Klivanov and Andrey V Kuzhuget. Profitable forecast of prices of stock options on real market data via the solution of an ill-posed problem for the Black-Scholes equation. *Preprint*, pages 1–13, 2015. iii, 1, 9, 11, 12, 28, 44, 45
- [30] Michael V Klivanov and Alexander V Tikhonravov. Estimates of initial conditions of parabolic equations and inequalities in infinite domains via lateral Cauchy data. *Journal of Differential Equations*, 237(1):198–224, 2007. 14
- [31] Michael V Klivanov and Aleksa Timonov. Carleman Estimates for Coefficient Inverse Problems and Numerical Applications. 2004. 14, 32
- [32] Ronald Lagnado and Stanley Osher. A technique for calibrating derivative security pricing models: numerical solution of an inverse problem. *Journal of computational finance*, 1(1):13–25, 1997. 32
- [33] Minqiang Li. You Don t Have to Bother Newton for Implied Volatility. *Available at SSRN 952727*, 2006. 32
- [34] Beth Marie, Campbell Hetrick, Rhonda Hughes, and Emily McNabb. Regularization of the backward heat equation via heatlets. *Electronic Journal of Differential Equations*, 2008(130):1–8, 2008. 11

-
- [35] Scott Mixon. The foreign exchange option market, 1917-1921. *Available at SSRN 1333442*, 2009. 3
- [36] Lyndon Moore and Steve Juh. Derivative pricing 60 years before black-scholes: Evidence from the Johannesburg stock exchange. *The Journal of Finance*, 61(6):3069–3098, 2006. 3
- [37] Christopher C Paige and Michael A Saunders. LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares. *ACM Transactions on Mathematical Software*, 8(1):43–71, 1982. 20, 21
- [38] Lothar Reichel, Fiorella Sgallari, and Qiang Ye. Tikhonov regularization based on generalized krylov subspace methods. *Applied Numerical Mathematics*, 62(9):1215–1228, 2012. 28
- [39] Christoph Reisinger. Calibration of volatility surfaces. (June), 2003. 32
- [40] Mark Rubinstein. Implied binomial trees. *The Journal of Finance*, 49(3):771–818, 1994. 31
- [41] Serap Sarikaya, Gerhard-Wilhelm Weber, and Yesim S Dogrusoz. Combination of conventional regularization methods and genetic algorithm for solving the inverse problem of electrocardiography, 2010. 1
- [42] Trevor J Saunders. *Aristotle: Politics: books I and II*. Clarendon Press, 1995. 2
- [43] Thomas I Seidman. Optimal Filtering for the Backward Heat Equation. *SIAM Journal on Numerical Analysis*, 33(1):162–170, 1996. 11
- [44] Sebastien Soulan, Maxime Besacier, Patrick Schiavone, et al. Resist trimming etch process control using dynamic scatterometry. *Microelectronic Engineering*, 86(4):1040–1042, 2009. 1
- [45] Andrey N Tikhonov. Solution of incorrectly formulated problems and the regularization method. In *Soviet Math. Dokl.*, volume 5, pages 1035–1038, 1963. 13
- [46] Andrey N Tikhonov and Vasilii Y Arsenin. *Solutions of ill-posed problems*. Vh Winston, 1977. 13
- [47] Joseph de la Vega. Confusion de confusiones. *Diálogos curiosos entre un filósofo agudo, un mercader discreto y un accionista erudito (1688)*. 2a ed. 1958, Madrid: Publicaciones Banco Urquijo, 1957. 1, 2
- [48] Anatoly G Yagola, Aleksandr S Leonov, and Vadim N Titarenko. Data Errors and an Error Estimation for Ill-Posed Problems. *Inverse Problems in Engineering*, 10(2):117–129, 2002. 11, 12, 13
- [49] Jorge P Zubelli. Inverse Problems in Finances A Short Survey of Calibration Techniques. pages 1–12, 2006. 32