

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Valuing Path-Dependent Options using the Finite Element Method, Duality Techniques, and Model Reduction

Georgios Foufas

CHALMERS |  UNIVERSITY OF GOTHENBURG



Department of Mathematical Sciences
Chalmers University of Technology and University of Göteborg
Göteborg, Sweden 2008

Valuing Path-Dependent Options using the Finite Element Method, Duality
Techniques, and Model Reduction

Georgios Foufas

ISBN 978-91-7385-112-1

©Georgios Foufas, 2008

Doktorsavhandlingar vid Chalmers Tekniska Högskola

Ny serie 2793

ISSN 0346-718X

Department of Mathematical Sciences

Chalmers University of Technology and University of Göteborg

SE-412 96 Göteborg

Sweden

Telephone +46 (0)31 772 1000

Printed in Göteborg, Sweden 2008

Valuing Path-Dependent Options using the Finite Element Method, Duality Techniques, and Model Reduction

Georgios Foufas

Department of Computational Mathematics
Chalmers University of Technology
Göteborg University

Abstract

In this thesis we develop an adaptive finite element method for pricing of several path-dependent options including barrier options, lookback options, and Asian options. The options are priced using the Black-Scholes PDE-model, and the resulting PDE:s are of parabolic type in one spatial dimension with different boundary conditions and jump conditions at monitoring dates.

The adaptive finite element method is based on piecewise polynomial approximations in space and time. We derive a posteriori estimates for the error in pointwise values of the solution and it's derivatives, using duality techniques. The estimates are used to determine suitable local resolution in space and time. The suggested adaptive finite element method is stable and gives fast and accurate results. In addition to option prices we also calculate certain sensitivity measures, or the so called Greeks, and present a new connection between some of the Greeks and the a posteriori error analysis.

We also develop an a posteriori error analysis for different SVD based model reduction techniques, and present a new model reduction technique. These techniques enables us to reduce the size of the problem, which radically improves the performance. The a posteriori error estimates are again derived using duality techniques. The model reduction techniques are tested on European and Asian options.

Keywords: finite element method, Galerkin, duality, a posteriori error estimation, adaptivity, option pricing, Greeks, Brownian motion, European option, barrier option, lookback option, Asian option, average option, POD, model reduction, balanced truncation

This thesis consists of the following papers:

- **Paper I:** *Valuing European, Barrier, and Lookback Options using the Finite Element Method and Duality Techniques*, (submitted)
- **Paper I:** *Valuing Fixed Strike Lookback Options using the Finite Element Method and Duality Techniques*, (submitted) (with Mats G. Larson)
- **Paper III:** *Valuing Asian Options using the Finite Element Method and Duality Techniques*, *Journal of Computational Finance and Applied Mathematics* (2007)(with Mats G. Larson)
- **Paper IV:** *A Note on the Connection Between the Greeks and A Posteriori Error Analysis*, (submitted) (with Mats G. Larson)
- **Paper V:** *A Posteriori Error Analysis of Weighted POD*, (submitted) (with Mats G. Larson)
- **Paper VI:** *Model Reduction in Option Pricing using Weighted POD*, (submitted) (with Mats G. Larson)
- **Paper VII:** *Option Manager: A Software Package for Calculating and Visualizing Exotic Option Prices and Greeks*

Contents

1	Introduction	1
1.1	Thesis Objectives	2
1.2	Main Results	2
1.3	Future Work	3
2	A Brief Introduction to Option Pricing	4
2.1	Background	5
2.2	Underlying Theory	5
2.3	Derivation of the Black-Scholes Formula	7
2.4	General Derivate Valuation Formula	10
2.5	Hedging and the Greeks	10
2.6	Dividends	11
3	The Finite Element Method	12
3.1	The Weak Form	13
3.2	Finite Element Approximation	14
3.3	Matrix Equations	15
3.4	Error Estimation	17
4	Model reduction	18
4.1	POD	19
5	Summary of Papers	20
5.1	Paper I	20
5.2	Paper II	20
5.3	Paper III	20
5.4	Paper IV	21
5.5	Paper V	21
5.6	Paper VI	22
5.7	Paper VII	22

Acknowledgements

First, and most of all, I would like to thank my supervisor Prof. Mats Larson at Umeå University for suggesting interesting topics, help with mathematical problems, reviewing the thesis, and for the many giving and inspiring discussions. His encouragement and help has been invaluable, already from the start when he was my advisor during the work with my master's thesis. For his financial support for the trips to Umeå, for the many nice dinners there, and for the interesting work related and especially non-work related discussions, I will always be grateful.

For giving me the opportunity of working as a PhD student I am indebted and thankful to Prof. Claes Johnson and Prof. Mats Larson. I would also like to thank Claes for the wonderful times at Tjörn and the many nice dinners and discussions at his home.

I am also thankful to my former supervisor during the work with my master thesis, Prof. Christer Borell, for introducing me to the area of financial mathematics.

For computer assistance, helping me with C++ matters, and the many enlightening discussions, I would like to thank my former colleague Anders Logg. I would also like to thank my former colleague and roommate at Asecoshuset, Axel Målqvist, for the daily discussions about both research and pleasure.

I would like to thank the Department of Mathematical Sciences for financial support and for providing such a great working environment. All my colleagues there have made the time so much more enjoyable. I also thank all the people at the Chalmers Finite Element Center for providing a pleasant working environment during my first years as a Phd student.

Finally I would like to thank my family and friends for their constant support during all these years.

Georgios Foufas
Göteborg, March 2008

1 Introduction

The valuation of different types of derivative contracts is very important in modern financial theory and practice. Exotic options have become very popular hedging and speculation instruments in recent years. At the same time a huge amount of literature has been devoted to the pricing and hedging of such instruments. The performance demands on the valuation process is usually very high. Many different methods have been applied to attack these problems. The demand for performance have led some to use approximations that produce closed form expressions. Others rely on numerical methods such as binomial and trinomial tree methods. Tree methods are easy to understand and can be applied to many types of problems, at the same time they sometimes work less well and they lack error analysis. Another frequently used method is Finite Difference (FD) method. In this thesis we use another method, the so called Finite Element (FE) method. The FE approach has several advantages compared to other numerical techniques such as finite differences techniques. For example, using the FE method one receives a solution in the entire domain, not only in isolated nodes as in FD codes. FE codes can also incorporate different kinds of boundary conditions in an easy way. Other important advantages of the FE technique are that it can easily deal with high curvature and irregular shapes of the computational domain. One of the most important advantages in practice is that the sensitivity measures, or the so called Greeks, can be calculated more exactly using the FE method.

Often exotic options have special features that needs to be taken into account when pricing them. Usually one is able to construct a PDE whose solution gives the price of the option. Some times one has to impose different types of restrictions to the solution when solving the PDE. Such as in the case of the discrete barrier option. In such cases dual techniques are shown to be especially powerful. In practice one is only interested in the price, and it's derivatives, in one or a few points. Using this criteria, the choice of computational mesh is based on a posteriori estimates of the error in desired quantities, which we derive using duality techniques. This makes it possible to calculate an optimal mesh for each type of option, which dramatically reduces the error without noticeably enhancing the computational effort. The technique is general and can be applied to many types of options.

The computationally most expensive phase for the FE method usually is the repeated solving of linear system of equations. Still, for high dimensional contracts (at least higher than four to five) there are no real alternatives to Monte Carlo, or Quasi Monte Carlo, simulations where much of the research is made today. In order to achieve better performance we have studied model reduction techniques. The goal of model reduction is to obtain a lower-

dimensional approximation to a high-dimensional dynamical system. There are two main sets of methods, singular value decomposition (SVD) based methods, and moment matching methods. Moment matching methods have been used in finance, but to our knowledge we are the first to try SVD based methods on option pricing problems. SVD based methods have error bounds and preserve stability, but moment matching methods have no global error bounds, and do not automatically preserve stability. In this thesis we use different model reduction techniques used in fluid and solid dynamics, such as the proper orthogonal decomposition (POD) method and our own extension of the POD method, the so called weighted POD method. These SVD based methods enables us to reduce the size of the problem, which radically improves the performance. The principal idea of dimensional model reduction is to find a small number of generalized co-ordinates in which to express the dynamics, ideally with some bounds on the truncation error. In the context of FE models this can be realized by using several linear combinations of the FE basis functions (modes or generalized coordinates) instead of the individual basis functions.

1.1 Thesis Objectives

The main objectives of the thesis are to:

- Develop a finite element method for computation of the values, and different sensitivity measures, of various path-dependent options.
- Develop an a posteriori error analysis for the different studied options and a framework for creating adaptive meshes.
- Implement the adaptive finite element method using C++, and also implement and verify the a posteriori error estimates.
- Develop an a posteriori error analysis for different SVD based model reduction techniques.
- Implement and test different SVD based model reduction techniques on option pricing problems.

1.2 Main Results

- We develop an adaptive finite element method for computation of the values, and different sensitivity measures, of European options, Asian

options, and Lookback options. The European option is used as a reference, since for the European option we have the exact solution when studying Black-Scholes equation. The method is implemented using C++. (Paper I, Paper II, and Paper III)

- We develop an a posteriori error analysis for the different studied options and a framework for creating adaptive meshes. This is accomplished by using dual techniques which are shown to be very powerful. The derivation is not always straight forward, since the dual equations can sometimes be complex and difficult to derive. The a posteriori error estimates are implemented and tested using C++. (Paper I, Paper II, and Paper III)
- We present a new connection between some of the Greeks and a finite element based a posteriori error analysis. At the same time we obtain a new way of calculating two of the Greeks. (Paper IV)
- We develop an a posteriori error analysis for the weighted POD method. The weighted POD method is an extension of the POD method which we derive studying the original derivation of the POD method. The ordinary POD method and the balanced truncation method are just special cases of the weighted POD method. (Paper V)
- We test the POD method, the balanced truncation, and the weighted POD method on European and Asian options. (Paper VI)
- We develop a software package with a graphical user interface (GUI) that gives the user the ability to easily calculate different exotic option prices and the corresponding Greeks. The software *Option Manager* is implemented in C++ with a GUI developed in *Matlab's Guide*. The program features the ability to show the option prices and Greeks graphically as evolutions in time or as a space-time plot for a specific time. (Paper VII)

1.3 Future Work

The finite element method has just recently become more popular in mathematical finance and option pricing problems. There are still many interesting

open problems left to study, some of them that we have touched on during our work.

- Extend the number of studied options and include other types, for instance American options. Some studies have been done on American options but are not included in the thesis.
- Use stochastic volatility models and better models for dividends. The Hobson-Rogers volatility model has been investigated and seems promising.
- Extend the software to two and three dimensional option pricing problems. The use of dual techniques are expected to be even more successful here. Also model reduction techniques is expected to give much better payoff in the multidimensional case.
- Study high-dimensional (higher than three) option pricing problems. This is a rather open problem today.
- If possible it would be nice to be able to connect the a posteriori error analysis to the choice of inner product in the weighted POD method.
- Study the calculation of Greeks more thoroughly. Compare different techniques and present numerical results using the connection between the Greeks and the a posteriori error analysis.
- Improve the C++ code and the software *Option Manager*.

2 A Brief Introduction to Option Pricing

This section gives a brief introduction to the theory of option pricing. A short background is presented and the mathematical model is explained, together with some useful tools for option pricing. For a more detailed discussion about option pricing we refer to Björk [2], Borell [3], or Wilmott [15].

2.1 Background

A contingent claim, or a derivative, is a contract the value of which depends on the values of other assets. One of the most common derivatives is the European call option. A European call option on a given stock with strike price K and maturity date T is the right, but not the obligation, for the holder of the option to buy one share of the stock at the price K at the time T . A European put option with strike price K and time of maturity T gives the holder the right, but not the obligation, to sell one share of the stock at the price K at maturity. The so called American option differs from the European option so that the holder can exercise the option at any time prior to the maturity date. Calls and puts are often called vanilla options.

Stocks and options have a long history. Stocks have existed for at least 750 years. Option contracts were used already during the Middle Ages. Valuing financial derivatives in a theoretical convincing way has been difficult throughout history. A very important contribution was given in 1973 when Black and Scholes presented their solution to the valuation of the European call option, based on the assumption that the stock log-price is governed by a so called Brownian motion. Their solution was based on the Itô calculus on Brownian motion. The concept arbitrage, that is risk free profit, is very central here. The most difficult part in this area is to understand the price dynamics of the underlying contracts.

Another kind of option is the exotic option with a payoff which does not just depend on its value on the maturity date, but on the history of the underlying asset price. There are many different kinds of exotic options. Some of them are easy to price and analytical pricing formulas exist, but most of them are more difficult to value. The average option, or the so called Asian option is an example of an option without a (known) closed form price formula.

2.2 Underlying Theory

Throughout this section we are working in the time interval $0 \leq t \leq T$. Let $B(t)$ denote the price of a risk free asset at time t governed by the equation $B(t) = B(0)e^{rt}$, where r is the constant interest rate. A common hypothesis about the behavior of asset prices is that they are given by geometric Brownian motions which implies that the asset prices are log-normally distributed (see e.g. Duffie [6] or Björk [2]). The price $S(t)$ of an asset at time t , solves the following stochastic differential equation

$$\begin{aligned} dS(t) &= S(t)(\mu dt + \sigma dW(t)), \\ S(0) &= S_0, \end{aligned} \quad (2.1)$$

where σ is the volatility, $\mu \in \mathbb{R}$ and $W(t)$ is a normalized Wiener process. Here σ is assumed to be a positive real number. The solution of (2.1) is

$$S(t) = S(0)e^{(\mu - \frac{\sigma^2}{2})t + \sigma W(t)}. \quad (2.2)$$

Now set

$$\tilde{W}(t) = \frac{\mu - r}{\sigma}t + W(t), \quad (2.3)$$

and note that

$$dS(t) = S(t)(r dt + \sigma d\tilde{W}(t)). \quad (2.4)$$

According to Cameron-Martin's theorem there exists another probability measure than the objective measure P , the risk neutral measure Q , such that \tilde{W} is a Q -Wiener process. The solution of (2.4) equals

$$S(t) = S(0)e^{(r - \frac{\sigma^2}{2})t + \sigma \tilde{W}(t)}, \quad (2.5)$$

and the measures P and Q are equivalent. The existence of the risk neutral measure Q assures that the market is free of arbitrage possibilities.

Because the Wiener process is not differentiable in the usual sense, the equation (2.1) is interpreted in the sense of stochastic differential calculus initiated by K. Itô. The most fundamental tool in stochastic calculus, Itô's lemma is given below. But first we state a definition. If the stochastic process $(h(t))_{0 \leq t \leq T}$ is progressively measurable and

$$\int_0^T |h(t)|^p dt < \infty \text{ almost surely,} \quad (2.6)$$

for some $p \in [1, \infty[$, then we say that h belongs to the class $L_W^p[0, T]$.

Lemma 2.1 (Itô's lemma). *Let the function $u(t, x_1, \dots, x_m)$ be two times continuously differentiable in $x_1, \dots, x_m \in \mathbb{R}$ and one time continuously differentiable in $t \in [0, T]$. Suppose we have m stochastic differentials*

$$dX_i(t) = a_i(t)dt + \sum_{k=1}^n b_{ik}(t)dW_k(t), \quad (2.7)$$

dependent on n stochastic independent Wiener Processes W_1, \dots, W_n . Let $\mathcal{F}_t = \sigma(W_1(\lambda), \dots, W_n(\lambda), \lambda \leq t)$. Let also the coefficients $a_i(t), b_{ik}(t)$ fulfil

$a_i(t) \in L^1_W[0, T]$, $b_{ik}(t) \in L^2_W[0, T]$ and so, especially, for fixed t the processes are \mathcal{F}_t -measurable. Let also $X(t) = (X_1(t), \dots, X_m(t))$. Then we have

$$\begin{aligned} du(t, X(t)) &= \frac{\partial u}{\partial t}(t, X(t))dt + \sum_{i=1}^m \frac{\partial u}{\partial x_i}(t, X(t))dX_i(t) \\ &+ \frac{1}{2} \sum_{i,j=1}^m \frac{\partial^2 u}{\partial x_i \partial x_j}(t, X(t))dX_i(t)dX_j(t). \end{aligned} \quad (2.8)$$

Note that

$$\begin{aligned} dt dt &= 0, \quad dt dW_i(t) = 0, \\ dW_i(t)dW_i(t) &= dt, \quad dW_i(t)dW_j(t) = 0 \quad \text{if } i \neq j. \end{aligned}$$

2.3 Derivation of the Black-Scholes Formula

Let $v(t, S(t))$ denote the value of the portfolio at time t , with the terminal condition $v(T, S(T)) = g(S(T))$, where the function g is piecewise continuous and fulfils

$$\sup_{x \in \mathbb{R}} (e^{-C|x|} |g(e^x)|) < \infty \quad (2.9)$$

for an appropriate constant $C > 0$. We then say that $g \in \mathcal{P}$. Suppose that the process $(v(t, S(t)))_{0 \leq t \leq T}$ is the value process of a self-financing strategy $(h_S(t), h_B(t))_{0 \leq t \leq T}$ in the stock and the risk free asset, that is

$$v(t, S(t)) = h_S(t)S(t) + h_B(t)B(t), \quad (2.10)$$

$$dv(t, S(t)) = h_S(t)dS(t) + h_B(t)dB(t). \quad (2.11)$$

By applying Ito's lemma and using (2.11) we get

$$\begin{aligned} dv(t, S(t)) &= v_t(t, S(t))dt + v_s(t, S(t))dS(t) + \frac{1}{2}v_{ss}(t, S(t))(dS(t))^2 \\ &= h_S(t)dS(t) + rh_B(t)B(t)dt. \end{aligned} \quad (2.12)$$

Identifying coefficients in (2.12) yields $h_S = v_s$. Rearranging the terms and using (2.10) we get the famous Black-Scholes differential equation

$$\begin{aligned} v_t(t, S(t)) + \frac{\sigma^2 S(t)^2}{2}v_{ss}(t, S(t)) + rS(t)v_s(t, S(t)) - rv(t, S(t)) &= 0, \quad (2.13) \\ t < T, \quad S(t) > 0. \end{aligned}$$

Together with the terminal condition $v(T, S(T)) = g(S(T))$, equation (2.13) has the following solution,

$$v(t, S(t)) = e^{-r\tau} E \left[g\left(se^{(r-\frac{\sigma^2}{2})\tau + \sigma W(\tau)}\right) \right], \quad (2.14)$$

where $s = S(t)$ and $\tau = T - t$.

Remark 2.1 Observe that (2.14) is independent of the drift coefficient μ .

We thus have the following important result.

Theorem 2.1 Let $g \in \mathcal{P}$. A simple European derivate with payoff $Y = g(S(T))$ at maturity T has the theoretical value $v(t, S(t))$ at time t , where

$$v(t, S(t)) = e^{-r\tau} E \left[g\left(se^{(r-\frac{\sigma^2}{2})\tau + \sigma W(\tau)}\right) \right], \quad (2.15)$$

and $\tau = T - t$.

We can simplify (2.15) using the risk neutral measure Q (see Geman, Karoui and Rochet [11], for a detailed discussion about changes of probability measure).

Theorem 2.2 The value $v(t, S(t))$ is equal to

$$e^{-r\tau} E^Q[g(S(T)) \mid \mathcal{F}_t].$$

Proof. According to (2.5) we have $S(T) = S(t)e^{(r-\frac{\sigma^2}{2})\tau + \sigma(\tilde{W}(T) - \tilde{W}(t))}$ and hence

$$E^Q[g(S(T)) \mid \mathcal{F}_t] = E^Q \left[g\left(S(t)e^{(r-\frac{\sigma^2}{2})\tau + \sigma(\tilde{W}(T) - \tilde{W}(t))}\right) \mid \mathcal{F}_t \right]. \quad (2.16)$$

But since $(\tilde{W}(T) - \tilde{W}(t))$ and \mathcal{F}_t are stochastic independent and \tilde{W} is a Q -Brownian motion, the right hand side of (2.16) becomes

$$E \left[g\left(se^{(r-\frac{\sigma^2}{2})\tau + \sigma(W(T) - W(t))}\right) \right]_{|s=S(t)} = e^{r\tau} v(t, S(t)),$$

□

We now state the famous Black-Scholes formula which gives the value of a European call option with payoff $Y = \max(0, S(T) - K)$ at maturity T .

Theorem 2.3 (Black-Scholes formula). *A European call option with maturity date T and strike price K has the value $c(t, S(t), K)$ at time $t < T$ where*

$$c(t, s, K) = s\Phi(d_1) - Ke^{-r\tau}\Phi(d_2), \quad (2.17)$$

$$d_1 = \frac{\ln \frac{s}{K} + (r + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}} \quad \text{and} \quad d_2 = d_1 - \sigma\sqrt{\tau},$$

and where Φ is the probability distribution function for a $N(0, 1)$ distributed stochastic variable.

Proof. Theorem 2.1 gives that

$$c(t, s, K) = e^{-r\tau} E \left[\max \left(0, se^{(r - \frac{\sigma^2}{2})\tau - \sigma\sqrt{\tau}G} - K \right) \right],$$

where $G \in N(0, 1)$. From this it follows that

$$c(t, s, K) = e^{-r\tau} E \left[se^{(r - \frac{\sigma^2}{2})\tau - \sigma\sqrt{\tau}G} - K; \quad G \leq \frac{\ln \frac{s}{K} + (r - \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}} \right]$$

$$= e^{-r\tau} \left(E \left[se^{(r - \frac{\sigma^2}{2})\tau - \sigma\sqrt{\tau}G}; \quad G \leq d_2 \right] - K\Phi(d_2) \right).$$

Here

$$e^{-r\tau} E \left[se^{(r - \frac{\sigma^2}{2})\tau - \sigma\sqrt{\tau}G}; \quad G \leq d_2 \right] = s \int_{x \leq d_2} e^{-\frac{\sigma^2}{2}\tau - \sigma\sqrt{\tau}x - \frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}$$

$$= s \int_{x \leq d_2} e^{-\frac{(\sigma\sqrt{\tau} + x)^2}{2}} \frac{dx}{\sqrt{2\pi}} = s\Phi(\sigma\sqrt{\tau} + d_2) = s\Phi(d_1),$$

which proves the theorem. \square

The price of the European put option can be derived in the same manner as the call price. Alternatively to derive the European put price one can use the so called call-put parity relation.

Theorem 2.4 (Call-put parity). *Let c and p be the value of an European call and put option respectively. Then we have*

$$p(t, s, K, T) = Ke^{-r\tau} - s + c(t, s, K, T). \quad (2.18)$$

Using Theorems 2.3 and 2.4 we can easily calculate the price of an European put option, $p(t, s, K, T)$.

$$p(t, s, K, T) = Ke^{-r\tau} - s + s\Phi(d_1) - Ke^{-r\tau}\Phi(d_2) \quad (2.19)$$

$$= Ke^{-r\tau}\Phi(-d_2) - s\Phi(-d_1).$$

2.4 General Derivate Valuation Formula

To be able to handle more complex derivatives we extend the previous valuation formula in Theorem 2 to European derivatives with payoff $X \in L^2(Q)$ and state the following theorem (for a more detailed discussion see Borell [3]).

Theorem 2.5 *A European derivative with payoff $X \in L^2(Q)$ at maturity T has the theoretical value*

$$v(t) = e^{-r\tau} E^Q[X | \mathcal{F}_t]. \quad (2.20)$$

Exotic derivatives may give rise to more complex PDE's than the ordinary European option. For example the Asian option. It can be shown, see [9], that the price of the Asian option $V = S_0 \bar{u}$ is given as the solution to the following PDE

$$\begin{aligned} \bar{u}_t + \frac{1}{2}(z - e^{-\gamma t})^2 \sigma^2 \bar{u}_{zz} &= 0, \\ \bar{u}(T, z) &= (z - K_1)^+. \end{aligned} \quad (2.21)$$

2.5 Hedging and the Greeks

Hedging is the reduction of the sensitivity of a portfolio to the movement of an underlying asset by taking opposite positions in different financial instruments. One simple way to hedge is the so called delta-hedging. With $V(S(t), t) = V(s, t)$ denoting the value of a portfolio or derivative, using Itô's lemma, (2.8), we have that

$$dV = \sigma s \frac{\partial V}{\partial s} dW + \left(\mu s \frac{\partial V}{\partial s} + \frac{1}{2} \sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} + \frac{\partial V}{\partial t} \right) dt. \quad (2.22)$$

Note that V must at least have one t derivative and two s derivatives. Let Π be a portfolio consisting of one option and $-\Delta$ number of the underlying assets,

$$\Pi = V - \Delta s. \quad (2.23)$$

Then

$$d\Pi = dV - \Delta ds, \quad (2.24)$$

which together with (2.22) and (2.1) gives that

$$d\Pi = \sigma s \left(\frac{\partial V}{\partial s} - \Delta \right) dW + \left(\mu s \frac{\partial V}{\partial s} + \frac{1}{2} \sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} + \frac{\partial V}{\partial t} - \mu \Delta ds \right) dt. \quad (2.25)$$

By choosing $\Delta = \frac{\partial V}{\partial s}$ we eliminate the randomness

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} \right) dt. \quad (2.26)$$

Delta hedging is a dynamic hedging strategy, that is, it must be continuously rebalanced to be a perfect hedge. Transaction costs makes this impossible in practice. When delta-hedging one eliminates the largest random part of the portfolio. One can also hedge away smaller effects due to, such as for instance, the curvature of the portfolio value with respect to the underlying asset. Then one needs the so called *gamma*, defined as

$$\Gamma = \frac{\partial^2 V}{\partial s^2}. \quad (2.27)$$

The decay of value in time of a portfolio is represented by the *theta*, where

$$\Theta = -\frac{\partial V}{\partial t}. \quad (2.28)$$

Sensitivity to volatility called the *vega* and is defined by

$$\frac{\partial V}{\partial \sigma}, \quad (2.29)$$

and sensitivity to interest rate is called *rho*, defined as

$$\rho = \frac{\partial V}{\partial r}. \quad (2.30)$$

The *speed* is the third derivative of V with regarding to s ,

$$\frac{\partial^3 V}{\partial s^3}. \quad (2.31)$$

2.6 Dividends

Many assets, such as equities, pay out dividends. These dividends affect the prices of options. There are several ways to model dividends. Dividends may be deterministic or stochastic, and may be made continuously or at discrete times. We will consider only deterministic dividends, whose amount and timing is known prior to the start of the option's life. This is a reasonable assumption if the options lifetime is not too long, since many companies have a similar payment from year to year. There are several ways to incorporate dividends into the Black-Sholes model. In this section we show how this is done in the simplest case, when we have a continuous and constant dividend yield. This is a good model for index options, where the many discrete

dividends can be approximated by a continuous yield without serious error. The model is also applicable to options on foreign currencies, though only for short dated options. For stocks, the dividends are often made at discrete times, and consequently this model is not suitable for stocks. For stocks the so called discrete dividend yield model or the fixed dividend model is more appropriate, see for example [12], [4], or [10]. In [14], Večer shows how to include discrete dividend payments, for the path-dependent Asian option, studied later in this thesis, in a very simple manner.

Suppose that the underlying pays out a dividend $D_0 s dt$ during the time dt , where D_0 is a constant. The dividend yield is then defined as the ratio of the dividend payment to the asset price. Thus the dividend $D_0 s dt$ represents a continuous constant dividend yield. Arbitrage considerations show that the asset price must fall the amount of the dividend payment, that is, the stock price stochastic differential equation (2.1) is modified to

$$dS(t) = S(t)((\mu - D_0)dt + \sigma dW(t)), \quad (2.32)$$

But as noted before, (2.1), the Black-Scholes equation is independent by the drift-coefficient μ in the stochastic differential equation. What changes is that we must now include the change due to dividends in our self-financing portfolio dynamics (2.11). Since we receive $D_0 S dt$ for every asset held and since we hold h_S number of the underlying, the change in value of our self-financed portfolio now reads

$$dv(t, S(t)) = h_S(t)dS(t) + h_S D_0 S(t)dt + h_B(t)dB(t). \quad (2.33)$$

The analysis proceeds exactly as before, but with new term arising from the dividend, and we find that the value of our portfolio solves the following equation

$$v_t + \frac{\sigma^2 S(t)^2}{2} v_{ss} + (r - D_0)S(t)v_s - rv = 0, \quad t < T, S(t) > 0. \quad (2.34)$$

We see that using a continuous dividend yield only corresponds to adjusting one coefficient in the partial differential equation.

3 The Finite Element Method

The finite element method is used for finding approximate solutions of partial differential equations (PDE's) as well as of integral equations. It was developed in the 1950's and 1960's by engineers, and was mainly used in structural mechanics, see e.g. [17] for an overview. The finite element method also has a strong mathematical foundation in functional analysis,

see [5]. The mathematical foundation provides the tools to derive analytical error estimates which can be used in a constructive way to improve the approximative solution.

As a model problem we choose to study the Black-Scholes equation presented in the previous Section. In order to construct a computational mesh we introduce a bounded interval $\Omega = [s_{min}, s_{max}] \subset \mathbb{R}^+$ with boundary $\partial\Omega = \{s_{min}, s_{max}\}$. We define the usual Hilbert space

$$H^1(\Omega) = \{v : \int_{\Omega} (|\nabla v|^2 + v^2) ds < \infty\}, \quad (3.1)$$

and let \mathcal{W} be the space of functions that are square integrable in time and belongs to $H^1(\Omega)$ in space, that is

$$\mathcal{W} = L^2([0, T], H^1(\Omega)). \quad (3.2)$$

We also use the notation $(u, v) = \int_{\Omega} uv ds$, and $(u, v)_{\partial\Omega} = u(s_{max})v(s_{max}) - u(s_{min})v(s_{min})$.

3.1 The Weak Form

The first step in formulating a finite element method is to rewrite the equation on weak form. Multiplying the Black-Scholes equation (2.13) by the test function $v \in \mathcal{W}$ and integrating on $\Omega \times [0, T]$ we obtain

$$\int_0^T \left((u_t, v) + (r - \nu)(su_s, v) + \frac{\sigma^2}{2} (s^2 u_{ss}, v) - r(u, v) \right) dt = 0. \quad (3.3)$$

Using integration by parts we get

$$(s^2 u_{ss}, v) = (s^2 u_s, v)_{\partial\Omega} - 2(su_s, v) - (s^2 u_s, v_s). \quad (3.4)$$

Thus equation (3.3) becomes

$$\int_0^T \left((u_t, v) + (r - \nu - \sigma^2)(su_s, v) - \frac{\sigma^2}{2} (s^2 u_s, v_s) + \frac{\sigma^2}{2} (s^2 u_s, v)_{\partial\Omega} - r(u, v) \right) dt = 0. \quad (3.5)$$

The boundary conditions for the European call option are $u(t, 0) = 0$ and $u(t, s) \sim se^{-\nu(T-t)}$ as $s \rightarrow \infty$, and for the corresponding put $u(t, 0) = Ke^{-r(T-t)}$ and $u(t, s) \sim 0$ as $s \rightarrow \infty$, see for example Wilmott, [16]. For simplicity of implementation we use the artificial boundary condition $u_{ss} = 0$ on $\partial\Omega$ for both the put and the call instead. This boundary condition works for all contracts if the payoff is at most linear in the underlying (see [16])

and does not affect the accuracy of the solution. Using equation (2.13) we can rewrite the boundary condition as

$$u_s = \frac{r}{s(r-\nu)}u - \frac{1}{s(r-\nu)}u_t, \quad (3.6)$$

and enforce it weakly by inserting identity (3.6) into equation (3.5). We thus want to solve the problem: find $u \in \mathcal{W}$ such that

$$\begin{cases} \int_0^T (m(u_t, v) + a(u, v)) dt = 0, \\ u(T, s) = \begin{cases} \max(s - K, 0), & \text{for a call,} \\ \max(K - s, 0), & \text{for a put,} \end{cases} \end{cases} \quad (3.7)$$

for every $v \in \mathcal{W}$, where

$$m(u_t, v) = (u_t, v) - \frac{\sigma^2}{2(r-\nu)}(su_t, v)_{\partial\Omega}, \quad (3.8)$$

and

$$\begin{aligned} a(u, v) &= (r - \nu - \sigma^2)(su_s, v) - \frac{\sigma^2}{2}(s^2u_s, v_s) \\ &\quad + \frac{\sigma^2 r}{2(r-\nu)}(su, v)_{\partial\Omega} - r(u, v). \end{aligned} \quad (3.9)$$

3.2 Finite Element Approximation

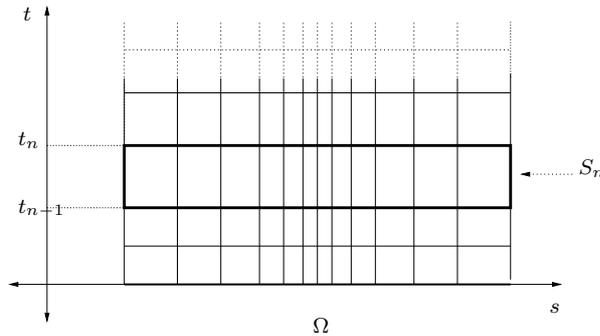
The finite element method is based on solution of the variational problem (3.7) with \mathcal{W} replaced by a finite dimensional function space of piecewise polynomials in space and time. For background on the finite element method see for instance [7].

We now partition $[0, T]$ as $0 = t_0 < t_1 < t_2 < \dots < t_N = T$, denoting each time interval by $I_n = (t_{n-1}, t_n)$ and each time step by $k_n = t_n - t_{n-1}$. Similarly we partition Ω as $s_{min} = s_0 < s_1 < s_2 < \dots < s_J = s_{max}$, denoting each spatial interval by $\kappa_j = [s_{j-1}, s_j)$ and the length of each interval by $h_j = s_j - s_{j-1}$.

In space, we let $\mathcal{V}^p \subset H^1(\Omega)$ denote the space of piecewise continuous functions of order p . On each space-time slab $S_n = I_n \times \Omega$, we define

$$\mathcal{W}_n^q = \{w(t, s) : w(t, s) = \sum_{j=0}^q t^j v_j(s), v_j \in \mathcal{V}^p, (t, s) \in S_n\}. \quad (3.10)$$

Let $\mathcal{W}^q \subset \mathcal{W}$ denote the space of functions defined on $[0, T] \times \Omega$ such that $v|_{S_n} \in \mathcal{W}_n^q$ for $1 \leq n \leq N$. For simplicity, we only give details for the continuous Galerkin method cG(p)-cG(q), (see e.g. [7] or [8]) which is defined

Figure 1: *Space-time discretization.*

by the following discrete version of equation (3.7). Find $U \in \mathcal{W}^q$ such that for $1 \leq n \leq N$

$$\begin{cases} \int_{I_n} (m(U_t, v) + a(U, v)) dt = 0 & \text{for all } v \in \mathcal{W}_n^{q-1}, \\ U^-(t_n) = U^+(t_n), & n = N - 1, \dots, 1, \\ U^-(t_N) = u_T, \end{cases} \quad (3.11)$$

where $U^\pm(t_n) = \lim_{\epsilon \rightarrow 0, \epsilon > 0} U(t_n \pm \epsilon)$. In the cG(1) method the approximation U of u is continuous piecewise linear in time and space, while the test functions v are continuous linear in space and piecewise constant in time. It is also possible to use a discontinuous method in time, we refer to [7], for details on the resulting discontinuous Galerkin method, cG(p)-dG(q). In Figure 2 we see a monthly sampled up and out barrier call option calculated using the cG(1)-dG(1) method.

3.3 Matrix Equations

We now derive the matrix equations for the case $p = q = 1$. Using the notation $U_n = U(t_n)$ and computing the time integral in equation (3.11) yields the scheme: for $1 \leq n \leq N$

$$m(U_n - U_{n-1}, v) + k_n a\left(\frac{U_n + U_{n-1}}{2}, v\right) = 0 \quad \text{for all } v \in \mathcal{W}_n^0, \quad (3.12)$$

which is the classical Crank-Nicolson method.

Let $\{\varphi_j\}_{j=0}^J$ be the standard nodal basis of \mathcal{P}_1 (see Figure 3). Then $U_n \in \mathcal{P}_1$ can be written as

$$U_n(s) = \sum_{j=0}^J \xi_{nj} \varphi_j(s), \quad 1 \leq n \leq N, \quad (3.13)$$

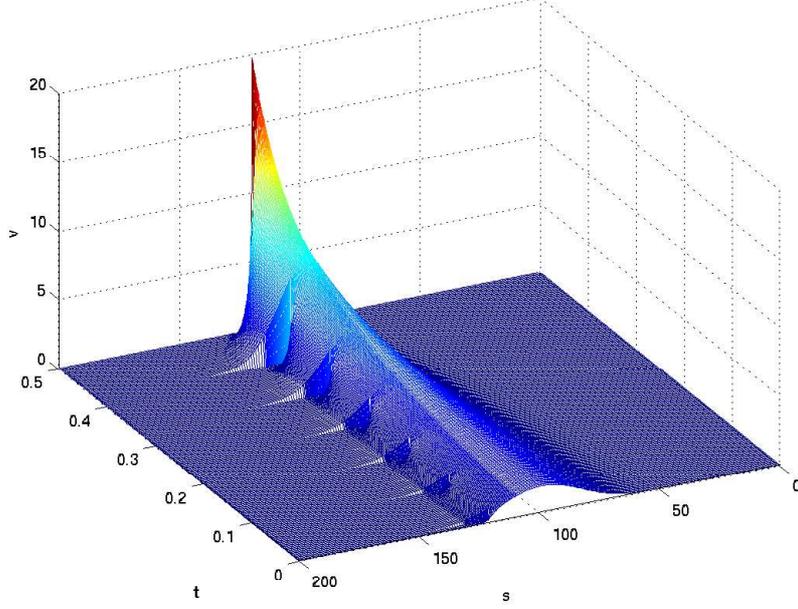


Figure 2: The monthly sampled up and out barrier call option computed using the $cG(1)$ - $dG(1)$ method with 200 time points and space points. Parameter values are $\sigma = 0.3$, $r = 0.1$, $q = 0.0$, $T = 0.5$, $t = 0.0$, $K = 100$, and $H = 120$.

and the test function v can be written as

$$v(s) = \sum_{i=0}^J \gamma_{ni} \varphi_i(s), \quad 1 \leq n \leq N, \quad (3.14)$$

for reals $\xi_{n0}, \dots, \xi_{nJ}, \gamma_{n0}, \dots, \gamma_{nJ}$.

Let now ξ_n be the vector of all $\xi_{n,j}$, $j = 0, 1, \dots, J$. If the expressions above for U and v are inserted into equation (3.12) we receive the matrix equation

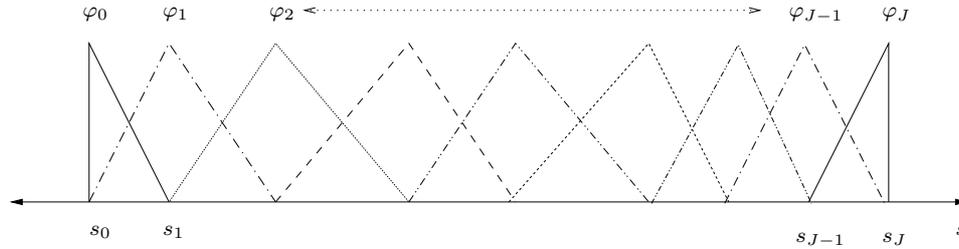
$$(\xi_n - \xi_{n-1})M + (\xi_n + \xi_{n-1})\frac{k_n A}{2} = 0, \quad 1 \leq n \leq N, \quad (3.15)$$

where

$$M = (\varphi_j, \varphi_i) - \frac{\sigma^2}{2(r - \nu)} (s\varphi_j, \varphi_i)_{\partial\Omega}, \quad 0 \leq i, j \leq J, \quad (3.16)$$

and

$$A = (r - \nu - \sigma^2)A_0 - \frac{\sigma^2}{2}A_1 - rA_2, \quad (3.17)$$

Figure 3: The hat-functions φ in the $cG(1)$ method.

where

$$\begin{aligned} A_0 &= (s\varphi_{j,s}, \varphi_i), & A_1 &= (s^2\varphi_{j,s}, \varphi_{i,s}) - \frac{r}{r-\nu}(s\varphi_j, \varphi_i)_{\partial\Omega}, \\ A_2 &= (\varphi_j, \varphi_i), & 0 \leq i, j \leq J. \end{aligned} \quad (3.18)$$

Rearranging the terms in equation (3.15) we get the matrix equation we need to solve successively backwards in time in order to obtain U_0 given U_N

$$\xi_{n-1} \left(M - \frac{k_n A}{2} \right) = \xi_n \left(M + \frac{k_n A}{2} \right), \quad 1 \leq n \leq N. \quad (3.19)$$

3.4 Error Estimation

There are two classes of finite element error estimates, *a priori* and *a posteriori*. The *a priori* estimate bounds the error $e = u - U$ in terms of data, u , and h , while a *a posteriori* estimates bounds the error in terms of data, U , and h . In this thesis we will only consider a *a posteriori* estimates since they are computable once you have calculated the solution U . We present such an estimate here of the Black-Scholes equation

$$(\psi, e(0, s)) = \int_0^T (m(U_t, \phi) + a(U, \phi)) dt, \quad (3.20)$$

where ϕ is the solution to the continuous dual problem for the Black-Scholes equation, $\psi = \phi(0, s)$, and m and a are the previously defined bilinear forms. These kind of error estimates can be used to create adaptive meshes. In Figure 4, we see a mesh resulting from using a mesh refinement algorithm based on a similar error estimate as the one above but in the case of a floating strike lookback put option with weekly sampling. The use of adapted meshes gives superior accuracy and performance with less degrees of freedom than using uniform meshes.

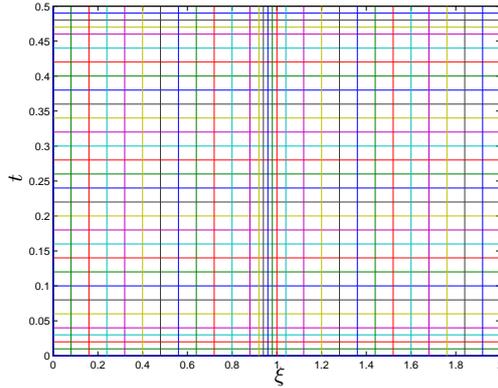


Figure 4: *The resulting mesh using a mesh refinement algorithm, calculated for a floating strike lookback put option with $T = 0.5$, $\sigma = 0.1$, and $r = 0.1$, when $\psi = \delta_1(\xi, \epsilon)$.*

4 Model reduction

Model reduction is the problem of obtaining a lower-dimensional approximation to a high-dimensional dynamical system. There are two main sets of methods, SVD based methods, and moment-matching methods. Moment matching methods have no global error bounds, and do not automatically preserve stability, whereas SVD based methods have error bounds and preserve stability. For a good survey of model reduction methods, see for example [1].

In this thesis we use the finite element method and adopt different SVD based model reduction techniques used in fluid and solid dynamics, which enables us to reduce the size of the problem, which radically improves the performance. The standard finite element basis is in some sense non-optimal, the question is what to use instead. As so elegantly described by [13], “The principal idea of dimensional model reduction is to find a small number of generalized co-ordinates in which to express the dynamics, ideally with some bounds on the truncation error”. In the context of FE models this can be realized by using several linear combinations of the FE basis functions (modes or generalized coordinates) instead of the individual basis functions. Many different generalized coordinates functions have been proposed.

4.1 POD

Here we present one of the simplest model reduction techniques used in the thesis. The other methods we use are based on the same ideas.

The idea is, given a set of data that lies in the vector space \mathcal{W} , to find a subspace \mathcal{W}_r of fixed dimension r such that the error in the projection onto the subspace is minimized. Suppose we have set of data $x(t) \in \mathbb{R}^n$, with $0 \leq t \leq T$. We then seek a projection $P_r : \mathbb{R}^n \rightarrow \mathbb{R}^n$ of fixed rank r , that minimizes the total error

$$\int_0^T \|x(t) - P_r x(t)\|^2 dt. \quad (4.1)$$

Now introduce the $n \times n$ matrix

$$R = \int_0^T x(t)x(t)^* dt, \quad (4.2)$$

where x^* denotes the transpose of x , and calculate the eigenvalues and eigenvectors of R given by

$$R\eta_k = \lambda_k \eta_k, \quad \lambda_1 \geq \dots \geq \lambda_n \geq 0. \quad (4.3)$$

Since the matrix R is symmetric, positive semidefinite, all the eigenvalues λ_k are real and non-negative, and the eigenvectors η_k may be chosen orthonormal. The main result of POD is that the optimal subspace of dimension r is spanned by $\{\eta_1, \eta_2, \dots, \eta_r\}$, and the optimal projection P_r is given by

$$P_r = \sum_{k=1}^r \eta_k \eta_k^*. \quad (4.4)$$

The vectors η_k is then used as the new basis and are called POD modes. These POD modes can then be used to form reduced order methods by applying Galerkin projection. In the thesis we discuss other ways of choosing these modes. For example incorporating dual information in the choice of the modes.

Writing the dynamics of our system as

$$\dot{x} = f(x(t)), \quad (4.5)$$

we define a new variable $x_r(t) \in \text{span}\{\eta_1, \dots, \eta_r\}$ by $\dot{x}_r(t) = P_r f(x_r(t))$. Let now

$$x_r(t) = \sum_{j=1}^r a_j(t) \eta_j. \quad (4.6)$$

Substituting this into equation (4.5) and multiplying by η_k^* we obtain

$$\dot{a}_k(t) = \eta_k^* f(x_r(t)), \quad k = 1, \dots, r, \quad (4.7)$$

a set of ODEs that describe the dynamics of $x_r(t)$.

5 Summary of Papers

5.1 Paper I

Valuing European, Barrier, and Lookback Options using the Finite Element Method and Duality Techniques, (submitted)

In this paper we develop an adaptive finite element method for computation of the values and different sensitivity measures of ordinary European options, barrier options, and lookback options.

The options are priced using the Black-Scholes PDE-model, and the resulting PDE:s are of parabolic type in one spatial dimension with different boundary conditions and jump conditions at monitoring dates. The adaptive finite element method is based on a posteriori estimates of the error in desired quantities, which we derive using duality techniques. The a posteriori error estimates are tested and also verified in the case of the European option. These estimates are then used to calculate optimal meshes for each type of option. The use of adapted meshes gives superior accuracy and performance with less degrees of freedom than using uniform meshes. The suggested adaptive finite element method is stable and gives fast and accurate results.

5.2 Paper II

Valuing Fixed Strike Lookback Options using the Finite Element Method and Duality Techniques, (submitted) (with Mats G. Larson)

In this paper is we present an adaptive finite element method for computation of the values and different sensitivity measures of fixed strike lookback options.

The fixed strike lookback options are priced using the Black-Scholes PDE-model, and a method developed by Andreasen. It consists of solving two coupled PDE:s that are of parabolic type in one spatial dimension with different boundary conditions and jump conditions at monitoring dates. The adaptive finite element method is based on a posteriori estimates of the error in desired quantities, which we derive using duality techniques. The derivation of the dual equations turns out to be a challenging problem. The a posteriori error estimates are tested and verified, and are used to calculate optimal meshes for each type of option. The use of adapted meshes gives superior accuracy and performance with less degrees of freedom than using uniform meshes. The suggested adaptive finite element method is stable and gives fast and accurate results.

5.3 Paper III

Valuing Asian Options using the Finite Element Method and Duality Tech-

niques, Journal of Computational Finance and Applied Mathematics (2007)
(with Mats G. Larson)

The Asian option is a popular and frequently traded pathdependent option which pricing problem has been studied a lot using many different techniques. The main objective of this paper is to develop an adaptive finite element method for computation of the values, and different sensitivity measures, of the Asian option with both fixed and floating strike. The pricing is based on Black-Scholes PDE-model and a method developed by Večer where the resulting PDE:s are of parabolic type in one spatial dimension and can be applied to both continuous and discrete Asian options. We propose using an adaptive finite element method which is based on a posteriori estimates of the error in desired quantities, which we derive using duality techniques. The a posteriori error estimates are tested and verified, and are used to calculate optimal meshes for each type of option. The use of adapted meshes gives superior accuracy and performance with less degrees of freedom than using uniform meshes. The suggested adaptive finite element method is stable, gives fast and accurate results, and can be applied to other types of options as well.

5.4 Paper IV

A Note on the Connection Between the Greeks and A Posteriori Error Analysis, (submitted) (with Mats G. Larson)

The sensitivity measures, also known as the Greeks, are very important tools in risk management. In this paper we present a new connection between some of the Greeks and a finite element based a posteriori error analysis. This is not only a nice feature of the a posteriori error analysis but it also gives us an alternative way of calculating the Greeks. The presented error estimation formula splits the error in parts originating from how good the numerical approximation is and in parts originating from how well the parameters are approximated. The study is based on the finite element method applied to the European option problem, but the technique is general and can be applied to other option valuation problems as well.

5.5 Paper V

A Posteriori Error Analysis of Weighted POD, (submitted) (with Mats G. Larson)

Model reduction is the problem of obtaining a lower-dimensional approximation to a high-dimensional dynamical system. The main objective of this paper is to develop an a posteriori error analysis for different model reduction techniques, such as the POD method and extensions of it. Here we use

the finite element method and adopt different SVD based model reduction techniques used in fluid and solid dynamics, which enables us to reduce the size of the problem, which radically improves the performance. The a posteriori error estimates are derived using duality techniques.

5.6 Paper VI

Model Reduction in Option Pricing using Weighted POD, (submitted) (with Mats G. Larson)

The main objective of this paper is apply different model reduction techniques, such as the POD method and a newly developed extension of it, Weighted POD, to the problem of pricing exotic options. Model reduction is the problem of obtaining a lower-dimensional approximation to a high-dimensional dynamical system. Here we use the finite element method and adopt SVD based model reduction techniques used in fluid and solid dynamics, which enables us to reduce the size of the problem, which radically improves the performance. The techniques are tested and compared on European and Asian options.

5.7 Paper VII

Option Manager: A Software Package for Calculating and Visualizing Exotic Option Prices and Greeks

In this report we present a software project that gives the user the ability to easily calculate different exotic option prices and the corresponding Greeks in a graphical user interface (GUI). The software *Option Manager* is implemented in C++ with a GUI developed in *Matlab's Guide*. The program features the ability to show the option prices and Greeks graphically as evolutions in time or as a space-time plot for a specific time. The valuation is done using the finite element method, and features dual techniques as well. The program is also equipped with the availability to calculate error estimations and show them graphically. This gives the user not just a tool for calculating prices and Greeks in an easy way, but at the same time it aids to the understanding with visualization of the prices, Greeks, and error plots.

References

- [1] A. ANTOULAS, D. SORENSEN, AND S. GUGERCIN, *A survey of model reduction methods for large-scale systems*, Contemp. Math., 280 (2001), pp. 193–219.
- [2] T. BJÖRK, *Arbitrage Theory in Continuous Time*, Oxford University Press, 1998.
- [3] C. BORELL, *Matematik och Optioner*, Matematiska inst. CTH och GU, 412 96 Gothenburg, 1999.
- [4] M. BOS AND S. VANDERMARK, *Finessing discrete dividends*, Risk Magazine, (2002).
- [5] S. BRENNER AND L. SCOTT, *The mathematical theory of finite element methods*, Springer Verlag, 1994.
- [6] D. DUFFIE, *Dynamic Asset Pricing Theory*, Princeton University Press, 1992.
- [7] K. ERIKSSON, D. ESTEP, P. HANSBO, AND C. JOHNSON, *Computational Differential Equations*, Studentlitteratur, 1996.
- [8] D. ESTEP, M. LARSON, AND R. WILLIAMS, *Estimating the error of numerical solutions of systems of reaction-diffusion equations*, MEMOIRS of the American Mathematical Society, 146 (2000).
- [9] G. FOUFAS AND M. G. LARSON, *Valuing asian options using the finite element method and duality techniques*, to appear in JCAM, (2008).
- [10] V. FRISHLING, *A discrete question*, Risk Magazine, (2002).
- [11] H. GEMAN, N. KAROUI, AND J. ROCHET, *Changes of numeraire, changes of probability measure and option pricing*, Journal of Applied Probability, 32 (1995), pp. 443–458.
- [12] R. HEATH AND P. JARROW, *Ex-dividend stock price behaviour and arbitrage opportunities*, Journal of Business, 61 (1988), pp. 95–108.
- [13] S. LALL, P. KRYSL, AND J. MARSDEN, *Dimensional model reduction in non-linear finite element dynamics of solids and structures*, International Journal for Numerical Methods in Engineering, (2001), pp. 479–504.
- [14] J. VEČEŘ, *Unified Asian pricing*, RISK, (2002), pp. 113–116.

- [15] P. WILMOTT, *Paul Wilmott on quantitative finance.*, Wiley, 2000.
- [16] P. WILMOTT, J. DEWYNNE, AND S. HOWISON, *Option pricing*, Oxford Financial Press, 1993.
- [17] C. ZIENKIEWICZ, *The birth of the finite element method and of computational mechanics*, Internat. J. Numer. Methods Engrg., 60 (2004), pp. 3–10.

Paper I

Valuing European, Barrier, and Lookback Options using the Finite Element Method and Duality Techniques

Georgios Foufas*

April 14, 2008

Abstract

The main objective of this paper is to develop an adaptive finite element method for computation of the values and different sensitivity measures of ordinary European options, barrier options, and lookback options.

The options are priced using the Black-Scholes PDE-model, and the resulting PDE:s are of parabolic type in one spatial dimension with different boundary conditions and jump conditions at monitoring dates. The adaptive finite element method is based on a posteriori estimates of the error in desired quantities, which we derive using duality techniques. The a posteriori error estimates are tested and verified, and are used to calculate optimal meshes for each type of option. The use of adapted meshes gives superior accuracy and performance with less degrees of freedom than using uniform meshes. The suggested adaptive finite element method is stable and gives fast and accurate results.

1 Introduction

The valuation of different types of derivative contracts is very important in modern financial theory and practice. Exotic options have become very popular hedging and speculation instruments in recent years. At the same time a huge amount of literature has been devoted to the pricing and hedging of such instruments. We now give a short introduction to the different options studied in this paper.

Vanillas: In 1973 Black and Scholes, [4], presented their solution to the European call option problem. Their famous partial differential equation can

*Research Assistant, Department of Mathematics, Chalmers University of Technology, S-412 96 Gothenburg , Sweden, *email*: foufas@math.chalmers.se

be used to value several simple options. Ordinary European options, also referred to as vanillas, are also used to hedge more complex and advanced options.

Barriers: Several analytical formulas for the different types of barrier options have appeared in the literature. Most of these analytical results makes some kind of limiting assumption, which have led many people to apply numerical methods instead. Most of these numerical methods have been binomial or trinomial tree methods. But tree-methods in general show poor convergence when the barrier is close to the initial stock price. Tree methods may be viewed as some type of explicit finite difference method for solving a parabolic partial differential equation, as noted by Zvan, Forsyth and Vetzal, [30]. Instead they propose using an implicit method which has superior convergence (when the barrier is close to the region of interest) and stability properties. Using the same technique they value both continuous and discrete barrier options, with or without American constraints, and with the possibility of time-varying barriers and discrete dividends. Superior accuracy is achieved in fewer time steps. See Section 3 about barrier options for a more detailed review of the literature.

Lookbacks: Closed-form solutions for continuous sampled lookback option prices have been obtained in [10], [17], and [18]. For the discretely sampled lookback option one has to rely on numerical methods. Most of them are again based on some binomial method, see for example [2], [3], [8], or [20]. A PDE approach is described in Chapter 12 of [29]. Andreasen, [1], uses a change of numeraire techniques to obtain option prices as function of time and a one-dimensional Markovian state variable only, applicable to both the fixed and the floating strike lookback options, as well as Asian options, whereas Wilmott, Dewynne, and Howison, [29], uses a two-dimensional state variable for the lookback with fixed strike. In [31] the same PDE-model as [29] is used but in a stochastic volatility setting.

New Contributions: The options are priced using the Black-Scholes PDE-model. The resulting PDE:s are of parabolic type in one spatial dimension with different boundary conditions and jump conditions at monitoring dates. All options are priced using an adaptive finite element method allowing variable resolution in space and time.

In practice one is only interested in the price, and it's derivatives, in one or a few points. Using this criteria, the choice of computational mesh is based on a posteriori estimates of the error in desired quantities, which we derive using duality techniques. These dual techniques are shown to be very useful and simple, and allows us to improve the various PDE methods already existing for different contracts. The presented a posteriori error estimation formula is tested and verified in the case of the European option.

It is then used to perform mesh refinements in both time and space for the other options. This makes it possible to calculate an optimal mesh for each type of option, which significantly reduces the error without noticeably enhancing the computational effort.

The duality approach is general and applicable to problems with all kinds of algebraic constraints. Other exotic options, such as the the fixed strike lookback option and the Asian option are also studied by the authors, see [14] and [15]. The suggested adaptive finite element method is stable and gives fast and accurate results.

Outline: In Section 2 we formulate the finite element method and derive an a posteriori error estimate for the ordinary European option. Then in Sections 3 and 4 we extend the framework to barrier options and lookback options respectively.

2 An Adaptive Finite Element Method for the European Option

For the ordinary European option there exists an analytical valuation formula. But for other options, such as the discrete barrier option and the discrete lookback option, studied later in this paper, we have to rely on numerical solutions. In this section we present the finite element method and develop the a posteriori error estimation framework, for the basic European option. Later we extend the techniques to the more exotic options.

2.1 Mathematical Background

We consider a continuous time trading economy on a bounded time horizon $[0, T]$. Probability is represented by the probability space $(\Omega_T, \mathcal{F}_T, P)$, where $\Omega_T = C[0, T]$, P is the corresponding Wiener measure, and $\mathcal{F}_T = \sigma(W(t); t \leq T)$. For simplicity we consider the standard Black-Scholes setting with a risk free asset and a dividend paying stock. Let $B(t)$ denote the price of a risk free asset at time t governed by the equation $B(t) = B(0)e^{rt}$, where r is the constant interest rate. Further we denote by $S(t)$ the value of an asset at time t . We assume the existence of an equivalent martingale measure Q , under which the discounted stock price $e^{-r(T-t)}S_t$ is an \mathcal{F}_t -martingale. The existence of the risk neutral measure Q assures that the market is free of arbitrage possibilities. Under Q the stock price follows the stochastic differential equation

$$dS(t) = (r - \nu)S(t)dt + S(t)\sigma dW(t), \quad (2.1)$$

where r is the constant interest rate, ν is the constant continuous dividend yield, σ is the volatility, and $W(t)$ is a Q Brownian motion process. Here σ

is assumed to be a positive real number. The solution of (2.1) is

$$S(t) = S(0)e^{(r-\nu-\frac{\sigma^2}{2})t+\sigma W(t)}. \quad (2.2)$$

2.2 The Black-Scholes PDE

The value of the ordinary European option, $u(t, S(t)) = u(t, s)$, is given as the solution to Black-Scholes equation

$$u_t(t, s) + \frac{\sigma^2 s^2}{2} u_{ss}(t, s) + (r - \nu) s u_s(t, s) - r u(t, s) = 0, \quad t < T, \quad (2.3)$$

which is valid for $s = S(t) \in \mathbb{R}^+$. In order to construct a computational mesh we introduce a bounded interval $\Omega = [s_{min}, s_{max}] \subset \mathbb{R}^+$ with boundary $\partial\Omega = \{s_{min}, s_{max}\}$. We define the usual Hilbert space

$$H^1(\Omega) = \{v : \int_{\Omega} (|\nabla v|^2 + v^2) ds < \infty\}, \quad (2.4)$$

and let \mathcal{W} be the space of functions that are square integrable in time and belongs to $H^1(\Omega)$ in space, that is

$$\mathcal{W} = L^2([0, T], H^1(\Omega)). \quad (2.5)$$

We also use the notation $(u, v) = \int_{\Omega} uv ds$, and $(u, v)_{\partial\Omega} = u(s_{max})v(s_{max}) - u(s_{min})v(s_{min})$.

2.3 Variational Formulation

Multiplying the Black-Scholes equation (2.3) by the test function $v \in \mathcal{W}$ and integrating on $\Omega \times [0, T]$ we obtain

$$\int_0^T \left((u_t, v) + (r - \nu)(s u_s, v) + \frac{\sigma^2}{2} (s^2 u_{ss}, v) - r(u, v) \right) dt = 0. \quad (2.6)$$

Using integration by parts we get

$$(s^2 u_{ss}, v) = (s^2 u_s, v)_{\partial\Omega} - 2(s u_s, v) - (s^2 u_s, v_s). \quad (2.7)$$

Thus equation (2.6) becomes

$$\begin{aligned} \int_0^T \left((u_t, v) + (r - \nu - \sigma^2)(s u_s, v) \right. \\ \left. - \frac{\sigma^2}{2} (s^2 u_s, v_s) + \frac{\sigma^2}{2} (s^2 u_s, v)_{\partial\Omega} - r(u, v) \right) dt = 0. \end{aligned} \quad (2.8)$$

The boundary conditions for the European call option are $u(t, 0) = 0$ and $u(t, s) \sim se^{-\nu(T-t)}$ as $s \rightarrow \infty$, and for the corresponding put $u(t, 0) = Ke^{-r(T-t)}$ and $u(t, s) \sim 0$ as $s \rightarrow \infty$, see for example Wilmott, [29]. For simplicity of implementation we use the artificial boundary condition $u_{ss} = 0$ on $\partial\Omega$ for both the put and the call instead. This boundary condition works well for all contracts where the payoff is at most linear in the underlying (see [29]) and does not affect the accuracy of the solution. Using equation (2.3) we can rewrite the boundary condition as

$$u_s = \frac{r}{s(r-\nu)}u - \frac{1}{s(r-\nu)}u_t, \quad (2.9)$$

and enforce it weakly by inserting identity (2.9) into equation (2.8). We thus want to solve the problem: find $u \in \mathcal{W}$ such that

$$\begin{cases} \int_0^T (m(u_t, v) + a(u, v)) dt = 0, \\ u(T, s) = \begin{cases} \max(s - K, 0), & \text{for a call,} \\ \max(K - s, 0), & \text{for a put,} \end{cases} \end{cases} \quad (2.10)$$

for every $v \in \mathcal{W}$, where

$$m(u_t, v) = (u_t, v) - \frac{\sigma^2}{2(r-\nu)}(su_t, v)_{\partial\Omega}, \quad (2.11)$$

and

$$\begin{aligned} a(u, v) &= (r - \nu - \sigma^2)(su_s, v) - \frac{\sigma^2}{2}(s^2u_s, v_s) \\ &\quad + \frac{\sigma^2 r}{2(r-\nu)}(su, v)_{\partial\Omega} - r(u, v). \end{aligned} \quad (2.12)$$

2.4 Finite Element Approximation

The finite element method is based on solution of the variational problem (2.10) with \mathcal{W} replaced by a finite dimensional function space of piecewise polynomials in space and time. For background on the finite element method see for instance [11].

We now partition $[0, T]$ as $0 = t_0 < t_1 < t_2 < \dots < t_N = T$, denoting each time interval by $I_n = (t_{n-1}, t_n)$ and each time step by $k_n = t_n - t_{n-1}$. Similarly we partition Ω as $s_{min} = s_0 < s_1 < s_2 < \dots < s_J = s_{max}$, denoting each spatial interval by $\kappa_j = [s_{j-1}, s_j)$ and the length of each interval by $h_j = s_j - s_{j-1}$.

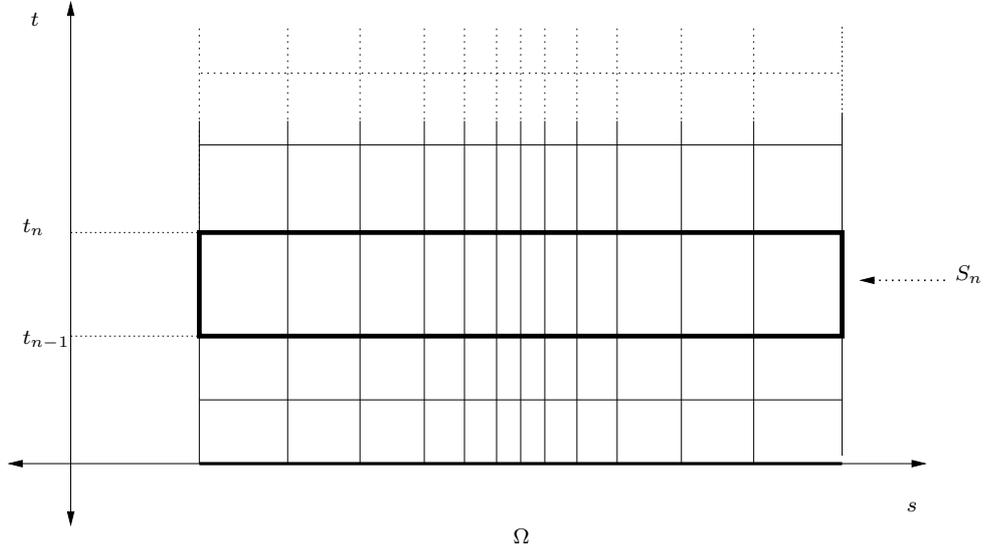


Figure 1: *Space-time discretization.*

In space, we let $\mathcal{V}^p \subset H^1(\Omega)$ denote the space of piecewise continuous functions of order p . On each space-time slab $S_n = I_n \times \Omega$, we define

$$\mathcal{W}_n^q = \{w(t, s) : w(t, s) = \sum_{j=0}^q t^j v_j(s), v_j \in \mathcal{V}^p, (t, s) \in S_n\}. \quad (2.13)$$

Let $\mathcal{W}^q \subset \mathcal{W}$ denote the space of functions defined on $[0, T] \times \Omega$ such that $v|_{S_n} \in \mathcal{W}_n^q$ for $1 \leq n \leq N$. For simplicity, we only give details for the continuous Galerkin method cG(p)-cG(q), (see e.g. [11] or [12]) which is defined by the following discrete version of equation (2.10). Find $U \in \mathcal{W}^q$ such that for $1 \leq n \leq N$

$$\begin{cases} \int_{I_n} (m(U_t, v) + a(U, v)) dt = 0 & \text{for all } v \in \mathcal{W}_n^{q-1}, \\ U^-(t_n) = U^+(t_n), & n = N-1, \dots, 1, \\ U^-(t_N) = u_T, \end{cases} \quad (2.14)$$

where $U^\pm(t_n) = \lim_{\epsilon \rightarrow 0, \epsilon > 0} U(t_n \pm \epsilon)$. In the cG(1) method the approximation U of u is continuous piecewise linear in time and space, while the test functions v are continuous linear in space and piecewise constant in time. It is also possible to use a discontinuous method in time, we refer to [11], for details on the resulting discontinuous Galerkin method, cG(p)-dG(q).

2.5 A Posteriori Error Estimation

2.5.1 Error Representation Formula

Since we are only interested in the solution, and it's derivatives, in one or a few points of Ω at time $t = 0$, we wish to find a mesh tailored for efficient and accurate solution at the points of interest. In order to find such a mesh we derive a posteriori error estimates of the error in the points of interest using duality techniques (see [11] or [12]).

To represent the error in a linear functional, $(u - U, \psi)$, we introduce the continuous dual problem for the Black-Scholes equation (2.3). Find $\phi \in \mathcal{W}$ such that

$$\begin{cases} -\phi_t + (\sigma^2 + \nu - 2r)\phi - (r - \nu - 2\sigma^2)s\phi_s + \frac{\sigma^2}{2}s^2\phi_{ss} = 0, \\ \phi(0, s) = \psi. \end{cases} \quad (2.15)$$

For simplicity we consider this equation over the whole space interval neglecting boundary conditions. Multiplying with the error $e = u - U \in \mathcal{W}$ and integrating in space and time we get

$$\int_0^T \left(-(\phi_t, e) + (\sigma^2 + \nu - 2r)(\phi, e) - (r - \nu - 2\sigma^2)(s\phi_s, e) + \frac{\sigma^2}{2}(s^2\phi_{ss}, e) \right) dt = 0. \quad (2.16)$$

The functions ϕ and ϕ_s are in principle zero close to $s = s_{min}$ and $s = s_{max}$ if the domain is large enough. Using integration by parts and neglecting the boundary terms we get

$$\begin{aligned} & -(\phi(T, s), e(T, s)) + (\phi(0, s), e(0, s)) \\ & + \int_0^T \left((\phi, e_t) + (\sigma^2 + \nu - 2r)(\phi, e) + (r - \nu - 2\sigma^2)(s\phi, e_s) \right) dt \\ & + \int_0^T \left((r - \nu - 2\sigma^2)(\phi, e) - \frac{\sigma^2}{2}(s^2\phi_s, e_s) - \sigma^2(s\phi_s, e) \right) dt = 0. \end{aligned} \quad (2.17)$$

Note that integration by parts gives

$$-\sigma^2(s\phi_s, e) = \sigma^2(s\phi, e_s) + \sigma^2(\phi, e), \quad (2.18)$$

using this identity, $\phi(0, s) = \psi$, and $e(T) = 0$, we get

$$\begin{aligned} (\psi, e(0, s)) = & \\ & - \int_0^T \left((\phi, e_t) - r(\phi, e) + (r - \nu - \sigma^2)(s\phi, e_s) - \frac{\sigma^2}{2}(s^2\phi_s, e_s) \right) dt. \end{aligned} \quad (2.19)$$

Recalling the earlier defined bilinear forms (2.11) and (2.12), and that we neglect the boundary terms we can also write

$$(\psi, e(0, s)) = - \int_0^T \left(m(e_t, \phi) + a(e, \phi) \right) dt. \quad (2.20)$$

Since $e = u - U$ and u solves equation (2.10) we get the error representation formula

$$\boxed{(\psi, e(0, s)) = \int_0^T \left(m(U_t, \phi) + a(U, \phi) \right) dt} \quad (2.21)$$

If we for example are interested in the error at $s = s_\alpha$, we choose $\psi = \delta_{s_\alpha}(s)$, and get the error representation formula

$$e(0, s_\alpha) = \int_0^T \left(m(U_t, \phi) + a(U, \phi) \right) dt. \quad (2.22)$$

If one instead is interested in derivatives of the solution, then a different ψ is chosen, as shown later on.

2.5.2 Estimating the Error

Let $\pi : \mathcal{W} \rightarrow \mathcal{W}^{q-1}$ be the L_2 projection in time, and let P be a suitable interpolation operator into \mathcal{V}^p in space. Thus πP is an interpolation operator such that $\pi P\phi \in \mathcal{W}^{q-1}$. Then using Galerkin orthogonality (2.14), we can replace ϕ by $\phi - \pi P\phi = \phi - P\phi + P\phi - \pi P\phi$. Equation (2.21) can then be written as

$$\begin{aligned} (\psi, e(0, s)) &= - \int_0^T \left(m(U_t, \phi - P\phi) + a(U, \phi - P\phi) \right) dt \quad (2.23) \\ &\quad - \int_0^T \left(m(U_t, P\phi - \pi P\phi) + a(U, P\phi - \pi P\phi) \right) dt \\ &= - \sum_n \sum_j \int_{I_n} \left(R_{\kappa_j}^s(U), \phi - P\phi \right) dt \\ &\quad - \sum_n \int_{I_n} \left(R^t(U), P\phi - \pi P\phi \right) dt, \end{aligned}$$

where

$$\begin{aligned} (R_{\kappa_j}^s(U), \phi - P\phi) &= - \frac{\sigma^2}{2} (s^2[U_s], \phi - P\phi)_{\partial\kappa_j} \quad (2.24) \\ &\quad + (U_t + (r - \nu)sU_s + \frac{\sigma^2}{2}s^2U_{ss} - rU, \phi - P\phi)_{\kappa_j} \end{aligned}$$

is the space residual, and

$$(R^t(U), P\phi - \pi P\phi) = (U_t + (r - \nu)sU_s + \frac{\sigma^2}{2}s^2U_{ss} - rU, P\phi - \pi P\phi) \quad (2.25)$$

is the time residual. Here we used the notation $[U_s]$ to denote the jump in U_s over element interfaces.

Finally, we present an algorithm for calculating the error.

Error Estimation Algorithm:

- Compute an approximation Φ of ϕ using an enriched finite element space, for instance higher order approximation.
- Compute $P\Phi$.
- Compute $\int_{I_n} (R_{\kappa_j}^s(U), \phi - P\phi) dt$ using quadrature in space and time for each element κ_j and time step.
- Compute $\pi P\Phi$.
- Compute $\int_{I_n} (R^t(U), P\phi - \pi P\phi) dt$ using quadrature in space and time for each time step.

2.5.3 Examples

Using the error estimation algorithm in the previous section we are able to calculate the error in desired quantities for different values of the parameters. This makes it possible to identify regions where a fine mesh is necessary.

Example 1. To estimate the error at $s = s_\alpha$ we let $\psi = \delta_{s_\alpha}(s)$ in (2.15). In order to implement this condition we use the approximation

$$\delta_{s_\alpha}(s) \approx \frac{1}{\epsilon\sqrt{\pi}} e^{-((s-s_\alpha)/\epsilon)^2} := \delta_{s_\alpha}(s, \epsilon), \quad (2.26)$$

where ϵ is a parameter that controls how well the delta function is approximated. In this example we have used $\epsilon = 1$. As seen from Figure 2, the solution to the dual problem differs from zero only within a short interval of Ω .

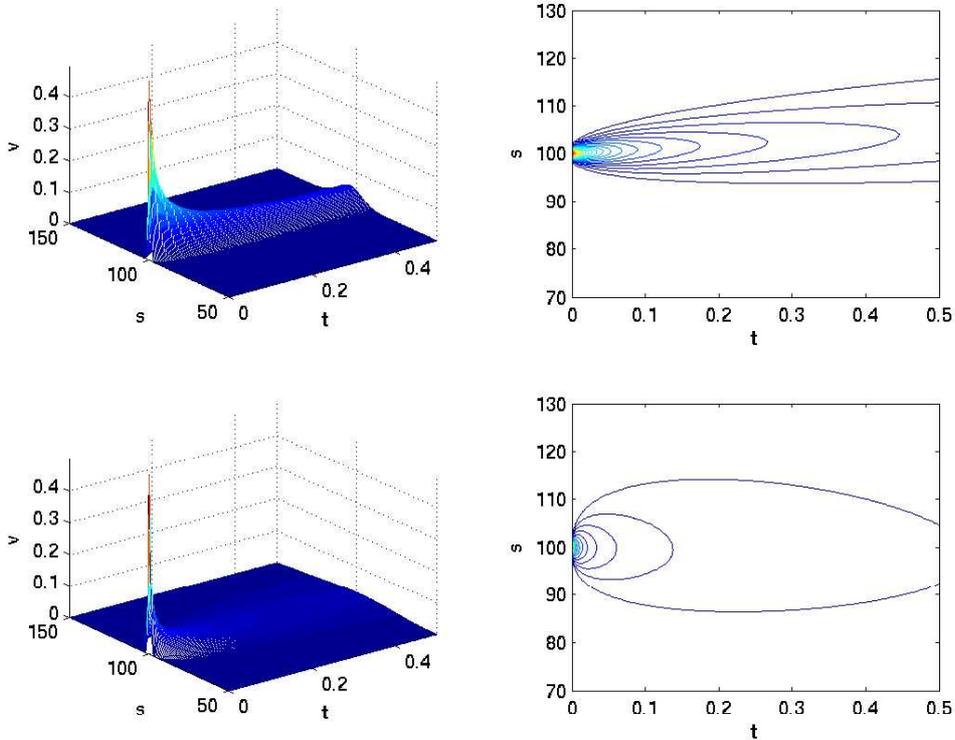


Figure 2: Above on the left, ϕ , for $\sigma = 0.1$, and $r = 0.1$, when $\psi = \delta_{100}(s, \epsilon)$. Below on the left, ϕ , for $\sigma = 0.3$. On the right, contour plots using 30 levels. Solutions computed using the cG(2)-dG(1) method with 200 space and time points.

We now check that the error representation formula really works. By using the error estimation algorithm in the previous section we can get an approximation of the functional of the error, that is an approximation of the right hand side of equation (2.21). This can then be compared to calculating the left hand side of equation (2.21) directly using the real error in the approximate solution, found by using Black-Scholes formula. The dual solution is calculated on a finer mesh, and using higher order approximations. In Figure 3, we see the contributions to error formula (2.21) from each space-time slab. The dual was calculated using the cG(2)-dG(1) method, and the primal using the cG(1)-cG(1) method. The dual mesh was thirty two times finer in each direction. The value of the functional of the error found by using the error representation formula was in this case 0.2033, in excellent agreement with the real value, that is the value of the left hand

side of equation (2.21), which was 0.2030. We also note that the contribution to the error differs from zero only within a short interval of Ω , just as the dual solution. This means that we may use a more sparse mesh where the contribution to the error is small and thus save computation time. The solution is larger near time $t = 0$, implying that one should use a finer time step there. Obviously the result depends on the value of the volatility σ , and the other parameters, which can be seen from the plot of the dual solution. We will later see how we can use the error representation formula to derive an optimal mesh for each problem.

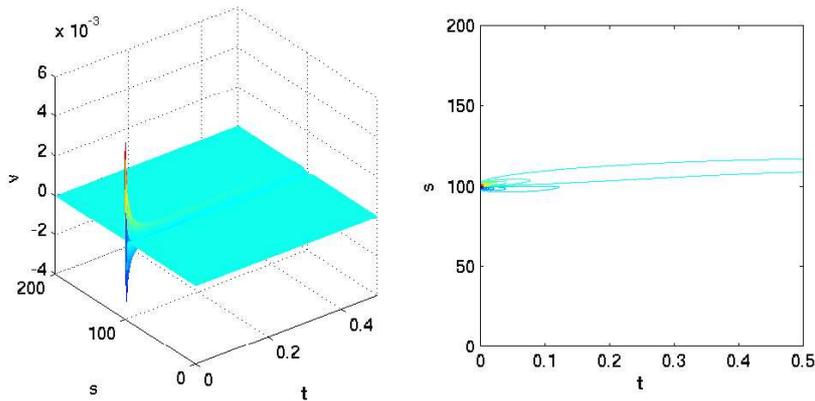


Figure 3: On the left, the contributions to the error of call option for $\sigma = 0.1$, $r = 0.1$, and $K = 100$ when $\psi = \delta_{100}(s, \epsilon)$. On the right, contour plot using 30 levels. The dual was computed using the $cG(2)$ - $dG(1)$ method with 400 space and time points, and the primal using the $cG(1)$ - $cG(1)$ method with 20 space and time points.

Example 2. In order to make a good estimation of the derivative of the solution, which is interesting when calculating the Greek *delta*, we need to study a different dual problem. We approximate the derivative using the

central difference formula

$$\frac{\partial u}{\partial s} \approx \frac{u(s + \mu) - u(s - \mu)}{2\mu} := \frac{\partial_h u}{\partial s}. \quad (2.27)$$

To estimate the error of the derivative of the solution at $s = s_\alpha$, $u_s(s_\alpha)$, we thus choose

$$\begin{aligned} \psi(s) &= \frac{\delta_{s_\alpha}(s - \mu) - \delta_{s_\alpha}(s + \mu)}{2\mu} \\ &\approx \frac{\delta_{s_\alpha}(s - \mu, \epsilon) - \delta_{s_\alpha}(s + \mu, \epsilon)}{2\mu} \end{aligned} \quad (2.28)$$

in (2.15), for an appropriate choice of μ . The error in our estimation of the derivative can be split into two parts

$$\left(\frac{\partial u}{\partial s} - \frac{\partial_h U}{\partial s} \right) = \left(\frac{\partial u}{\partial s} - \frac{\partial_h u}{\partial s} \right) + \left(\frac{\partial_h u}{\partial s} - \frac{\partial_h U}{\partial s} \right). \quad (2.29)$$

The first term corresponds to the error in (2.27), while the second can be estimated using the a posteriori estimate. Figure 4 shows the dual solution for this choice of ψ when $\mu = 1$ and $\epsilon = 1$. Figure 5 shows the contributions to the error estimation formula from each space-time slab. We see that this solution is even more centrally oriented than the previous one, implying that the derivative has a local dependence.

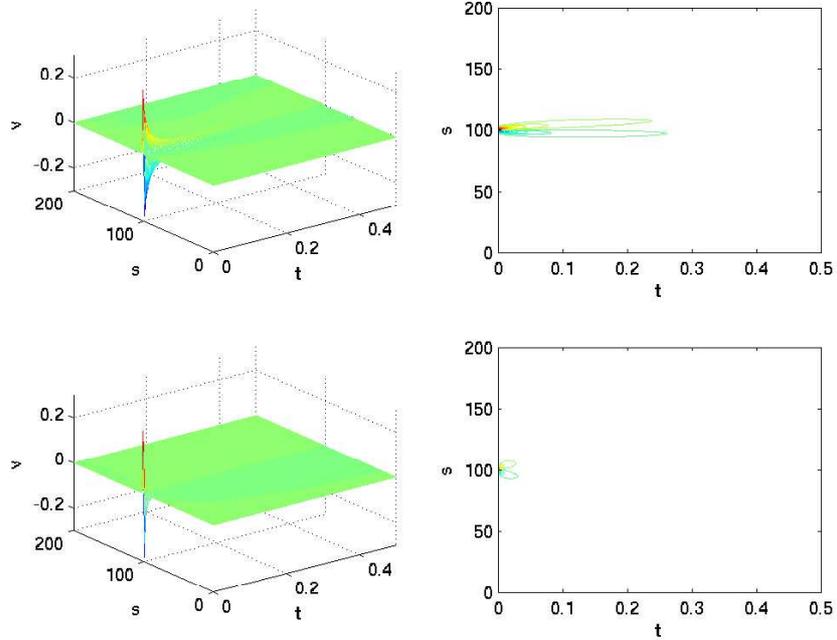


Figure 4: Above on the left, ϕ , for $\sigma = 0.1$, and $r = 0.1$, when ψ is chosen as in example 2. Below on the left, ϕ , for $\sigma = 0.3$. On the right, contour plots using 30 levels. Solutions computed using the $cG(2)$ - $dG(1)$ method with 200 space and time points.

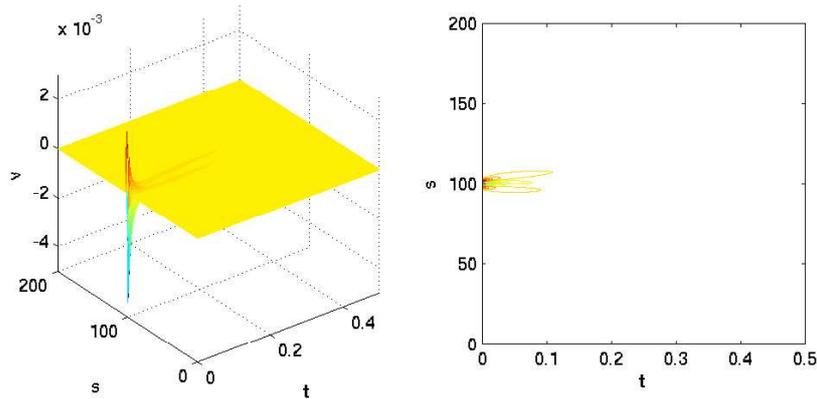


Figure 5: On the left, the contributions to the functional of the error of call option for $\sigma = 0.1$, $r = 0.1$, and $K = 100$ when ψ is chosen as in example 2. On the right, contour plot using 30 levels. The dual was computed using the $cG(2)$ - $dG(1)$ method with 640 space and time points, and the primal using the $cG(1)$ - $cG(1)$ method with 20 space and time points.

2.6 Adaptive Mesh Refinement

Adaptive mesh refinement may be accomplished in many different ways. Our goal not is to create the best adaptive method, since adaptivity would be to slow to use in reality. Rather we wish to create an optimal mesh in advance for each case, so that when valuing an option we simply use a suited pre-calculated mesh. This gives superior performance. In this section we show how these meshes are calculated and what typical meshes look like.

Mesh Refinement Algorithm:

- Compute an approximation U of u using the FE method on a coarse mesh.

- Compute the error in desired quantities by using the a posteriori error estimation algorithm.
- Calculate the time and space averages of the contributions to the error from each space-time slab. This gives us two vectors, one with time averages and one with space averages.
- Identify the $Q\%$ largest elements in the space average vector, and refine the corresponding time steps by dividing them in half.
- Identify the $Q\%$ largest elements in the time average vector, and refine the corresponding spatial steps by dividing them in half.
- Compute a new FE approximation U on the refined mesh.
- Repeat until minimum mesh size is reached.

In Figure 6, we see a typical mesh resulting from using the mesh refinement algorithm above. In this case Q was set to 10%. Three successive refinements were made, starting from a sparse mesh with 20 nodes in time and space. The final mesh has only 27 nodes in each direction, but the error has decreased by a factor 70. The dual was calculated using a fine mesh with 640 nodes in time and space.

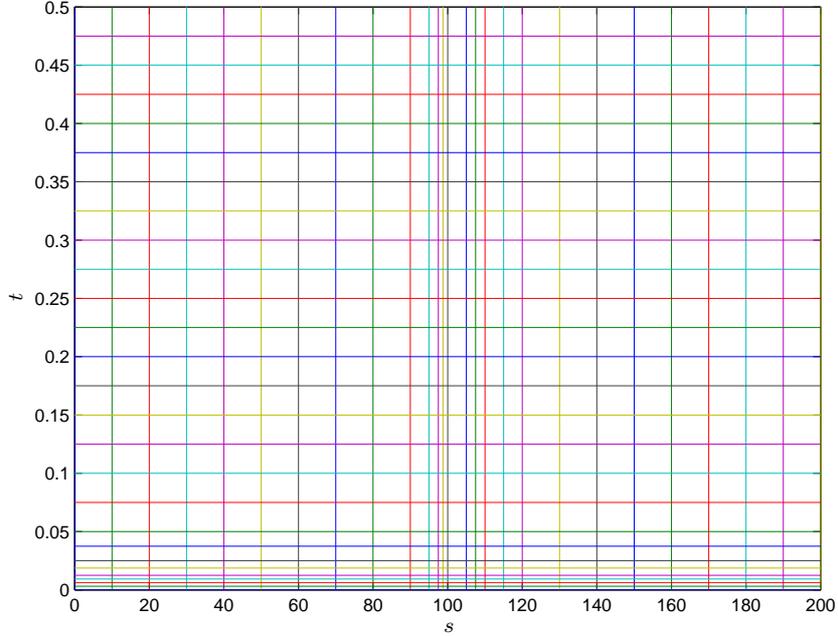


Figure 6: *The resulting mesh using the mesh refinement algorithm, calculated for a call option with $\sigma = 0.1$, $r = 0.1$, and $K = 100$ when $\psi = \delta_{100}(s, \epsilon)$. The dual was computed using the $cG(2)$ - $dG(1)$ and the primal using the $cG(1)$ - $cG(1)$ method. Three successive refinements were made.*

2.7 The Greeks

In order to hedge our option, we need the sensitivity measures, or the so called Greeks. The most common one is the so called *delta*

$$\Delta = \frac{\partial u}{\partial s}. \quad (2.30)$$

The second derivative is called *gamma*

$$\Gamma = \frac{\partial^2 u}{\partial s^2}. \quad (2.31)$$

The decay of value in time is represented by the *theta*, where

$$\Theta = -\frac{\partial u}{\partial t}. \quad (2.32)$$

Sensitivity to volatility called the *vega* and is defined by

$$\frac{\partial u}{\partial \sigma}, \quad (2.33)$$

and sensitivity to interest rate is called *rho*, defined as

$$\rho = \frac{\partial u}{\partial r}. \quad (2.34)$$

In Figure 7 we see the delta and gamma of a European call at time $t = 0$.

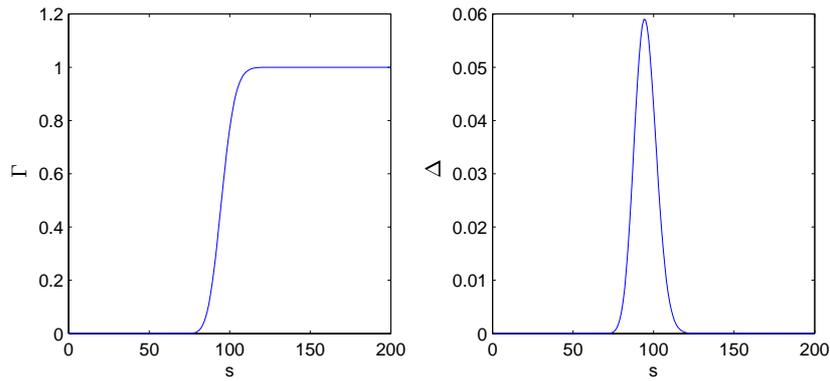


Figure 7: *The delta (on the left) and gamma (on the right) of a European call at time $t = 0$, with $\sigma = 0.1$, $T = 0.5$, $K = 100$ and $r = 0.1$.*

2.8 Results

We begin by validating our method against the known exact solution for the European call option. Recalling the previous calculations in Example 1 in Section 2.5.3 we know that the error representation formula works and is accurate. The value of the functional of the error found by using the

error representation formula in the test example on the European option was 0.2033, in excellent agreement with the real value of the functional found by using Black-Scholes formula, which was 0.2030. Table 1 compares values of the European call calculated using the cG(1)-cG(1) finite element method mentioned above, with the analytical value derived by Black-Scholes formula. We see that the FE method is very stable and has a maximum relative error of 0.1 percent when 400 time points are used. Figure 8 shows

σ	S(0)	FE(200)	FE(400)	Black-Scholes	Relative error (%)
0.10	90	0.8067	0.8093	0.8101	0.107
	100	5.8478	5.8496	5.8503	0.011
	110	14.9287	14.9297	14.9300	0.002
0.20	90	3.0487	3.0500	3.0504	0.014
	100	8.2767	8.2775	8.2778	0.003
	110	16.0177	16.0184	16.0187	0.002
0.30	90	5.5198	5.5206	5.5209	0.005
	100	10.9058	10.9063	10.9065	0.002
	110	18.0464	18.0468	18.0469	0.0008

Table 1: *The European call calculated using the cG(1)-cG(1) method compared to Black-Scholes analytical value when $r = 0.1$, $q = 0.0$, $T = 0.5$, $K = 100$, and $t = 0$. The number of time and space points is given in parenthesis. The relative error is between the FE(400) solution and the analytical solution.*

the finite element solution calculated using a the adapted mesh in the previous section. The mesh is finer close to time $t = 0$ and close to the strike price, but it is not centered around the strike price. In this way the same accuracy is achieved in less degrees of freedom. The original uniform mesh has 20 nodes in time and space. By using the error representation formula the error was calculated to 0.2 for the uniform mesh. The adapted mesh has only 27 nodes in the spatial direction, but the error has decreased by a factor 70 to 0.0028.

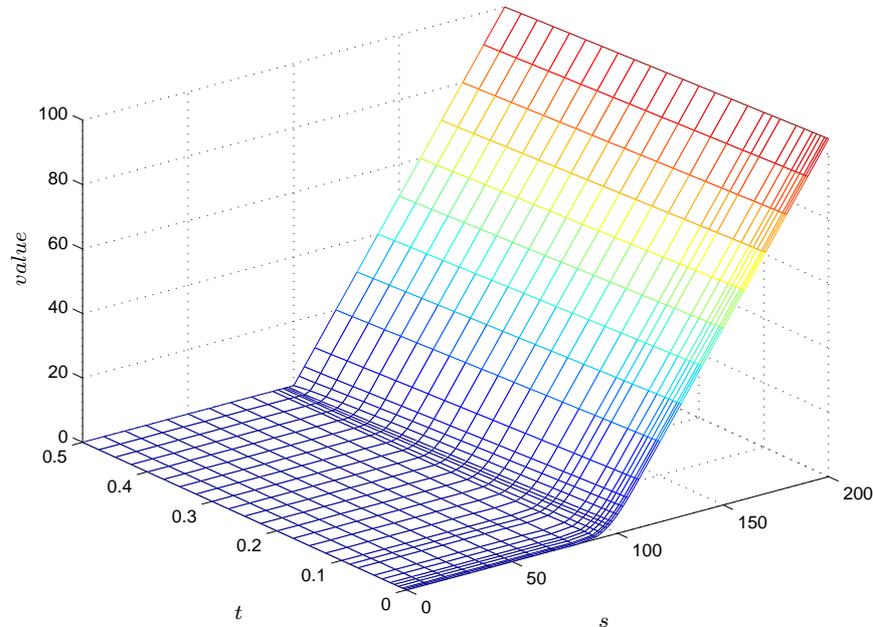


Figure 8: *The finite element solution U , when $\sigma = 0.1$, $q = 0.0$, $K = 100$, and $r = 0.10$. Computed using the $cG(1)$ - $cG(1)$ method on an adapted mesh with 27 time and space points.*

3 Barrier Options

3.1 Classification

Barrier options are path-dependent options. They have a payoff depending on whether or not the underlying asset crosses a predetermined level, called the barrier. There are two main types of barrier option, the knock-out option and the knock-in option. The knock-out option gives a payoff unless the underlying asset crosses the barrier during the lifetime of the option. The knock-in option gives a payoff as long as the barrier is reached before the expiry. We also separate between up and down options. If the barrier is above the initial value of the underlying asset, we have an up option, otherwise we have a down option. For example, the up-and-out call option with barrier H and strike K gives the usual payoff $\max(S - K, 0)$ at expiry unless the underlying asset crosses the barrier during the lifetime of the

option, that is unless $S_t \geq H$ for some $t \leq T$. Thus the up-and-out call option gives the payoff

$$\max(S_T - K, 0)1_{\{\max_{t \in [0, T]} S_t < H\}}, \quad (3.1)$$

at maturity T . In the same way we conclude that the down-and-out call option gives the payoff

$$\max(S_T - K, 0)1_{\{\min_{t \in [0, T]} S_t > H\}}. \quad (3.2)$$

at maturity T . There are also variants of the barrier options mentioned above. The barrier could be time-dependent, usually piecewise constant, or one could allow for early exercise. Another type of barrier option is the double barrier, with both a lower and an upper barrier. Sometimes a rebate is paid if the barrier is reached. All of the different types of barrier options mentioned above exist both with continuous and discrete monitoring of the barrier. The latter is perhaps the most natural and used one. There are many other more exotic types of barrier options, such as outside barrier options where the barrier depends on an other asset, or soft barrier options which allows the contract to be gradually knocked-in or out, or Parisian options that have barriers that are triggered only if the underlying asset has been beyond the barrier for more than a specified time. We refer the interested reader to [28].

Barrier options are popular options, mainly because they are cheaper than the corresponding options without barriers. If an investor believes that it is unlikely that the underlying asset will fall below a certain level, then it is natural to buy a knock-out option with the barrier at that level. Barrier options gives investors the opportunity to avoid paying for scenarios that they believe are unlikely. Off course this also involves a certain risk.

3.2 Review of Literature

The first known literature on the pricing of barrier options dates back to Merton in 1973, [23], who presented a closed-form solution for the price of the continuously monitored down-and-out European call. After that several analytical formulas for the different types of continuous barrier options have appeared in the literature. Both Rich, [25], and Rubenstein and Reiner, [27], presented pricing formulas for a variety of standard barrier options in 1991. More exotic barrier options, such as the partial barrier option and the rainbow barrier option, have also been analytically valued by Heynen and Kat, [19], and Carr, [7]. Several people have also presented analytical results on the continuous double barrier option, see [16], [21], and [22].

In general the analytical approach relies on limiting assumptions, for instance that the monitoring is assumed to be continuous, and if barriers

change over time they are assumed to change as an exponential function of time. These limitations have led many to apply numerical methods instead. Most of these numerical methods have been binomial or trinomial tree methods. The binomial method shows very poor convergence unless the number of time steps is chosen in such a way as to ensure that the barrier lies on a horizontal layer of nodes in the tree, see [5] and [24]. Ritchken, [26], applies a trinomial tree method which performs better than the binomial, but it may still need a very large number of time steps if the initial stock price is close to a barrier (see also [6]). Cheuk and Vorst, [9], improve Ritchken's method by incorporating a time-dependent shift in the trinomial tree, but it still requires a fairly large number of time steps if the barrier lies close to the initial stock price. Figlewski and Gao, [13], uses an adaptive mesh in their trinomial tree, which gives them a more flexible and efficient method. Each of these tree methods may be viewed as some type of explicit finite difference method for solving a parabolic partial differential equation, as noted by Zvan, Forsyth and Vetzal, [30]. Instead they propose using an implicit method which has superior convergence (when the barrier is close to the region of interest) and stability properties. Using the same technique they value both continuous and discrete barrier options, with or without American constraints, and with the possibility of time-varying barriers and discrete dividends. This is accomplished with superior accuracy in fewer time steps than the methods mentioned earlier.

3.3 Pricing Barrier Options

For the continuously monitored barrier option there exists various analytical pricing formulas, but as mentioned in the previous section, for the discretely monitored barrier option one has to rely on numerical techniques. Therefore we concentrate on the valuation of discrete barrier options.

Barrier options are only weakly path-dependent, their value depends on whether the barrier is reached or not, not on any other information about the path. In an other paper, [15], we study Asian options, which are strongly path-dependent. For simplicity we only study out-options, in-options can be handled through an in-out parity argument shown later.

The value of an out barrier option still satisfies the Black-Scholes equation as long as the underlying has not crossed the barrier. The continuous barrier option is only to be solved on a partition of \mathbb{R}^+ , since if the underlying asset crosses the barrier the option is worthless. But for the discrete barrier option we must solve the Black-Scholes equation over the whole of $S \in \mathbb{R}^+$, since the barrier can be crossed between monitoring points and still not be knocked-in or out.

3.3.1 Barrier Constraint

For the value of the discretely monitored up-and-out call, $u(t, s)$, with monitoring dates $D = \{t_k^*\}_{k=0}^K \subset \{t_n\}_{n=0}^N$, where $t_0 = 0$ and $t_N = T$, we have the barrier constraint

$$u^-(t_k^*, s_j) = \tag{3.3}$$

$$BC(u^+(t_k^*)) := \begin{cases} 0 & \text{if } s_j \geq h(t_k^*)H, \quad j = 0, 1, \dots, J, \\ u^+(t_k^*, s_j) & \text{if } s_j < h(t_k^*)H, \quad j = 0, 1, \dots, J, \end{cases}$$

where H is the barrier, and h is a time dependent positive function which allows the barrier to move in time. The constraint is easily changed to handle a rebate, we simply set value to the rebate instead of zero if the barrier is crossed.

3.3.2 The Finite Element Method

Since the pricing partial differential equation is the same as for the ordinary European option, the finite element method is also same, except that we apply the barrier constraint at each monitoring date. With the same discretization as for the European option, we let $U(t_n)$ denote the approximate solution at time step n . Naturally a barrier constraint similar to (3.3) then holds for U . Thus we arrive at the problem, find $U \in \mathcal{W}^q$ such that for $1 \leq n \leq N$

$$\begin{cases} \int_{I_n} (m(U_t, v) + a(U, v)) dt = 0 & \text{for all } v \in \mathcal{W}_n^{q-1} \\ U^-(t_n) = U^+(t_n), \quad n = N-1, \dots, 1 & t_n \notin D, \\ U^-(t_n) = BC(U^+(t_n)), \quad t_n \in D, \\ U^-(t_N) = u_T, \end{cases} \tag{3.4}$$

where

$$m(U_t, v) = (U_t, v) - \frac{\sigma^2}{2(r - \nu)} (sU_t, v)_{\partial\Omega}, \tag{3.5}$$

and

$$\begin{aligned} a(U, v) &= (r - \nu - \sigma^2)(sU_s, v) - \frac{\sigma^2}{2}(s^2U_s, v_s) \\ &\quad + \frac{\sigma^2 r}{2(r - \nu)}(sU, v)_{\partial\Omega} - r(U, v). \end{aligned} \tag{3.6}$$

Again we use the artificial boundary condition $u_{ss} = 0$ at $\partial\Omega$.

3.3.3 In Barrier Options

The in barrier option is activated when the barrier is hit. When the barrier is hit we actually receive another derivative, namely the corresponding vanilla contract. One therefore refers to the in barrier option as a second-order contract. When solving for the value of an in barrier option one must first solve for the value of the vanilla. It therefore takes roughly twice as long time to calculate the value as for the out option.

If there are no rebates the relationship between in barrier options and out barrier options is very simple. By considering a portfolio consisting of one in-option and one out-option with the same barrier, time to maturity and expiry date, it is obvious from a financial point of view that the value of the portfolio is equal to the value of the corresponding vanilla option. This is because only one of the two barrier options can have hit the barrier at expiry, and the value of that barrier option then equals the value of the vanilla.

3.4 A Posteriori Error Estimation

3.4.1 Error Representation Formula

We now introduce the continuous dual problem for the barrier option, which differs from the European option case only in that we now have to include the barrier constraints. That is, find $\phi \in \mathcal{W}$

$$\begin{cases} -\phi_t + (\sigma^2 + \nu - 2r)\phi - (r - \nu - 2\sigma^2)s\phi_s + \frac{\sigma^2}{2}s^2\phi_{ss} = 0, \\ \phi(0, s) = \delta_{s_\alpha}, \\ \phi^+(t_k^*) = \begin{cases} 0 & \text{if } s_j \geq h(t_k^*)H, \\ \phi^-(t_k^*, s_j) & \text{if } s_j < h(t_k^*)H, \end{cases} \end{cases} \quad j = 0, 1, \dots, J, \quad t_k^* \in D. \quad (3.7)$$

For simplicity we consider this equation over the whole space interval neglecting boundary conditions. Multiplying with the error $e = u - U \in \mathcal{W}$ and integrating over space and time we get

$$\begin{aligned} \sum_k \int_{t_{k-1}^*}^{t_k^*} \left(-(\phi_t, e) + (\sigma^2 + \nu - 2r)(\phi, e) \right. \\ \left. - (r - \nu - 2\sigma^2)(s\phi_s, e) + \frac{\sigma^2}{2}(s^2\phi_{ss}, e) \right) dt = 0. \end{aligned} \quad (3.8)$$

Just as in the European case we now want to move derivatives from ϕ to the error e , using integration by parts. This gives the error representation formula

$$\boxed{e(0, s_\alpha) = - \sum_k \int_{t_{k-1}^*}^{t_k^*} (m(e_t, \phi) + a(e, \phi))} \quad (3.9)$$

where the bilinear forms $m(u_t, v)$ and $a(u, v)$ are defined exactly as before. We now prove this formula. The difference from the case of the European option is that we now have jumps in ϕ at the monitoring dates, affecting only the first term in (3.8). Studying this term in detail we see that

$$\begin{aligned}
& - \sum_k \int_{t_{k-1}^*}^{t_k^*} (\phi_t, e) dt \tag{3.10} \\
&= \sum_k \left(\int_{t_{k-1}^*}^{t_k^*} (\phi, e_t) dt - (\phi^-(t_k^*), e^-(t_k^*)) + (\phi^+(t_{k-1}^*), e^+(t_{k-1}^*)) \right) \\
&= \sum_k \left(\int_{t_{k-1}^*}^{t_k^*} (\phi, e_t) dt \right) - (\phi(T), e(T)) + (\phi(t_0), e(t_0)) \\
&\quad - \sum_k \left((\phi^-(t_k^*), e^-(t_k^*)) - (\phi^+(t_k^*), e^+(t_k^*)) \right).
\end{aligned}$$

Next we note that

$$\begin{aligned}
& (\phi^-(t_k^*), e^-(t_k^*)) - (\phi^+(t_k^*), e^+(t_k^*)) \tag{3.11} \\
&= \int_{s < h(t_k^*)H} (\phi^-(t_k^*)e^-(t_k^*) - \phi^+(t_k^*)e^+(t_k^*)) ds \\
&\quad + \int_{s \geq h(t_k^*)H} (\phi^-(t_k^*)e^-(t_k^*) - \phi^+(t_k^*)e^+(t_k^*)) ds = 0,
\end{aligned}$$

where the first term on the right is zero since $\phi^+(t_k^*) = \phi^-(t_k^*)$, and $e^-(t_k^*) = e^+(t_k^*)$ for $s < h(t_k^*)H$, and second term is zero since $\phi^+(t_k^*) = e^-(t_k^*) = 0$, for $s \geq h(t_k^*)H$.

Integrating the other terms in (3.8) we get

$$\begin{aligned}
& \int_{t_{k-1}^*}^{t_k^*} \left((\sigma^2 + \nu - 2r)(\phi, e) - (r - \nu - 2\sigma^2)(s\phi_s, e) + \frac{\sigma^2}{2}(s^2\phi_{ss}, e) \right) dt \tag{3.12} \\
&= \int_{t_{k-1}^*}^{t_k^*} \left(-r(\phi, e) + (r - \nu - \sigma^2)(s\phi, e_s) - \frac{\sigma^2}{2}(s^2\phi_s, e_s) \right) dt
\end{aligned}$$

just as for the case of the European option studied before. Now summing up, considering equations (3.10) and (3.12), we can rewrite (3.8) as

$$\begin{aligned}
0 &= -(\phi(T), e(T)) + (\phi(t_0), e(t_0)) \tag{3.13} \\
&\quad + \sum_k \int_{t_{k-1}^*}^{t_k^*} \left((\phi, e_t) - r(\phi, e) + (r - \nu - \sigma^2)(s\phi, e_s) - \frac{\sigma^2}{2}(s^2\phi_s, e_s) \right) dt,
\end{aligned}$$

or equivalently, using that $e(T) = 0$ and the boundary condition $\phi(0, s) = \delta_{s_\alpha}$, we have the error representation formula

$$e(0, s_\alpha) = - \sum_k \int_{t_{k-1}^*}^{t_k^*} (m(e_t, \phi) + a(e, \phi)), \quad (3.14)$$

where the bilinear forms $m(u_t, v)$ and $a(u, v)$ are defined exactly as before.

3.4.2 Examples

Using the same error estimation algorithm as for the European option we are able to calculate the error in desired quantities for different values of the parameters. This makes it possible to identify regions where a fine mesh is necessary.

Looking at the dual solution for some barrier options, Figures 9 and 10, we see again the same phenomena as for the vanilla option, namely that the dual is concentrated within a narrow area close to the center of the domain.

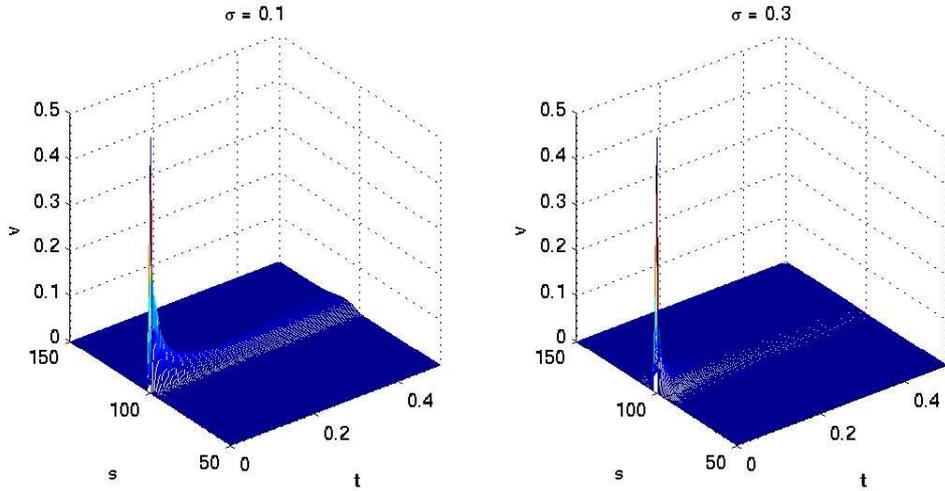


Figure 9: ϕ for two different values of σ when $r = 0.10$ and $q = 0.0$ for the weekly sampled down-and-out barrier call option with $K = 100$ and barrier $H = 99.9$. Computed with space step 0.1 and time step 0.005 , using the boundary condition $\phi(0, s) = \delta_{100}(s, \epsilon)$.

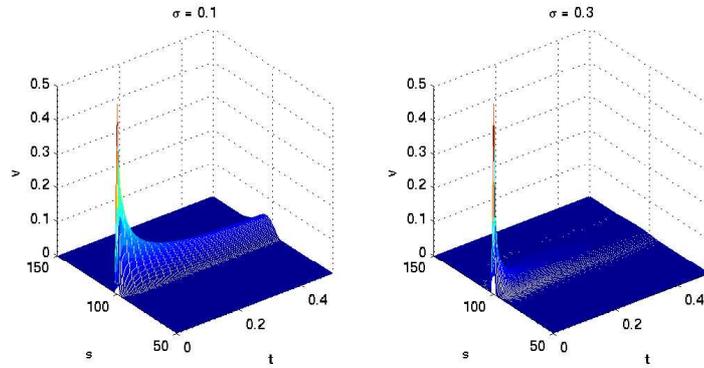


Figure 10: ϕ for two different values of σ when $r = 0.10$ and $q = 0.0$ for the weekly sampled double barrier call option with $K = 100$ and barriers $H_{low} = 95$ and $H_{high} = 125$. Computed with space 0.5 and time step 0.0025 , using the boundary condition $\phi(0, s) = \delta_{100}(s, \epsilon)$.

The sampling frequency clearly affects the dual solution, as can easily be seen from a contour plot of the solution, figure (11). It suggests that one should be careful near the sampling-dates since the solution changes rapidly there because of the monitoring constraint.

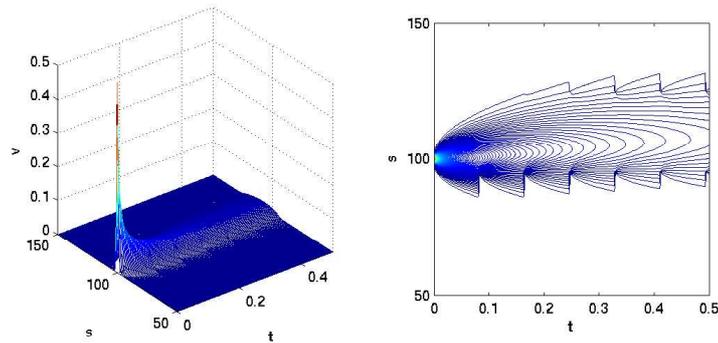


Figure 11: ϕ for the monthly sampled double barrier call option with $K = 100$ and barriers $H_{low} = 95$ and $H_{high} = 125$ when $\sigma = 0.2$, $r = 0.10$ and $q = 0.0$. Computed with space 0.5 and time step 0.0025 , using the boundary condition $\phi(0, s) = \delta_{100}(s, \epsilon)$.

In Figure 12, we see the contributions to the error representation formula (3.9) from each space-time slab. The dual was calculated using the cG(2)-dG(1) method, and the primal using the cG(1)-dG(1) method. The dual mesh was thirty two times finer in each direction. The value of the functional of the error found by using the error representation formula was in this case 1.96. We also note that the contribution to the error differs from zero only within a short interval of Ω , just as the dual solution. We now proceed to calculate adaptive meshes.

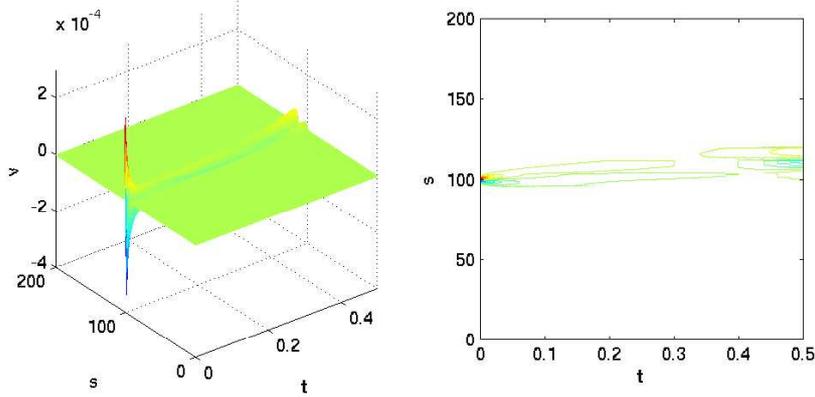


Figure 12: *The contributions to the functional of the error for the weekly sampled double barrier call option with $K = 100$ and barriers $H_{low} = 96$ and $H_{high} = 120$ when $\sigma = 0.1$, $r = 0.10$, $K = 100$, and $q = 0.0$. The dual was computed using the cG(2)-dG(1) method with 800 space and time points, and the primal using the cG(1)-dG(1) method with 25 space and time points, using the boundary condition $\phi(0, s) = \delta_{100}(s, \epsilon)$.*

3.4.3 Adaptive Mesh Refinement

As in the case of the European option we use the error representation formula to derive an optimal mesh for each problem. In Figure 13, we see a mesh

resulting from using the mesh refinement algorithm in the case of a weekly sampled double barrier option with barriers at $s = 96$ and $s = 120$ and strike price $K = 100$. In this case Q was set to 15%. Three successive refinements were made, starting from a sparse mesh with 25 nodes in time and space. The final mesh has only 37 nodes in each direction, but the functional of the error has decreased by a factor 5. The dual was calculated using a fine mesh with 800 nodes in time and space.

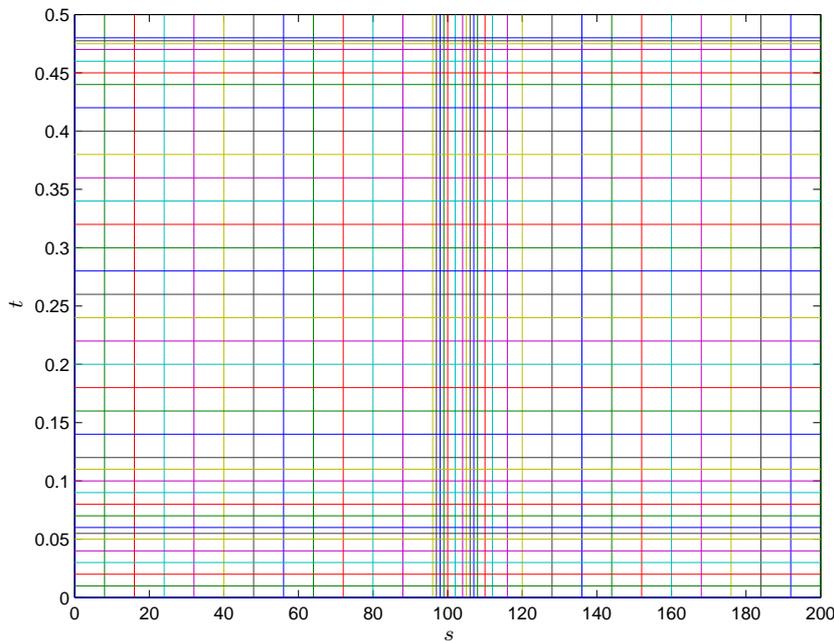


Figure 13: *The resulting mesh using the mesh refinement algorithm, calculated for a weekly sampled double barrier call option with $H_{low} = 96$ and $H_{high} = 120$, $T = 0.5$, $\sigma = 0.1$, and $r = 0.1$, when $\psi = \delta_{100}(s, \epsilon)$. The dual was computed using the $cG(2)$ - $dG(1)$ and the primal using the $cG(1)$ - $dG(1)$ method. Three successive refinements were made.*

3.5 Results

As noted by Zvan, Forsyth, and Vetzal, [30], implicit methods are more suitable for barrier options. Explicit methods may give rise to spurious numerical oscillations. We therefore use the discontinuous Galerkin method, which is more or less implicit.

As mentioned earlier, tree-methods show very poor convergence when the barrier lies close to the point of interest. Table 2 compares values of the down-and-out barrier call, when the barrier is close to the point of interest, for the finite element method developed in this paper with the finite difference method in [30]. As shown the methods are in agreement. As a comparison the value of the the corresponding continuous option is also given. Note the significant difference in price between the continuous option and the corresponding discrete option. This difference will be even larger if sampling rate is reduced.

	Continuous	Daily	Weekly
ZFV	0.16	1.51	3.00
FE	not calc.	1.51	3.01

Table 2: *The down-and-out barrier call when the barrier level $H = 99.9$. ZFV refers to the finite difference solution in [30], and FE refers to the $cG(2)$ - $dG(1)$ finite element solution. Parameter values are $r = 0.10$, $q = 0.0$, $\sigma = 0.2$, $T = 0.5$, $t = 0$, $K = 100$, and $S(0) = 100$.*

Figure 14 shows the finite element solution for a double barrier call option with weekly monitoring calculated using the adapted mesh in the previous section.

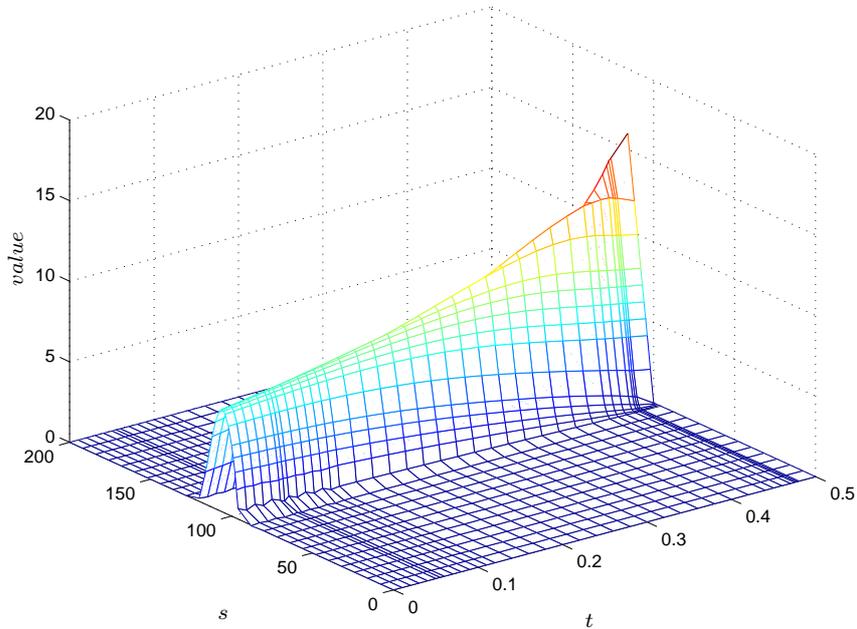


Figure 14: *The weekly sampled double barrier call computed using the $cG(1)$ - $dG(1)$ method, on an adapted mesh with 37 time points and space points. Parameter values are $\sigma = 0.1$, $r = 0.1$, $q = 0.0$, $T = 0.5$, $t = 0.0$, $K = 100$, and $H_{low} = 96$ and $H_{high} = 120$.*

The mesh is finer close to the strike price and towards lower and upper the barrier. We also note that the mesh is finer close to time $t = 0$ and $t = T$. In this way the same accuracy is achieved in less degrees of freedom. The original uniform mesh has 25 nodes in time and space. By using the error representation formula the functional of the error was calculated to 1.96 for the uniform mesh. The adapted mesh has only 37 nodes in the spatial direction, but the functional of the error has decreased by a factor 5 to 0.40.

In Figure 15 we see the value and the delta of a double barrier call at time $t = 0$.

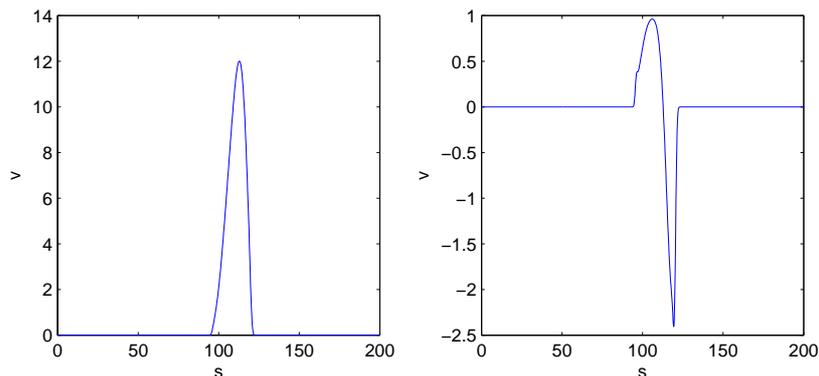


Figure 15: *To the left, the value of an double barrier call, and to the right the corresponding delta, at time $t = 0$, with $\sigma = 0.1$, $T = 1$, $K = 100$, and $r = 0.1$.*

4 Lookback Options

4.1 Classification

The lookback option has a payoff that depends on the maximum or minimum of the underlying stock price over some given interval in time. There are two types of lookback options, the lookback rate option and the lookback strike option, also known as the fixed strike and the floating strike lookback option respectively. Both of these options exists as puts and calls. If we denote the maximum asset price over the time interval $0 \leq t \leq T$ by M , the lookback strike put option has almost the same payoff as the vanilla put, but with M replacing the exercise price K , that is

$$\max(M - S(T), 0). \quad (4.1)$$

In the same way the corresponding call option gives almost the same payoff as the vanilla call, but with M , now being the minimum asset price during the lifetime of the option, replacing the exercise price K , that is

$$\max(S(T) - M, 0). \quad (4.2)$$

Similarly the lookback rate put has almost the same payoff as the vanilla put, but with M replacing $S(T)$, i.e

$$\max(K - M, 0), \quad (4.3)$$

and naturally the corresponding call option then has the payoff

$$\max(M - K, 0), \quad (4.4)$$

where M is defined as

$$M = \begin{cases} \max_{0 \leq t \leq T}(S(t)), & \text{for a call} \\ \min_{0 \leq t \leq T}(S(t)), & \text{for a put.} \end{cases} \quad (4.5)$$

As for the barrier option there also exists discrete variants of the lookback options mentioned above. If the maximum (or minimum) is measured at the discrete times t_i , then the updating rule is

$$M_i = \begin{cases} \max(S(t_i), M(t_{i-1})), & \text{for a call,} \\ \min(S(t_i), M(t_{i-1})), & \text{for a put.} \end{cases} \quad (4.6)$$

4.2 Review of Literature

Closed-form solutions have been obtained for continuous sampled lookback option prices by [10], [17], and [18]. For the discretely sampled lookback option one has to rely on numerical methods. Most of them are based on some binomial method, see for example [2], [3], [8], or [20]. A PDE approach is described by Wilmott, Dewynne and Howison in Chapter 12 of [29]. They show that the price of lookback options is given as the solution to the ordinary Black-Scholes equation but with the maximum of the asset price entering as a parameter, and with different boundary and final conditions. They also show that for the floating strike lookback option it is possible to reduce the dimension of the problem by a change of variables, so that the price is a function of one state-variable and time. The same change of variables was previously used by Babbs, [2]. Andreasen, [1], uses a change of numeraire techniques to obtain option prices as function of time and a one-dimensional Markovian state variable only, for both the fixed and the floating strike lookback options, as well as Asian options. In [31] the same PDE-model as [29] is used but in a stochastic volatility setting.

4.3 Pricing Lookback Options

Lookback options fulfill the dream of every investor, selling at the highest or buying at the lowest price during the lifetime of the option. Naturally this makes lookback options expensive. Discrete sampling decreases the value of the contract and at the same time it is more natural to use. Therefore we concentrate on discrete lookbacks. For the floating strike lookback option we will use Wilmott, Dewynne and Howison's method. For a thorough analysis of this pricing method we refer to [29]. The fixed strike lookback option is examined in another paper by Foufas and Larson [14].

4.4 Floating Strike Lookbacks

Consider the discrete floating strike lookback option with monitoring dates $D = \{t_k^*\}_{k=0}^K \subset \{t_n\}_{n=0}^N$, where $t_0 = 0$ and $t_N = T$. Discrete sampling implies that we must have jump conditions across the sampling dates, since arbitrage considerations show that the realized value of the option cannot be discontinuous. For the value of discrete lookback option $V(S, M, t)$ we have the jump condition

$$V^-(S, M, t_k^*) = \begin{cases} V^+(S, \max(S, M), t_k^*), & \text{for a put,} \\ V^+(S, \min(S, M), t_k^*), & \text{for a call,} \end{cases} \quad (4.7)$$

across monitoring dates $t_k^* \in D$, where M is defined as

$$M = \begin{cases} \max_k(S(t_k^*)), & \text{for a put,} \\ \min_k(S(t_k^*)), & \text{for a call,} \end{cases} \quad (4.8)$$

As described by [29], M can be written as

$$M = \begin{cases} \lim_{n \rightarrow \infty} \left(\int_0^t f(\tau) S(\tau)^n \right)^{1/n}, & \text{for a call,} \\ \lim_{n \rightarrow \infty} \left(\int_0^t f(\tau) (1/S(\tau))^n \right)^{-1/n}, & \text{for a put,} \end{cases} \quad (4.9)$$

where

$$f(t) = \sum_k \delta_{t_k^*}(t), \quad (4.10)$$

and δ is the delta function.

Following [29], the pricing equation for the value of the lookback option is just the Black-Scholes partial differential equation. The independent variable M only enters as a parameter in the equation, but it also appears in the boundary and final conditions.

For the lookback strike put and call option we can reduce the dimension of the problem by introducing the similarity transformation (see [29])

$$\xi = S/M, \quad (4.11)$$

$$V(S, M, t) = Mu(\xi, t). \quad (4.12)$$

The partial differential equation for $u(\xi, t)$ then reads

$$\frac{\partial u}{\partial t} + \frac{1}{2}\sigma^2\xi^2\frac{\partial^2 u}{\partial \xi^2} + (r - \nu)\xi\frac{\partial u}{\partial \xi} - ru = 0, \quad (4.13)$$

and the final condition becomes

$$u(\xi, T) = u_T := \begin{cases} \max(\xi - 1), & \text{for a call,} \\ \max(1 - \xi), & \text{for a put.} \end{cases} \quad (4.14)$$

The jump condition across sampling dates $t_k^* \in D$ becomes

$$\begin{aligned} u^-(t_k^*) & \quad (4.15) \\ & = JC(u^+(t_k^*)) := \begin{cases} \max(\xi, 1)u^+(\min(\xi, 1), t_k^*), & \text{for a put,} \\ \min(\xi, 1)u^+(\max(\xi, 1), t_k^*), & \text{for a call.} \end{cases} \end{aligned}$$

The boundary condition at $\xi = 0$ are, (see [29])

$$u(0, t) = e^{-r(T-t)}, \quad (4.16)$$

for a put, and

$$u(0, t) = 0, \quad (4.17)$$

for a call. Concerning the boundary condition at $\xi \rightarrow \infty$ all one can, and need to, say is that the option value can grow at most linearly with ξ as $\xi \rightarrow \infty$ (see [29]).

4.4.1 The Finite Element Method

Again, we use the same discretization as for the vanilla option, and the artificial boundary condition $u_{\xi\xi} = 0$ at $\partial\Omega = \{\xi_{min}, \xi_{max}\}$. Since the pricing partial differential equation is the same as for the ordinary European option, the finite element method is same, except that we have to apply the jump conditions at each monitoring date. We let $U(t_n)$ denote the approximate solution at time step n , which fulfills jump conditions similar to (4.15). We thus want to solve the problem, find $U \in \mathcal{W}^q$ such that for $1 \leq n \leq N$

$$\begin{cases} \int_{I_n} (m(U_t, v) + a(U, v)) dt = 0 & \text{for all } v \in \mathcal{W}_n^{q-1}, \\ U^-(t_n) = U^+(t_n), & n = N-1, \dots, 1 \quad t_n \notin D, \\ U^-(t_n) = JC(U^+(t_n)), & t_n \in D, \\ U^-(t_N) = u_T, \end{cases} \quad (4.18)$$

where $m(U_t, v)$ and $a(U, v)$ are defined exactly as before.

4.4.2 Error Representation Formula

We now introduce the continuous dual problem for the floating strike look-back option, which differs only slightly from the barrier option case studied before. For clarity of exposition we here only give the details for the put option, the dual problem for the call option is almost the same.

$$\begin{cases} -\phi_t + (\sigma^2 + \nu - 2r)\phi - (r - \nu - 2\sigma^2)\xi\phi_\xi + \frac{\sigma^2}{2}\xi^2\phi_{\xi\xi} = 0, \\ \phi(0, \xi) = \delta_{\xi\alpha}, \\ \phi^+(t_k^*) = \begin{cases} \phi^-(t_k^*), & \xi < 1, t_k^* \in D, \\ \delta_1(\xi) \int_{\eta \geq 1} \phi^-(t_k^*) \eta d\eta, & \xi \geq 1, t_k^* \in D. \end{cases} \end{cases} \quad (4.19)$$

For simplicity we consider this equation over the whole space interval neglecting boundary conditions. Multiplying with the error $e = u - U \in \mathcal{W}$ and integrating in space and time we get

$$\begin{aligned} \sum_k \int_{t_{k-1}^*}^{t_k^*} \left(-(\phi_t, e) + (\sigma^2 + \nu - 2r)(\phi, e) \right. \\ \left. - (r - \nu - 2\sigma^2)(\xi\phi_\xi, e) + \frac{\sigma^2}{2}(\xi^2\phi_{\xi\xi}, e) \right) dt = 0. \end{aligned} \quad (4.20)$$

Moving derivatives from ϕ to the error e , this equation gives us the error representation formula

$$\boxed{e(0, s_\alpha) = - \sum_k \int_{t_{k-1}^*}^{t_k^*} (m(e_t, \phi) + a(e, \phi))} \quad (4.21)$$

where the bilinear forms $m(u_t, v)$ and $a(u, v)$ are defined exactly as before. We now present the details deriving this formula. Just as in the case of the barrier option we have to be extra careful with the first term (ϕ_t, e) in equation (4.20). Studying this term in detail we see that

$$\begin{aligned} - \sum_k \int_{t_{k-1}^*}^{t_k^*} (\phi_t, e) dt \\ = \sum_k \left(\int_{t_{k-1}^*}^{t_k^*} (\phi, e_t) dt \right) - (\phi(T), e(T)) + (\phi(t_0), e(t_0)) \\ - \sum_k ((\phi^-(t_k^*), e^-(t_k^*)) - (\phi^+(t_k^*), e^+(t_k^*))). \end{aligned} \quad (4.22)$$

Expanding the last two terms on the right we obtain

$$\begin{aligned} (\phi^-(t_k^*), e^-(t_k^*)) - (\phi^+(t_k^*), e^+(t_k^*)) \\ = \int_{\xi \geq 1} (\phi^-(t_k^*)e^-(t_k^*) - \phi^+(t_k^*)e^+(t_k^*)) d\xi, \end{aligned} \quad (4.23)$$

since $\phi^+(t_k^*) = \phi^-(t_k^*)$, and $e^-(t_k^*) = e^+(t_k^*)$ for $\xi < 1$, according to the boundary conditions for ϕ and U . Using the jump condition for ϕ we note that

$$\begin{aligned}
& \int_{\xi \geq 1} (\phi^-(t_k^*)e^-(t_k^*) - \phi^+(t_k^*)e^+(t_k^*)) d\xi & (4.24) \\
& = \int_{\xi \geq 1} \phi^-(t_k^*)e^-(t_k^*) d\xi - \int_{\xi \geq 1} \left(\delta_1(\xi) \int_{\eta \geq 1} \phi^-(t_k^*)\eta d\eta \right) e^+(t_k^*) d\xi \\
& = e^-(t_k^*, 1) \int_{\xi \geq 1} \phi^-(t_k^*)\xi d\xi - e^+(t_k^*, 1) \int_{\xi \geq 1} \phi^-(t_k^*)\xi d\xi = 0,
\end{aligned}$$

since $e^-(t_k^*, \xi) = \xi e^-(t_k^*, 1) = \xi e^+(t_k^*, 1)$ for $\xi \geq 1$. Finally, moving derivatives from ϕ to e in equation (4.20) using integration by parts, we arrive at the same error representation formula as for the European option and the barrier option studied before

$$e(0, s_\alpha) = - \sum_k \int_{t_{k-1}^*}^{t_k^*} (m(e_t, \phi) + a(e, \phi)), \quad (4.25)$$

where the bilinear forms $m(u_t, v)$ and $a(u, v)$ are defined exactly as before.

4.4.3 Examples

Using the same error estimation algorithm as in the previous cases we are able to calculate the error in desired quantities for different values of the parameters. This makes it possible to identify regions where a fine mesh is necessary.

Figures 17 and 16 show dual solutions for the daily and monthly sampled floating strike lookback put options respectively. In all cases we have used the boundary condition $\phi(0, s) = \delta_1(\xi, \epsilon)$, where $\epsilon = 5000$. We see that the sampling frequency has a significant effect on the dual solution.

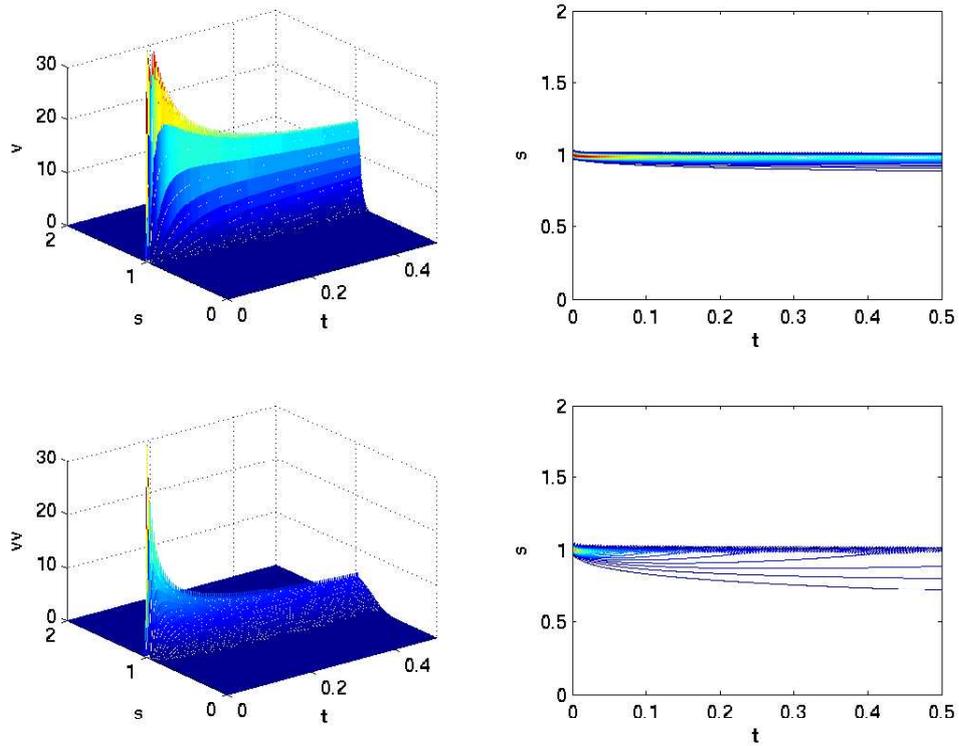


Figure 16: Above on the left, ϕ , for $\sigma = 0.1$, $r = 0.1$, and $q = 0.0$ with daily sampling. Below on the left, ϕ , for $\sigma = 0.3$. On the right, contour plots using 30 levels. Solutions computed using the $cG(2)$ - $dG(1)$ method with 200 space points and 400 time points.

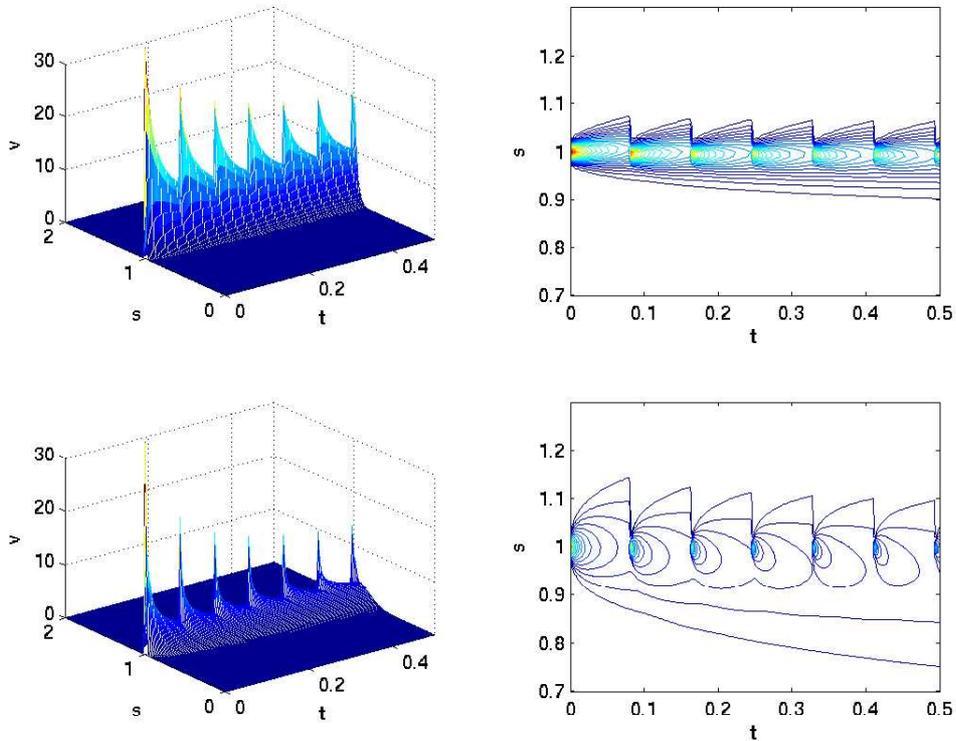


Figure 17: Above on the left, ϕ , for $\sigma = 0.1$, $r = 0.1$, and $q = 0.0$, with monthly sampling. Below on the left, ϕ , for $\sigma = 0.3$. On the right, contour plots using 30 levels. Solutions computed using the $cG(2)$ - $dG(1)$ method with 200 space and time points.

In Figure 18, we see the contributions to the error representation formula (4.21) from each space-time slab. The dual was calculated using the $cG(2)$ - $dG(1)$ method, and the primal using the $cG(1)$ - $cG(1)$ method. The dual mesh was thirty two times finer in each direction. The value of the functional of the error found by using the error representation formula was in this case 0.0042. We also note that the contribution to the error differs from zero only within a short interval of Ω , just as the dual solution. We now proceed to calculate adaptive meshes.

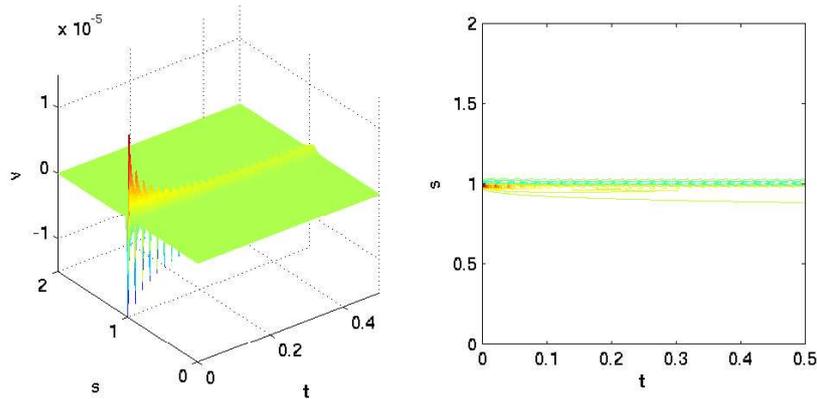


Figure 18: *The contributions to the functional of the error for the weekly sampled floating strike lookback put when $\sigma = 0.1$, $r = 0.10$ and $q = 0.0$. The dual was computed using the $cG(2)$ - $dG(1)$ method with 800 space and time points, and the primal using the $cG(1)$ - $dG(1)$ method with 25 space and time points.*

4.4.4 Adaptive Mesh Refinement

As in the case of the European option we use the error representation formula to derive an optimal mesh for each problem. In Figure 19, we see a mesh resulting from using the mesh refinement algorithm in the case of a floating strike lookback put option with weekly sampling. In this case Q was set to 10%. Two successive refinements were made, starting from a sparse mesh with 25 nodes in time and space. The final mesh has only 30 nodes in each direction, but the functional of the error has decreased by a factor 14. The dual was calculated using a fine mesh with 800 nodes in time and space.

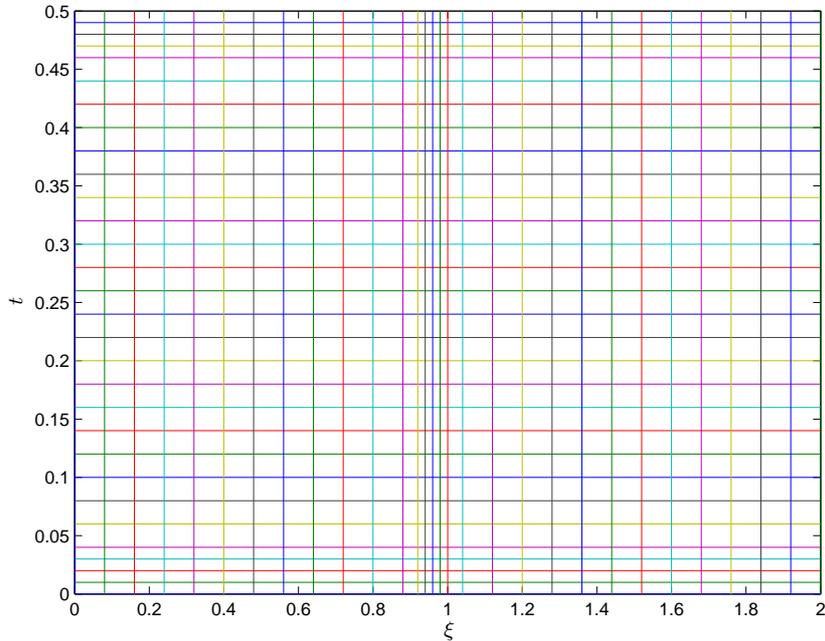


Figure 19: *The resulting mesh using the mesh refinement algorithm, calculated for a floating strike lookback put option with $T = 0.5$, $\sigma = 0.1$, and $r = 0.1$, when $\psi = \delta_1(\xi, \epsilon)$. The dual was computed using the cG(2)-cG(1) and the primal using the cG(1)-dG(1) method. Two successive refinements were made.*

4.4.5 The Greeks

Using the chain rule and equations (4.12) and (4.12) we can derive expressions for the Greeks in our new variable ξ

$$\Delta = \frac{\partial V}{\partial s} = \frac{\partial u}{\partial \xi} \quad (4.26)$$

$$\Gamma = \frac{\partial^2 V}{\partial s^2} = \frac{1}{M} \frac{\partial^2 u}{\partial \xi^2}, \quad (4.27)$$

$$\Theta = -\frac{\partial V}{\partial t} = -\frac{\partial u}{\partial t}. \quad (4.28)$$

In Figure 20 we see the delta and gamma of a weekly sampled floating strike lookback put at time $t = 0$.

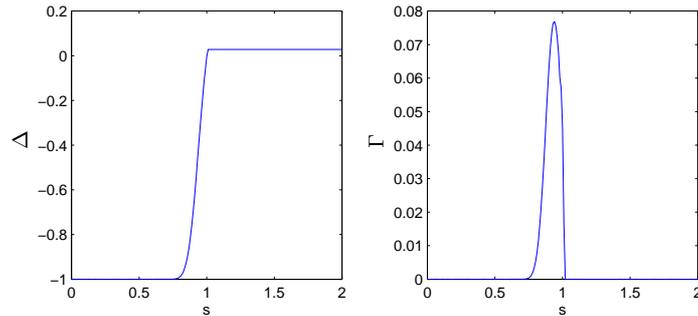


Figure 20: The delta (on the left) and gamma (on the right) of a weekly sampled floating strike lookback put at time $t = 0$, with $\sigma = 0.1$, $T = 1$, $S_0 = 100$ and $r = 0.1$.

4.4.6 Results

Table 3 compares the method used in this work with the ones in [29] and [30], for the discrete floating strike lookback put. All three methods give fairly the same result. Note that the values given in the table are not prices, but values of U . To get the option price we use equation (4.12).

	Wilmott	ZFV(389)	FE(400)
$\xi = 0.9$	0.101	0.10025	0.10035
$\xi = 1.0$	0.089	0.08885	0.08887
$\xi = 1.1$	0.095	0.09546	0.09502

Table 3: Comparison of discrete floating strike lookback put option values when $r = 0.1$, $q = 0.0$, $\sigma = 0.2$, $T = 1.0$, and $t = 0$. The number of nodes in the ξ direction is given in parenthesis. Wilmott refers to [29], ZFV refers to [30], and FE refers to the finite element method used in this work. Sampling was made at times 0.5, 1.5, 2.5, ..., 10.5, 11.5 months.

By using the mesh refinement algorithm an adapted mesh was calculated for the example of the weekly sampled lookback put option, which was finer close to the center of Ω and towards time $t = 0$ and $t = T$. In this way the same accuracy is achieved in less degrees of freedom. The original uniform mesh has 25 nodes in time and space. By using the error representation formula the functional of the error was calculated to 0.0042 for the uniform mesh. The adapted mesh has only 30 nodes in each direction, but the functional of the error has decreased by a factor 14 to 0.000305.

Figure 21 shows the value of U for the floating strike lookback put with monthly sampling.

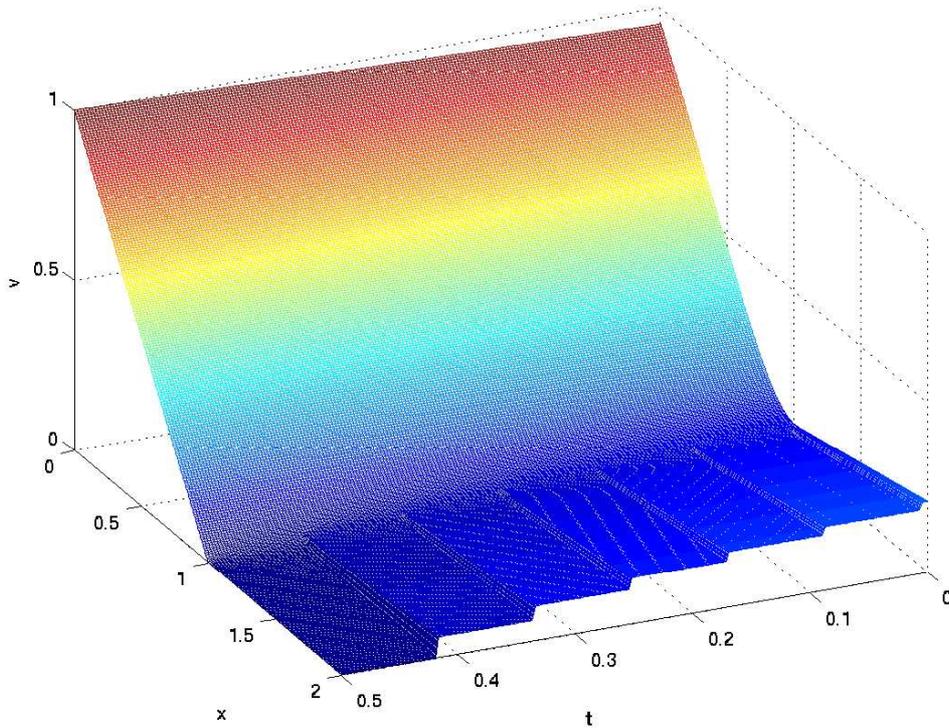


Figure 21: *The floating strike lookback put option with monthly sampling, when $\sigma = 0.3$, $r = 0.1$, $q = 0.0$, $T = 0.5$, and $t = 0$. Solution computed using the $cG(1)$ - $cG(1)$ method with a uniform mesh with 200 space points and time points.*

References

- [1] J. ANDREASEN, *The pricing of discretely sampled Asian and lookback options: a change of numeraire approach*, The Journal of Computational Finance, 2 (Fall 1998), pp. 5–30.
- [2] S. BABBS, *Binomial valuation of lookback options*. Working paper, Midland Montague, Capital Markets, London, 1992.
- [3] J. BARRAQUAND AND T. PUDET, *Pricing of of American path-dependent contingent claims*, Math. Fin., (1996), pp. 17–51.
- [4] F. BLACK AND M. SCHOLES, *The pricing of options and corporate liabilities.*, Journal of Political Economy, (1973), pp. 637–654.
- [5] P. BOYLE AND I. LAU, *Bumping up against the barrier with the binomial method.*, Journal of Derivatives, (1994), pp. 6–14.
- [6] M. BROADIE, P. GLASSERMAN, AND S. KOU, *Connecting discrete and continuous path-dependent options*, Finance Stochastics, (1999), pp. 55–82.
- [7] P. CARR, *Two extensions to barrier option evaluation.*, Applied Mathematical Finance, (1995), pp. 173–209.
- [8] T. CHEUK AND T. VORST, *Lookback options and the observation frequency: a binomial approach*. Working paper, Erasmus University, 1994.
- [9] ———, *Complex barrier options.*, Journal of Derivatives, (1996), pp. 8–22.
- [10] A. CONZE AND R. VISWANATHAN, *Path-dependent options: The case of lookback options.*, Journal of Finance, (1991), pp. 1893–1907.
- [11] K. ERIKSSON, D. ESTEP, P. HANSBO, AND C. JOHNSON, *Computational Differential Equations*, Studentlitteratur, 1996.
- [12] D. ESTEP, M. LARSON, AND R. WILLIAMS, *Estimating the error of numerical solutions of systems of reaction-diffusion equations*, MEMOIRS of the American Mathematical Society, 146 (2000).
- [13] S. FIGLEWSKI AND B. GAO, *The adaptive mesh model: a new approach to efficient option pricing*. Working paper, Stern School of Business, New York University., 1997.

- [14] G. FOUFAS AND M. G. LARSON, *Valuing fixed strike lookback options using the finite element method and duality techniques*. 2007.
- [15] ———, *Valuing asian options using the finite element method and duality techniques*, to appear in JCAM, (2008).
- [16] H. GEMAN AND M. YOR, *Pricing and hedging double-barrier options: a probabilistic approach*, Math. Finance, 6 (1996), pp. 365–378.
- [17] M. GOLDMAN, H. SOSIN, AND M. GATTO, *Path-dependent options: buy at the low, sell at the high.*, Journal of Finance, (1979), pp. 1111–1128.
- [18] M. GOLDMAN, H. SOSIN, AND L. SHEPP, *On contingent claims that insure ex-post optimal stock market timing*, Journal of Finance, (1979), pp. 401–414.
- [19] P. HEYNEN AND H. KAT, *Discrete partial barrier options with a moving barrier*, Journal of Financial Engineering, (1996), pp. 199–209.
- [20] J. HULL AND A. WHITE, *Efficient procedures for valuing European and American path-dependent options*, J. Derivatives, (1993), pp. 21–31.
- [21] A. KOLKIEWICZ, *Pricing and hedging more general double barrier options*. Working paper, Department of Statistics and Actuarial Science, University of Waterloo.
- [22] N. KUNIMOTO AND M. IKEDA, *Pricing options with curved boundaries*, Mathematical Finance, (1992), pp. 275–298.
- [23] R. MERTON, *The theory of rational option pricing*, Bell Journal of Economics and Management Science., (1973), pp. 141–183.
- [24] M. REIMER AND K. SANDMANN, *A discrete time approach for European and American barrier options*. Working paper, Department of Statistics, Rheinische Friedrich-Wilhelms-Universität, Bonn, 1995.
- [25] D. RICH, *The mathematical foundations of barrier option-pricing theory*, Advances in Futures and Options Research, (1991), pp. 267–311.
- [26] P. RITCHKEN, *On pricing barrier options*, Journal of Derivatives, (1995), pp. 19–28.
- [27] M. RUBINSTEIN AND E. REINER, *Breaking down the barriers.*, RISK, 4 (1991), pp. 28–35.
- [28] P. WILMOTT, *Paul Wilmott on quantitative finance.*, Wiley, 2000.

- [29] P. WILMOTT, J. DEWYNNE, AND S. HOWISON, *Option pricing*, Oxford Financial Press, 1993.
- [30] R. ZVAN, P. FORSYTH, AND K. VETZAL, *PDE methods for pricing barrier options*. To appear in the Journal of Economic Dynamics and Control.
- [31] ———, *A finite element approach to the pricing of discrete lookbacks with stochastic volatility*, Applied Mathematical Finance, (1999), pp. 87–106.

Paper II

Valuing Fixed Strike Lookback Options using the Finite Element Method and Duality Techniques

Georgios Foufas* Mats G. Larson†

April 14, 2008

Abstract

The main objective of this paper is to develop an adaptive finite element method for computation of the values and different sensitivity measures of fixed strike lookback options.

The fixed strike lookback options are priced using the Black-Scholes PDE-model, and a method developed by Andreasen. It consists of solving two coupled PDE:s that are of parabolic type in one spatial dimension with different boundary conditions and jump conditions at monitoring dates. The adaptive finite element method is based on a posteriori estimates of the error in desired quantities, which we derive using duality techniques. The a posteriori error estimates are tested and verified, and are used to calculate optimal meshes for each type of option. The use of adapted meshes gives superior accuracy and performance with less degrees of freedom than using uniform meshes. The suggested adaptive finite element method is stable and gives fast and accurate results.

1 Introduction

The valuation of different types of derivative contracts is very important in modern financial theory and practice. Exotic options have become very popular hedging and speculation instruments in recent years. At the same time a huge amount of literature has been devoted to the pricing and hedging of such instruments.

*Research Assistant, Department of Mathematics, Chalmers University of Technology, S-412 96 Gothenburg, Sweden, *email*: foufas@math.chalmers.se

†Professor of Applied Mathematics, Corresponding author, Department of Mathematics, Umeå University, S-901 87 Umeå, Sweden, *email*: mats.larson@math.umu.se

Closed-form solutions have been obtained for continuous sampled lookback option prices by [5], [11], and [12]. For the discretely sampled lookback option one has to rely on numerical methods. Most of them are based on some binomial method, see for example [2], [3], [4], or [13]. A PDE approach is described by Wilmott, Dewynne and Howison in Chapter 12 of [15]. They show that the price of lookback options is given as the solution to the ordinary Black-Scholes equation but with the maximum of the asset price entering as a parameter, and with different boundary and final conditions. They also show that for the floating strike lookback option it is possible to reduce the dimension of the problem by a change of variables, so that the price is a function of one state-variable and time. The same change of variables was previously used by Babbs, [2]. Andreasen, [1], uses a change of numeraire techniques to obtain option prices as function of time and a one-dimensional Markovian state variable only, for both the fixed and the floating strike lookback options, as well as Asian options. In [16] the same PDE-model as [15] is used but in a stochastic volatility setting.

Classification: The lookback option has a payoff that depends on the maximum or minimum of the underlying stock price over some given interval in time. There are two types of lookback options, the lookback rate option and the lookback strike option, also known as the fixed strike and the floating strike lookback option respectively. Both of these options exists as puts and calls. If we denote the maximum asset price over the time interval $0 \leq t \leq T$ by M , the lookback strike put option has almost the same payoff as the vanilla put, but with M replacing the exercise price K , that is

$$\max(M - S(T), 0). \quad (1.1)$$

In the same way the corresponding call option gives almost the same payoff as the vanilla call, but with M , now being the minimum asset price during the lifetime of the option, replacing the exercise price K , that is

$$\max(S(T) - M, 0). \quad (1.2)$$

Similarly the lookback rate put has almost the same payoff as the vanilla put, but with M replacing $S(T)$, i.e

$$\max(K - M, 0), \quad (1.3)$$

and naturally the corresponding call option then has the payoff

$$\max(M - K, 0), \quad (1.4)$$

where M is defined as

$$M = \begin{cases} \max_{0 \leq t \leq T}(S(t)), & \text{for a call} \\ \min_{0 \leq t \leq T}(S(t)), & \text{for a put.} \end{cases} \quad (1.5)$$

As for the barrier option there also exists discrete variants of the lookback options mentioned above. If the maximum (or minimum) is measured at the discrete times t_i , then the updating rule is

$$M_i = \begin{cases} \max(S(t_i), M(t_{i-1})), & \text{for a call,} \\ \min(S(t_i), M(t_{i-1})), & \text{for a put.} \end{cases} \quad (1.6)$$

New Contributions: The fixed strike lookback options are priced using the Black-Scholes PDE-model and a method developed by Andreasen [1]. It consists of solving two coupled PDE:s that are of parabolic type in one spatial dimension with different boundary conditions and jump conditions at monitoring dates. All pricing is done using an adaptive finite element method allowing variable resolution in space and time.

In practice one is only interested in the price, and it's derivatives, in one or a few points. Using this criteria, the choice of computational mesh is based on a posteriori estimates of the error in desired quantities, which we derive using duality techniques. These dual techniques are shown to be very useful and simple, and allows us to improve the PDE methods already existing. The presented a posteriori error estimation formula is tested and verified in the case of the European option. It is then used to perform mesh refinements in both time and space for the fixed strike lookback options. This makes it possible to calculate an optimal mesh for each type of option, which significantly reduces the error without noticeably enhancing the computational effort.

The duality approach is general and applicable to problems with all kinds of algebraic constraints. Other exotic options, such as the barrier option, the floating strike lookback option, and the Asian option are also studied by the authors, see [9] and [10]. The suggested adaptive finite element method is stable and gives fast and accurate results.

Outline: In Section 2 we formulate the finite element method and derive an a posteriori error estimate for the ordinary European option. Then in Section 3 we derive an error representation formula for the European option, present and error estimation algorithm, and give some numerical examples. In Section 4 we perform adaptive mesh refinements for the European option based on the previous results. In Section 5 we present a pricing technique for the fixed strike lookback option and apply the finite element method. Section 6 includes a derivation of an a posteriori error representation formula for the fixed strike lookback option, and gives some examples were it is applied. Adaptive mesh refinement for the fixed strike lookback option is presented in Section 7. In Section 8 we present some results, and finally in Section 9 we state some conclusions.

2 An Adaptive Finite Element Method for the European Option

For the ordinary European option there exists an analytical valuation formula. But for other options, such as the discrete fixed strike lookback option, studied later in this paper, we have to rely on numerical solutions. In this section we present the finite element method and develop the a posteriori error estimation framework, for the basic European option. Later we extend the techniques to the fixed strike lookback option.

2.1 Mathematical Background

We consider a continuous time trading economy on a bounded time horizon $[0, T]$. Probability is represented by the probability space $(\Omega_T, \mathcal{F}_T, P)$, where $\Omega_T = C[0, T]$, P is the corresponding Wiener measure, and $\mathcal{F}_T = \sigma(W(t); t \leq T)$. For simplicity we consider the standard Black-Scholes setting with a risk free asset and a dividend paying stock. Let $B(t)$ denote the price of a risk free asset at time t governed by the equation $B(t) = B(0)e^{rt}$, where r is the constant interest rate. Further we denote by $S(t)$ the value of an asset at time t . We assume the existence of an equivalent martingale measure Q , under which the discounted stock price $e^{-r(T-t)}S_t$ is an \mathcal{F}_t -martingale. The existence of the risk neutral measure Q assures that the market is free of arbitrage possibilities. Under Q the stock price follows the stochastic differential equation

$$dS(t) = (r - \nu)S(t)dt + S(t)\sigma dW(t), \quad (2.1)$$

where r is the constant interest rate, ν is the constant continuous dividend yield, σ is the volatility, and $W(t)$ is a Q Brownian motion process. Here σ is assumed to be a positive real number. The solution of (2.1) is

$$S(t) = S(0)e^{(r-\nu-\frac{\sigma^2}{2})t+\sigma W(t)}. \quad (2.2)$$

2.2 The Black-Scholes PDE

The value of the ordinary European option, $u(t, S(t)) = u(t, s)$, is given as the solution to Black-Scholes equation

$$u_t(t, s) + \frac{\sigma^2 s^2}{2}u_{ss}(t, s) + (r - \nu)su_s(t, s) - ru(t, s) = 0, \quad t < T, \quad (2.3)$$

which is valid for $s = S(t) \in \mathbb{R}^+$. In order to construct a computational mesh we introduce a bounded interval $\Omega = [s_{min}, s_{max}] \subset \mathbb{R}^+$ with boundary

$\partial\Omega = \{s_{min}, s_{max}\}$. We define the usual Hilbert space

$$H^1(\Omega) = \{v : \int_{\Omega} (|\nabla v|^2 + v^2) ds < \infty\}, \quad (2.4)$$

and let \mathcal{W} be the space of functions that are square integrable in time and belongs to $H^1(\Omega)$ in space, that is

$$\mathcal{W} = L^2([0, T], H^1(\Omega)). \quad (2.5)$$

We also use the notation $(u, v) = \int_{\Omega} uv ds$, and $(u, v)_{\partial\Omega} = u(s_{max})v(s_{max}) - u(s_{min})v(s_{min})$.

2.3 Variational Formulation

Multiplying the Black-Scholes equation (2.3) by the test function $v \in \mathcal{W}$ and integrating on $\Omega \times [0, T]$ we obtain

$$\int_0^T \left((u_t, v) + (r - \nu)(su_s, v) + \frac{\sigma^2}{2} (s^2 u_{ss}, v) - r(u, v) \right) dt = 0. \quad (2.6)$$

Using integration by parts we get

$$(s^2 u_{ss}, v) = (s^2 u_s, v)_{\partial\Omega} - 2(su_s, v) - (s^2 u_s, v_s). \quad (2.7)$$

Thus equation (2.6) becomes

$$\int_0^T \left((u_t, v) + (r - \nu - \sigma^2)(su_s, v) - \frac{\sigma^2}{2} (s^2 u_s, v_s) + \frac{\sigma^2}{2} (s^2 u_s, v)_{\partial\Omega} - r(u, v) \right) dt = 0. \quad (2.8)$$

The boundary conditions for the European call option are $u(t, 0) = 0$ and $u(t, s) \sim se^{-\nu(T-t)}$ as $s \rightarrow \infty$, and for the corresponding put $u(t, 0) = Ke^{-r(T-t)}$ and $u(t, s) \sim 0$ as $s \rightarrow \infty$, see for example Wilmott, [15]. For simplicity of implementation we use the artificial boundary condition $u_{ss} = 0$ on $\partial\Omega$ for both the put and the call instead. This boundary condition works for all contracts if the payoff is at most linear in the underlying (see [15]) and does not affect the accuracy of the solution. Using equation (2.3) we can rewrite the boundary condition as

$$u_s = \frac{r}{s(r - \nu)} u - \frac{1}{s(r - \nu)} u_t, \quad (2.9)$$

and enforce it weakly by inserting identity (2.9) into equation (2.8). We thus want to solve the problem: find $u \in \mathcal{W}$ such that

$$\begin{cases} \int_0^T (m(u_t, v) + a(u, v)) dt = 0, \\ u(T, s) = \begin{cases} \max(s - K, 0), & \text{for a call,} \\ \max(K - s, 0), & \text{for a put,} \end{cases} \end{cases} \quad (2.10)$$

for every $v \in \mathcal{W}$, where

$$m(u_t, v) = (u_t, v) - \frac{\sigma^2}{2(r - \nu)} (su_t, v)_{\partial\Omega}, \quad (2.11)$$

and

$$\begin{aligned} a(u, v) &= (r - \nu - \sigma^2)(su_s, v) - \frac{\sigma^2}{2}(s^2u_s, v_s) \\ &\quad + \frac{\sigma^2 r}{2(r - \nu)}(su, v)_{\partial\Omega} - r(u, v). \end{aligned} \quad (2.12)$$

2.4 Finite Element Approximation

The finite element method is based on solution of the variational problem (2.10) with \mathcal{W} replaced by a finite dimensional function space of piecewise polynomials in space and time. For background on the finite element method see for instance [6].

We now partition $[0, T]$ as $0 = t_0 < t_1 < t_2 < \dots < t_N = T$, denoting each time interval by $I_n = (t_{n-1}, t_n]$ and each time step by $k_n = t_n - t_{n-1}$. Similarly we partition Ω as $s_{min} = s_0 < s_1 < s_2 < \dots < s_J = s_{max}$, denoting each spatial interval by $\kappa_j = [s_{j-1}, s_j]$ and the length of each interval by $h_j = s_j - s_{j-1}$.

In space, we let $\mathcal{V}^p \subset H^1(\Omega)$ denote the space of piecewise continuous functions of order p . On each space-time slab $S_n = I_n \times \Omega$, we define

$$\mathcal{W}_n^q = \{w(t, s) : w(t, s) = \sum_{j=0}^q t^j v_j(s), v_j \in \mathcal{V}^p, (t, s) \in S_n\}. \quad (2.13)$$

Let $\mathcal{W}^q \subset \mathcal{W}$ denote the space of functions defined on $[0, T] \times \Omega$ such that $v|_{S_n} \in \mathcal{W}_n^q$ for $1 \leq n \leq N$. For simplicity, we only give details for the continuous Galerkin method cG(p)-cG(q), (see e.g. [6] or [7]) which is defined by the following discrete version of equation (2.10). Find $U \in \mathcal{W}^q$ such that for $1 \leq n \leq N$

$$\begin{cases} \int_{I_n} (m(U_t, v) + a(U, v)) dt = 0 & \text{for all } v \in \mathcal{W}_n^{q-1}, \\ U^-(t_n) = U^+(t_n), & n = N - 1, \dots, 1, \\ U^-(t_N) = u_T, \end{cases} \quad (2.14)$$

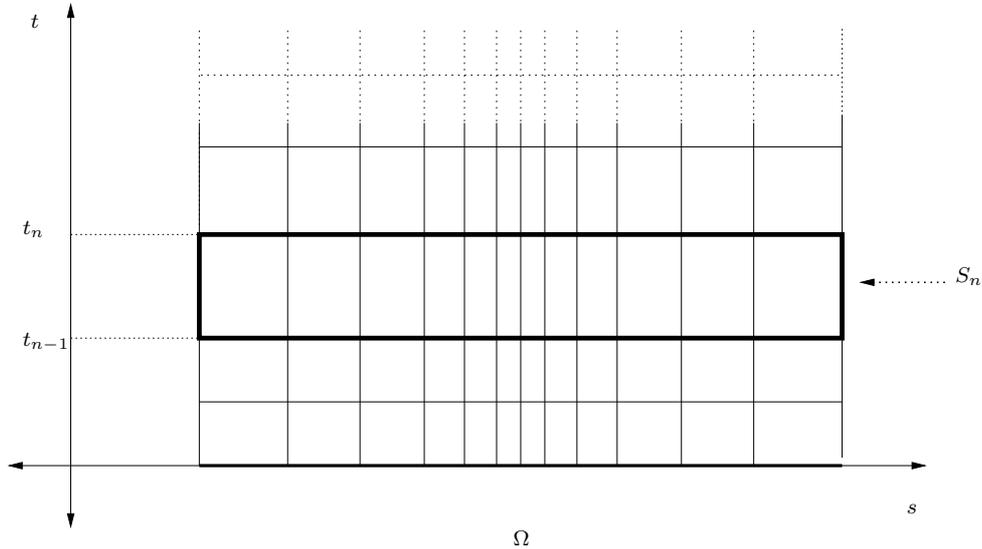


Figure 1: *Space-time discretization.*

where $U^\pm(t_n) = \lim_{\epsilon \rightarrow 0, \epsilon > 0} U(t_n \pm \epsilon)$. In the cG(1) method the approximation U of u is continuous piecewise linear in time and space, while the test functions v are continuous linear in space and piecewise constant in time. It is also possible to use a discontinuous method in time, we refer to [6], for details on the resulting discontinuous Galerkin method, cG(p)-dG(q).

3 A Posteriori Error Estimation for the European Option

3.1 Error Representation Formula

Since we are only interested in the solution, and its derivatives, in one or a few points of Ω at time $t = 0$, we wish to find a mesh tailored for efficient and accurate solution at the points of interest. In order to find such a mesh we derive a posteriori error estimates of the error in the points of interest using duality techniques (see [6] or [7]).

To represent the error in a linear functional, $(u - U, \psi)$, we introduce the continuous dual problem for the Black-Scholes equation (2.3). Find $\phi \in \mathcal{W}$

such that

$$\begin{cases} -\phi_t + (\sigma^2 + \nu - 2r)\phi - (r - \nu - 2\sigma^2)s\phi_s + \frac{\sigma^2}{2}s^2\phi_{ss} = 0, \\ \phi(0, s) = \psi. \end{cases} \quad (3.1)$$

For simplicity we consider this equation over the whole space interval neglecting boundary conditions. Multiplying with the error $e = u - U \in \mathcal{W}$ and integrating in space and time we get

$$\int_0^T \left(-(\phi_t, e) + (\sigma^2 + \nu - 2r)(\phi, e) - (r - \nu - 2\sigma^2)(s\phi_s, e) + \frac{\sigma^2}{2}(s^2\phi_{ss}, e) \right) dt = 0. \quad (3.2)$$

The functions ϕ and ϕ_s are in principle zero close to $s = s_{min}$ and $s = s_{max}$ if the domain is large enough. Using integration by parts and neglecting the boundary terms we get

$$\begin{aligned} & -(\phi(T, s), e(T, s)) + (\phi(0, s), e(0, s)) \\ & + \int_0^T \left((\phi, e_t) + (\sigma^2 + \nu - 2r)(\phi, e) + (r - \nu - 2\sigma^2)(s\phi, e_s) \right) dt \\ & + \int_0^T \left((r - \nu - 2\sigma^2)(\phi, e) - \frac{\sigma^2}{2}(s^2\phi_s, e_s) - \sigma^2(s\phi_s, e) \right) dt = 0. \end{aligned} \quad (3.3)$$

Note that integration by parts gives

$$-\sigma^2(s\phi_s, e) = \sigma^2(s\phi, e_s) + \sigma^2(\phi, e), \quad (3.4)$$

using this identity, $\phi(0, s) = \psi$, and $e(T) = 0$, we get

$$\begin{aligned} & (\psi, e(0, s)) = \\ & - \int_0^T \left((\phi, e_t) - r(\phi, e) + (r - \nu - \sigma^2)(s\phi, e_s) - \frac{\sigma^2}{2}(s^2\phi_s, e_s) \right) dt. \end{aligned} \quad (3.5)$$

Recalling the earlier defined bilinear forms (2.11) and (2.12), and that we neglect the boundary terms we can also write

$$(\psi, e(0, s)) = - \int_0^T \left(m(e_t, \phi) + a(e, \phi) \right) dt. \quad (3.6)$$

Since $e = u - U$ and u solves equation (2.10) we get the error representation formula

$$\boxed{(\psi, e(0, s)) = \int_0^T \left(m(U_t, \phi) + a(U, \phi) \right) dt} \quad (3.7)$$

If we for example are interested in the error at $s = s_\alpha$, we choose $\psi = \delta_{s_\alpha}(s)$, and get the error representation formula

$$e(0, s_\alpha) = \int_0^T \left(m(U_t, \phi) + a(U, \phi) \right) dt. \quad (3.8)$$

If one instead is interested in derivatives of the solution, then a different ψ is chosen, as shown later on.

3.2 Estimating the Error

Let $\pi : \mathcal{W} \rightarrow \mathcal{W}^{q-1}$ be the L_2 projection in time, and let P be a suitable interpolation operator into \mathcal{V}^p in space. Thus πP is an interpolation operator such that $\pi P\phi \in \mathcal{W}^{q-1}$. Then using Galerkin orthogonality (2.14), we can replace ϕ by $\phi - \pi P\phi = \phi - P\phi + P\phi - \pi P\phi$. Equation (3.7) can then be written as

$$\begin{aligned} (\psi, e(0, s)) &= - \int_0^T \left(m(U_t, \phi - P\phi) + a(U, \phi - P\phi) \right) dt \quad (3.9) \\ &\quad - \int_0^T \left(m(U_t, P\phi - \pi P\phi) + a(U, P\phi - \pi P\phi) \right) dt \\ &= - \sum_n \sum_j \int_{I_n} \left(R_{\kappa_j}^s(U), \phi - P\phi \right) dt \\ &\quad - \sum_n \int_{I_n} \left(R^t(U), P\phi - \pi P\phi \right) dt, \end{aligned}$$

where

$$\begin{aligned} (R_{\kappa_j}^s(U), \phi - P\phi) &= - \frac{\sigma^2}{2} (s^2 [U_s], \phi - P\phi)_{\partial\kappa_j} \quad (3.10) \\ &\quad + (U_t + (r - \nu)sU_s + \frac{\sigma^2}{2}s^2U_{ss} - rU, \phi - P\phi)_{\kappa_j} \end{aligned}$$

is the space residual, and

$$(R^t(U), P\phi - \pi P\phi) = (U_t + (r - \nu)sU_s + \frac{\sigma^2}{2}s^2U_{ss} - rU, P\phi - \pi P\phi) \quad (3.11)$$

is the time residual. Here we used the notation $[U_s]$ to denote the jump in U_s over element interfaces.

Finally, we present an algorithm for calculating the error.

Error Estimation Algorithm:

- Compute an approximation Φ of ϕ using an enriched finite element space, for instance higher order approximation.
- Compute $P\Phi$.
- Compute $\int_{I_n} \left(R_{\kappa_j}^s(U), \phi - P\phi \right) dt$ using quadrature in space and time for each element κ_j and time step.
- Compute $\pi P\Phi$.
- Compute $\int_{I_n} \left(R^t(U), P\phi - \pi P\phi \right) dt$ using quadrature in space and time for each time step.

3.3 Examples

Using the error estimation algorithm in the previous section we are able to calculate the error in desired quantities for different values of the parameters. This makes it possible to identify regions where a fine mesh is necessary.

Example 1. To estimate the error at $s = s_\alpha$ we let $\psi = \delta_{s_\alpha}(s)$ in (3.1). In order to implement this condition we use the approximation

$$\delta_{s_\alpha}(s) \approx \frac{1}{\epsilon\sqrt{\pi}} e^{-((s-s_\alpha)/\epsilon)^2} := \delta_{s_\alpha}(s, \epsilon), \quad (3.12)$$

where ϵ is a parameter that controls how well the delta function is approximated. In this example we have used $\epsilon = 1$. As seen from Figure 2, the solution to the dual problem differs from zero only within a short interval of Ω .

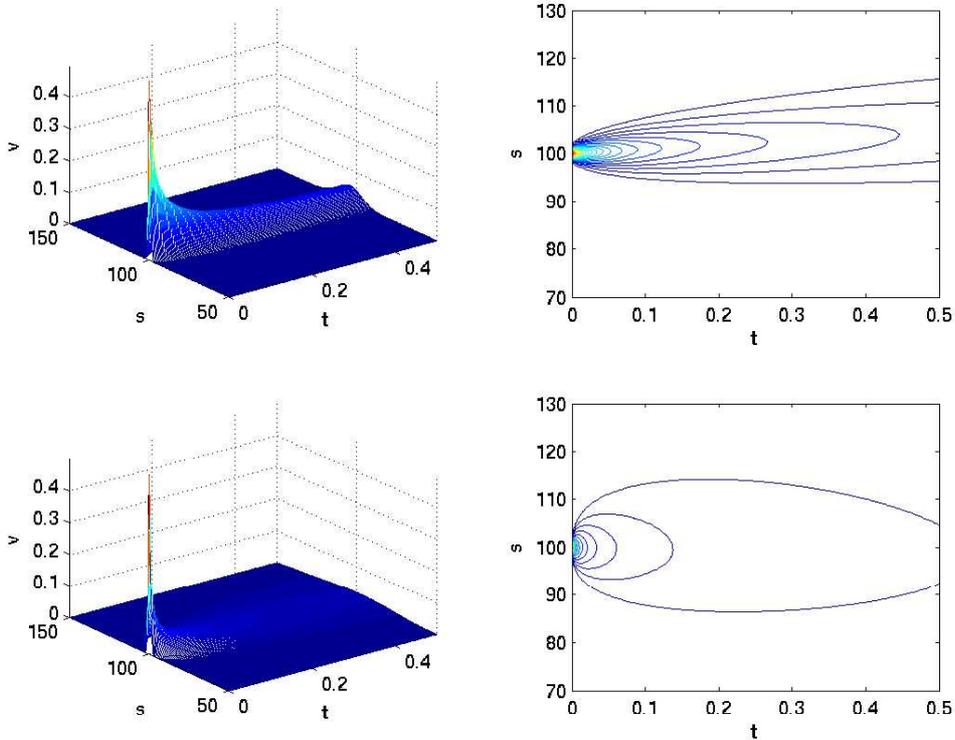


Figure 2: Above on the left, ϕ for a European call option when $\sigma = 0.1$, $r = 0.1$, and $\psi = \delta_{100}(s, \epsilon)$. Below on the left, ϕ , for $\sigma = 0.3$. On the right, contour plots using 30 levels. Solutions computed using the cG(2)-dG(1) method with 200 space and time points.

We now check that the error representation formula really works. By using the error estimation algorithm in the previous section we can get an approximation of the functional of the error, that is an approximation of the right hand side of equation (3.7). This can then be compared to calculating the left hand side of equation (3.7) directly using the real error in the approximate solution, found by using Black-Scholes formula. The dual solution is calculated on a finer mesh, and using higher order approximations. In Figure 3, we see the contributions to error formula (3.7) from each space-time slab. The dual was calculated using the cG(2)-dG(1) method, and the primal using the cG(1)-cG(1) method. The dual mesh was thirty-two times finer in each direction. The value of the functional of the error found by using the error representation formula was in this case 0.2033, in excellent agreement with the real value, that is the value of the left hand

side of equation (3.7), which was 0.2030. We also note that the contribution to the error differs from zero only within a short interval of Ω , just as the dual solution. This means that we may use a more sparse mesh where the contribution to the error is small and thus save computation time. The solution is larger near time $t = 0$, implying that one should use a finer time step there. Obviously the result depends on the value of the volatility σ , and the other parameters, which can be seen from the plot of the dual solution. We will later see how we can use the error representation formula to derive an optimal mesh for each problem.

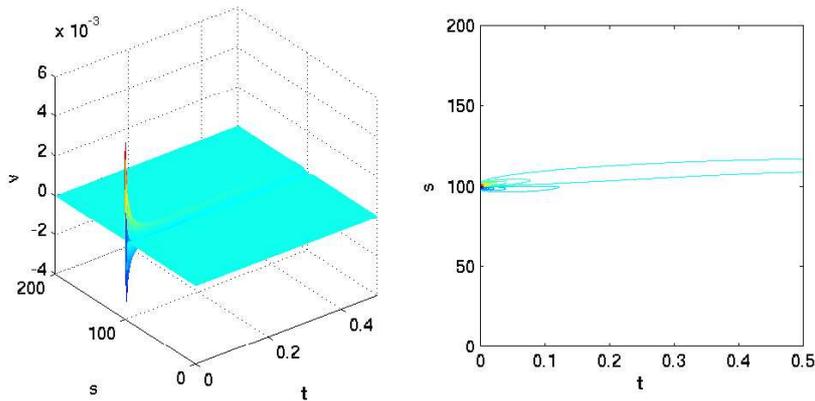


Figure 3: On the left, the contributions to the error of call option for $\sigma = 0.1$, $r = 0.1$, and $K = 100$ when $\psi = \delta_{100}(s, \epsilon)$. On the right, contour plot using 30 levels. The dual was computed using the $cG(2)$ - $dG(1)$ method with 400 space and time points, and the primal using the $cG(1)$ - $cG(1)$ method with 20 space and time points.

Example 2. In order to make a good estimation of the derivative of the solution, which is interesting when calculating the Greek *delta*, we need to study a different dual problem. We approximate the derivative using the

central difference formula

$$\frac{\partial u}{\partial s} \approx \frac{u(s + \mu) - u(s - \mu)}{2\mu} := \frac{\partial_h u}{\partial s}. \quad (3.13)$$

To estimate the error of the derivative of the solution at $s = s_\alpha$, $u_s(s_\alpha)$, we thus choose

$$\begin{aligned} \psi(s) &= \frac{\delta_{s_\alpha}(s - \mu) - \delta_{s_\alpha}(s + \mu)}{2\mu} \\ &\approx \frac{\delta_{s_\alpha}(s - \mu, \epsilon) - \delta_{s_\alpha}(s + \mu, \epsilon)}{2\mu} \end{aligned} \quad (3.14)$$

in (3.1), for an appropriate choice of μ . The error in our estimation of the derivative can be split into two parts

$$\left(\frac{\partial u}{\partial s} - \frac{\partial_h U}{\partial s} \right) = \left(\frac{\partial u}{\partial s} - \frac{\partial_h u}{\partial s} \right) + \left(\frac{\partial_h u}{\partial s} - \frac{\partial_h U}{\partial s} \right). \quad (3.15)$$

The first term corresponds to the error in (3.13), while the second can be estimated using the a posteriori estimate. Figure 4 shows the dual solution for this choice of ψ when $\mu = 1$ and $\epsilon = 1$. Figure 5 shows the contributions to the error estimation formula from each space-time slab. We see that this solution is even more centrally oriented than the previous one, implying that the derivative has a local dependence.

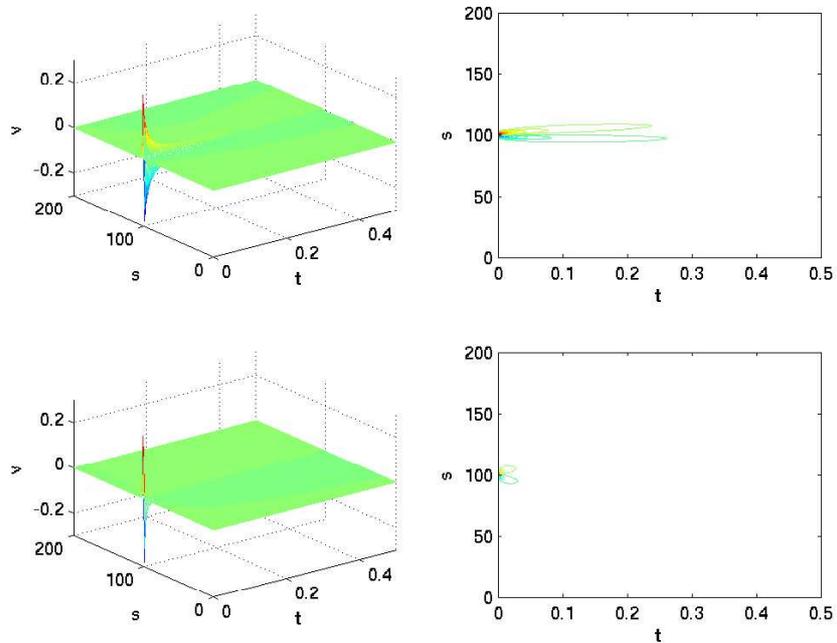


Figure 4: Above on the left, ϕ for a European call option with $\sigma = 0.1$, and $r = 0.1$, when ψ is chosen as in example 2. Below on the left, ϕ , for $\sigma = 0.3$. On the right, contour plots using 30 levels. Solutions computed using the $cG(2)$ - $dG(1)$ method with 200 space and time points.

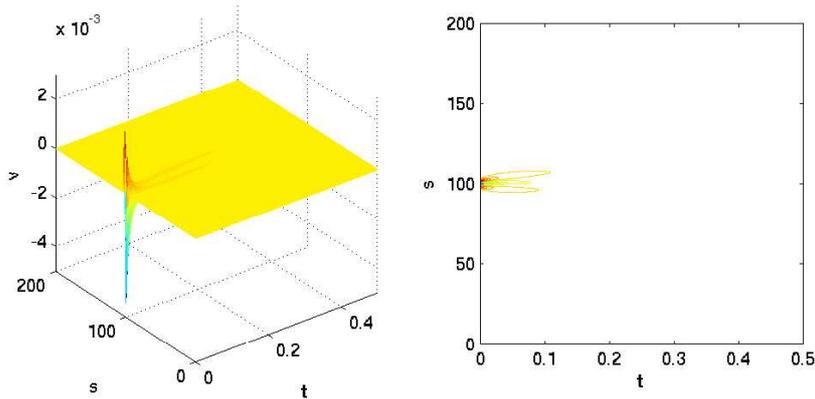


Figure 5: On the left, the contributions to the functional of the error of European call option for $\sigma = 0.1$, $r = 0.1$, and $K = 100$ when ψ is chosen as in example 2. On the right, contour plot using 30 levels. The dual was computed using the $cG(2)$ - $dG(1)$ method with 640 space and time points, and the primal using the $cG(1)$ - $cG(1)$ method with 20 space and time points.

4 Adaptive Mesh Refinement for the European option

Adaptive mesh refinement may be accomplished in many different ways. Our goal not is to create the best adaptive method, since adaptivity would be too slow to use in reality. Rather we wish to create an optimal mesh in advance for each case, so that when valuing an option we simply use a suited pre calculated mesh. This gives superior performance. In this section we show how these meshes are calculated and what typical meshes look like.

Mesh Refinement Algorithm:

- Compute an approximation U of u using the FE method on a coarse mesh.

- Compute the error in desired quantities by using the a posteriori error estimation algorithm.
- Calculate the time and space averages of the contributions to the error from each space-time slab. This gives us two vectors, one with time averages and one with space averages.
- Identify the $Q\%$ largest elements in the space average vector, and refine the corresponding time steps by dividing them in half.
- Identify the $Q\%$ largest elements in the time average vector, and refine the corresponding spatial steps by dividing them in half.
- Compute a new FE approximation U on the refined mesh.
- Repeat until minimum mesh size is reached.

In Figure 6, we see a typical mesh resulting from using the mesh refinement algorithm above. In this case Q was set to 10%. Three successive refinements were made, starting from a sparse mesh with 20 nodes in time and space. The final mesh has only 27 nodes in each direction, but the error has decreased by a factor 70. The dual was calculated using a fine mesh with 640 nodes in time and space.

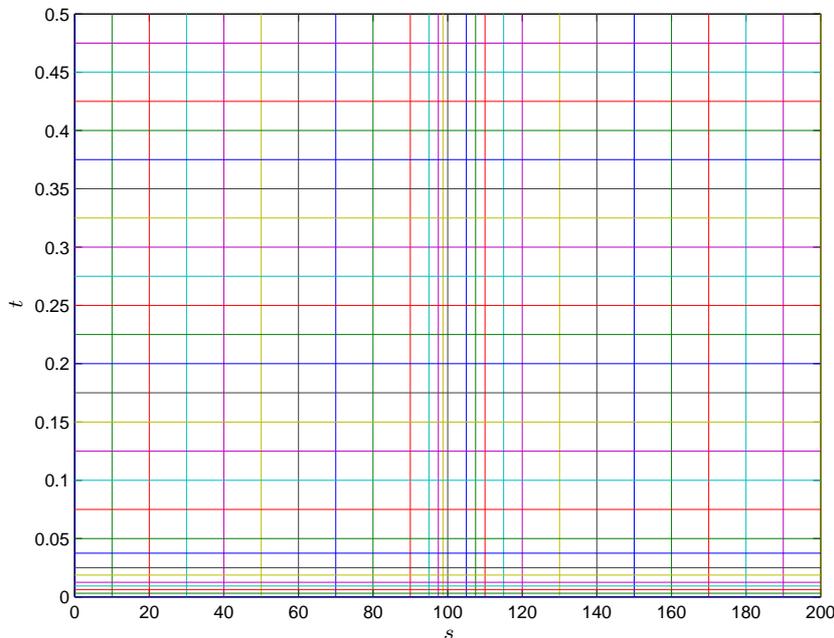


Figure 6: *The resulting mesh using the mesh refinement algorithm, calculated for a European call option with $\sigma = 0.1$, $r = 0.1$, and $K = 100$ when $\psi = \delta_{100}(s, \epsilon)$. The dual was computed using the $cG(2)$ - $dG(1)$ and the primal using the $cG(1)$ - $cG(1)$ method. Three successive refinements were made.*

5 An Adaptive Finite Element Method for the Fixed Strike Lookback Option

Lookback options fulfill the dream of every investor, selling at the highest or buying at the lowest price during the lifetime of the option. Naturally this makes lookback options expensive. Discrete sampling decreases the value of the contract and at the same time it is more natural to use. Therefore we concentrate on discrete fixed strike lookbacks. The floating strike lookback option is treated in another paper by the authors, [9].

5.1 Pricing Partial Differential Equations

Here we will give a brief presentation of Andreasen's method for pricing of the fixed strike lookback option, for further details we refer to [1].

Consider the discrete fixed strike looback call option with monitoring dates $D = \{t_k^*\}_{k=0}^N \subset \{t_n\}_{n=0}^N$, where $t_0 = 0$ and $t_N = T$. This option gives the payoff

$$\max(M(T) - K, 0), \quad (5.1)$$

where M is defined as

$$M(t) = \sup_{1 \leq k \leq m(t)} S(t_k^*), \quad (5.2)$$

where $m(t) = \sup\{1 \leq k \leq N : t_k^* \leq t, t_k^* \in D\}$, and we use the convention $M(t) = 0$ for $t = 0$. As noted by Andreasen, the evaluation of the price is a two-step procedure. First, we solve the option price at time t when $M(t) \geq K$. We then solve for the case when $M(t) < K$ by observing that in this case the option can be considered as a first passage problem of S to the level K where the reward is equal to the option value in the previous case where $M(t) \geq K$.

Consider first the case $M(t) \geq K$. The option price is then given by

$$V(t) = E_t \left[e^{-r(T-t)} (M(T) - K)^+ \right] \quad (5.3)$$

$$= E_t \left[e^{-r(T-t)} (M(T) - K) \right]. \quad (5.4)$$

Andreasen then applies the following change of numeraire,

$$d\mathcal{Q}' = \frac{S(T)}{S(t)e^{(r-q)(T-t)}} d\mathcal{Q}, \quad (5.5)$$

where \mathcal{Q} is the risk neutral measure. By Girsanov theorem it follows that, under \mathcal{Q}'

$$W'(t) = W(t) - \sigma t, \quad (5.6)$$

and

$$\frac{dS(t)}{S(t)} = (r - q + \sigma^2)dt + \sigma dW'(t). \quad (5.7)$$

The option price can then be written as

$$V(t) = S(t)E_t' \left[e^{-q(T-t)} \frac{M(T)}{S(T)} \right] - e^{-r(T-t)} K. \quad (5.8)$$

Define

$$x(t) = \frac{M(t)}{S(t)}, \quad (5.9)$$

then for each sampling date $t_k^* \in D$ we have that

$$x^+(t_k^*) = \begin{cases} 1 & \text{if } x^-(t_k^*) \leq 1, \\ x^-(t_k^*) & \text{if } x^-(t_k^*) > 1. \end{cases} \quad (5.10)$$

Applying Ito's formula we get

$$dx(t) = -(r - q)x(t^-)dt - \sigma x(t^-)dW'(t) + (1 - x(t^-))^+ dm(t), \quad (5.11)$$

$$x(t_0) = 1. \quad (5.12)$$

Now define

$$f(t) = E'_t \left[e^{-q(T-t)} \frac{M(T)}{S(T)} \right] \quad (5.13)$$

$$= E'_t \left[e^{-q(T-t)} x(T) \right] \quad (5.14)$$

$$= E' \left[e^{-q(T-t)} x(T) \mid x(t) \right]. \quad (5.15)$$

Then $f(t) = f(t, x(t))$ is given as the solution to the following partial differential equation

$$f_t - (r - q)xf_x + \frac{1}{2}\sigma^2 x^2 f_{xx} - qf = 0, \quad (5.16)$$

with boundary conditions

$$f^-(t_k^*) = BC_f(f^+(t_k^*)) := \begin{cases} f^+(t_k^*, 1), & x \leq 1, t_k^* \in D, \\ f^+(t_k^*, x), & x > 1, t_k^* \in D, \end{cases} \quad (5.17)$$

$$f^+(t_N, x) = f(T, x) = x. \quad (5.18)$$

Thus, for $M(t) \geq K$, $t > t_1$, we have

$$V(t) = S(t)f(t, x(t)) - e^{-r(T-t)}K. \quad (5.19)$$

Now consider the case when $M(t) < K$. The first time $t_k^* > t$, with S larger than K , we get a reward of

$$V(t_k^*) = S(t_k^*)f(t_k^*, x(t_k^*)) - e^{-r(T-t_k^*)}K = S(t_k^*)f(t_k^*, 1) - e^{-r(T-t_k^*)}K, \quad (5.20)$$

where the second equality follows since t_k^* is the first time $M(t) \leq K$. So for $M(t) < K$ the value of the option is

$$V(t) = E \left[e^{-r(\tau-t)} (S(\tau)f(\tau, 1) - e^{-r(T-\tau)}K) 1_{\tau \leq t_n} \mid S(t) \right], \quad (5.21)$$

where

$$\tau = \inf\{t_k^* \in D : S(t_k^*) \geq K\}. \quad (5.22)$$

The solution to the first passage problem (5.20) can be found by solving the corresponding partial differential equation for $g = V(t, S(t))$

$$g_t + (r - \nu)sg_s + \frac{1}{2}\sigma^2 s^2 g_{ss} - rg = 0, \quad (5.23)$$

with boundary conditions

$$g^-(t_k^*) = BC_g(g^+(t_k^*)) := \begin{cases} g^+(t_k^*, s), & s < K, t_k^* \in D, \\ sf(t_k^*, 1) - e^{-r(T-t_k^*)}K, & s \geq K, t_k^* \in D, \end{cases} \quad (5.24)$$

$$g^+(t_N, s) = 0. \quad (5.25)$$

5.2 The Finite Element Method

Let F denote the approximate finite element solution corresponding to equation (5.16). As for the European option we use the artificial boundary condition $F_{ss} = 0$ on $\partial\Omega$. The only difference in equation (5.16) from the Black-Scholes formula is that the coefficient of F is ν instead of r and the sign of the second coefficient, so there is only a minor change in the FEM problem. What differs are the boundary conditions (5.17) and (5.18). Going through exactly the same calculations as for the European option with the changes mentioned above we get the FEM problem, find $U \in \mathcal{W}^q$ such that for $1 \leq n \leq N$

$$\begin{cases} \int_{I_n} (m_f(F_t, v) + a_f(F, v)) dt = 0 & \text{for all } v \in \mathcal{W}_n^{q-1} \\ F^-(t_n) = F^+(t_n), & n = N-1, \dots, 1 \quad t_n \notin D, \\ F^-(t_n) = BC_f(F^+(t_n)), & t_n \in D, \\ F^-(t_N) = f_T, \end{cases} \quad (5.26)$$

where

$$m_f(F_t, v) = (F_t, v) + \frac{\sigma^2}{2(r-\nu)}(sF_t, v)_{\partial\Omega}, \quad (5.27)$$

and

$$\begin{aligned} a_f(F, v) = & -(r-\nu+\sigma^2)(xF_x, v) - \frac{\sigma^2}{2}(x^2F_x, v_x) \\ & - \frac{\sigma^2\nu}{2(r-\nu)}(xF, v)_{\partial\Omega} - \nu(F, v). \end{aligned} \quad (5.28)$$

Similarly we let G denote the approximate finite element solution corresponding to equation (5.23). The finite element problem now reads, find $G \in \mathcal{W}^q$ such that for $1 \leq n \leq N$

$$\begin{cases} \int_{I_n} (m_g(G_t, v) + a_g(G, v)) dt = 0 & \text{for all } v \in \mathcal{W}_n^{q-1} \\ G^-(t_n) = G^+(t_n), & n = N-1, \dots, 1 \quad t_n \notin D, \\ G^-(t_n) = BC_g(G^+(t_n)), & t_n \in D, \\ G^-(t_N) = g_T, \end{cases} \quad (5.29)$$

where

$$m_g(G_t, v) = (G_t, v) - \frac{\sigma^2}{2(r - \nu)}(sG_t, v)_{\partial\Omega}, \quad (5.30)$$

and

$$\begin{aligned} a_g(G, v) &= (r - \nu - \sigma^2)(sG_s, v) - \frac{\sigma^2}{2}(s^2G_s, v_s) \\ &\quad + \frac{\sigma^2 r}{2(r - \nu)}(sG, v)_{\partial\Omega} - r(G, v). \end{aligned} \quad (5.31)$$

6 A Posteriori Error estimation for the Fixed Strike Lookback Option

6.1 Error Representation Formula

Since the primal problem involves a two-step procedure, so does the dual. We will now examine the dual problems for the two pricing PDE:s for the fixed strike lookback option. Since one in most cases seeks the value at time $t = 0$, we are mainly interested in calculating the approximate solution G above, since $M(0) = 0 < K$. Doing so, we have to solve also the equation for F , but we only need the value of F at one point in space. Keeping this in mind, we construct the dual problems according to this criteria. If one is interested in the solution F in more points in space, then one naturally studies a different dual problem.

We begin by introducing the dual problem to the first passage problem (5.23) for g . Find $\phi \in \mathcal{W}$ such that

$$\begin{cases} -\phi_t + (\sigma^2 + \nu - 2r)\phi - (r - \nu - 2\sigma^2)s\phi_s + \frac{\sigma^2}{2}s^2\phi_{ss} = 0, \\ \phi(0, s) = \delta_{s\alpha}, \\ \phi^+(t_k^*) = \begin{cases} \phi^-(t_k^*), & s < K, t_k^* \in D, \\ 0, & s \geq K, t_k^* \in D, \end{cases} \end{cases} \quad (6.1)$$

which we for simplicity consider over the whole space interval, neglecting boundary terms (cf the European problem). Secondly, we introduce the dual problem to the equation for f , (5.16), which is coupled to the previous dual problem. Again neglecting boundary terms, we wish to solve the problem: find $\varphi \in \mathcal{W}$ such that

$$\begin{cases} -\varphi_t + (\sigma^2 + r - 2\nu)\varphi + (r - \nu + 2\sigma^2)x\varphi_x + \frac{\sigma^2}{2}x^2\varphi_{xx} = 0, \\ \varphi(t_0, x) = \varphi(0, x) = \delta_1(x)I_\phi(0), \\ \varphi^+(t_k^*) = \begin{cases} \delta_1(x) \left(\int_{\eta \leq 1} \varphi^-(t_k^*) d\eta + I_\phi(t_k^*) \right), & x \leq 1, t_k^* \in D. \\ \varphi^-(t_k^*), & x > 1, t_k^* \in D, \end{cases} \end{cases} \quad (6.2)$$

where $I_\phi(t_k^*) := \int_{s \geq K} \phi^-(t_k^*, s) s ds$. Using these two coupled dual problems we derive an error representation formula for the solution G , that we are interested in. We now state this formula and then proceed to derive it. Let $e = g - G$ and $\tilde{e} = f - F$, then

$$e(0, s_\alpha) = - \sum_k \int_{t_{k-1}^*}^{t_k^*} \left(m_g(\phi_t, e) + a_g(\phi, e) + m_f(\tilde{e}_t, \varphi) + a_f(\tilde{e}, \varphi) \right) \quad (6.3)$$

where m_g , a_g , m_f , and a_f are the bilinear forms derived in the previous section. To prove this formula we begin by studying equation (6.1). Multiplying with the error $e = g - G \in \mathcal{W}$ and integrating over space and time we get

$$\begin{aligned} \sum_k \int_{t_{k-1}^*}^{t_k^*} \left(-(\phi_t, e) + (\sigma^2 + \nu - 2r)(\phi, e) \right. \\ \left. - (r - \nu - 2\sigma^2)(s\phi_s, e) + \frac{\sigma^2}{2}(s^2\phi_{ss}, e) \right) dt = 0. \end{aligned} \quad (6.4)$$

Examining the first term (ϕ_t, e) in equation (6.4) in detail, we see that

$$\begin{aligned} - \sum_k \int_{t_{k-1}^*}^{t_k^*} (\phi_t, e) dt \\ = \sum_k \left(\int_{t_{k-1}^*}^{t_k^*} (\phi, e_t) dt - (\phi^-(t_k^*), e^-(t_k^*)) + (\phi^+(t_{k-1}^*), e^+(t_{k-1}^*)) \right) \\ = \sum_k \left(\int_{t_{k-1}^*}^{t_k^*} (\phi, e_t) dt \right) - (\phi(T), e(T)) + (\phi(t_0), e(t_0)) \\ - \sum_k \left((\phi^-(t_k^*), e^-(t_k^*)) - (\phi^+(t_k^*), e^+(t_k^*)) \right). \end{aligned} \quad (6.5)$$

Expanding the last two terms on the right we get

$$\begin{aligned} (\phi^-(t_k^*), e^-(t_k^*)) - (\phi^+(t_k^*), e^+(t_k^*)) \\ = \int_{s \geq K} (\phi^-(t_k^*) e^-(t_k^*) - \phi^+(t_k^*) e^+(t_k^*)) ds, \end{aligned} \quad (6.6)$$

since $\phi^+(t_k^*) = \phi^-(t_k^*)$, and $e^-(t_k^*) = e^+(t_k^*)$ for $s < K$, according to the boundary conditions for ϕ and G . Using the boundary condition for ϕ we see that equation (6.6) equals

$$\begin{aligned} \int_{s \geq K} (\phi^-(t_k^*, s) e^-(t_k^*, s) - \phi^+(t_k^*, s) e^+(t_k^*, s)) ds \\ = \int_{s \geq K} \phi^-(t_k^*, s) e^-(t_k^*, s) ds = \tilde{e}^-(t_k^*, 1) \int_{s \geq K} \phi^-(t_k^*, s) s ds, \end{aligned} \quad (6.7)$$

where the second equality follows since for $s \geq K$ it holds that

$$\begin{aligned} e^-(t_k^*, s) &= g^-(t_k^*, s) - G^-(t_k^*, s) \\ &= \left(s f^-(t_k^*, 1) - e^{-r(T-t_k^*)} K \right) - \left(s F^-(t_k^*, 1) - e^{-r(T-t_k^*)} K \right) \\ &= s(f^-(t_k^*, 1) - F^-(t_k^*, 1)) := s\tilde{e}^-(t_k^*, 1), \end{aligned} \quad (6.8)$$

according to the boundary conditions for g and G . Moving derivatives from ϕ to e in equation (6.4) using integration by parts and equation (6.5), we arrive at the error representation formula

$$\begin{aligned} e(0, s_\alpha) &= \sum_k \left(\tilde{e}^-(t_k^*, 1) \int_{s \geq K} \phi^-(t_k^*, s) s ds \right) \\ &\quad - \sum_k \int_{t_{k-1}^*}^{t_k^*} \left((\phi, e_t) - r(\phi, e) + (r - \nu - \sigma^2)(s\phi, e_s) - \frac{\sigma^2}{2}(s^2\phi_s, e_s) \right) dt \\ &= \sum_k \left(\tilde{e}^-(t_k^*, 1) \int_{s \geq K} \phi^-(t_k^*, s) s ds \right) - \sum_k \int_{t_{k-1}^*}^{t_k^*} \left(m_g(\phi_t, e) + a_g(\phi, e) \right), \end{aligned} \quad (6.9)$$

where m_g and a_g are defined above. We thus need to control over the term

$$\sum_k \left(\tilde{e}^-(t_k^*, 1) \int_{s \geq K} \phi^-(t_k^*, s) s ds \right) = \sum_k \left(\tilde{e}^-, \delta_1 \int_{s \geq K} \phi^-(t_k^*, s) s ds \right), \quad (6.10)$$

involving the error of f at the point $x = 1$. We therefore constructed the dual problem for f , (6.2), with the term $\delta_1 \int_{s \geq K} \phi^-(t_k^*, s) s ds = \delta_1 I_\phi(t_k^*)$ as input at each monitoring date $t_k^* \in D$. Remembering that we neglect boundary conditions, and multiplying equation (6.2) with the error $\tilde{e} = f - F \in \mathcal{W}$ and integrating in space and time we get

$$\begin{aligned} \sum_k \int_{t_{k-1}^*}^{t_k^*} \left(-(\varphi_t, \tilde{e}) + (\sigma^2 + r - 2\nu)(\varphi, \tilde{e}) \right. \\ \left. + (r - \nu + 2\sigma^2)(x\varphi_x, \tilde{e}) + \frac{\sigma^2}{2}(x^2\varphi_{xx}, \tilde{e}) \right) dt = 0. \end{aligned} \quad (6.11)$$

Just as in the previous cases we have to be extra careful with the first term (φ_t, \tilde{e}) . Studying this term in detail we see that

$$\begin{aligned} - \sum_k \int_{t_{k-1}^*}^{t_k^*} (\varphi_t, \tilde{e}) dt \\ = \sum_k \left(\int_{t_{k-1}^*}^{t_k^*} (\varphi, \tilde{e}_t) dt \right) - (\varphi(T), \tilde{e}(T)) + (\varphi(t_0), \tilde{e}(t_0)) \\ - \sum_k \left((\varphi^-(t_k^*), \tilde{e}^-(t_k^*)) - (\varphi^+(t_k^*), \tilde{e}^+(t_k^*)) \right). \end{aligned} \quad (6.12)$$

Examining the last two terms on the right, using the boundary conditions for φ and \tilde{e} we see that

$$\begin{aligned}
& (\varphi^-(t_k^*), \tilde{e}^-(t_k^*)) - (\varphi^+(t_k^*), \tilde{e}^+(t_k^*)) \tag{6.13} \\
&= \int_{x \leq 1} (\varphi^-(t_k^*) \tilde{e}^-(t_k^*) - \varphi^+(t_k^*) \tilde{e}^+(t_k^*)) dx \\
&\quad + \int_{x > 1} (\varphi^-(t_k^*) \tilde{e}^-(t_k^*) - \varphi^+(t_k^*) \tilde{e}^+(t_k^*)) dx \\
&= \tilde{e}^-(t_k^*, 1) \int_{x \leq 1} \varphi^-(t_k^*) dx - \int_{x \leq 1} \delta_1(x) \left(\int_{\eta \leq 1} \varphi^-(t_k^*) d\eta + I_\phi(t_k^*) \right) \tilde{e}^+(t_k^*) dx \\
&\quad + \int_{x > 1} (\varphi^-(t_k^*) \tilde{e}^-(t_k^*) - \varphi^-(t_k^*) \tilde{e}^+(t_k^*)) dx \\
&= \tilde{e}^-(t_k^*, 1) \int_{x \leq 1} \varphi^-(t_k^*) dx - \tilde{e}^+(t_k^*, 1) \int_{x \leq 1} \varphi^-(t_k^*) dx - \tilde{e}^+(t_k^*, 1) I_\phi(t_k^*) \\
&= -\tilde{e}^-(t_k^*, 1) I_\phi(t_k^*)
\end{aligned}$$

Now, using integration by parts, moving derivatives from φ to e in equation (6.11), and using equations (6.12) and (6.13), we get

$$\begin{aligned}
& \sum_k (\tilde{e}^-(t_k^*, 1) I_\phi(t_k^*)) \tag{6.14} \\
&= - \sum_k \int_{t_{k-1}^*}^{t_k^*} \left((\varphi, e_t) - \nu(\varphi, e) - (r - \nu + \sigma^2)(x\varphi, e_x) - \frac{\sigma^2}{2}(x^2\varphi_x, e_x) \right) dt,
\end{aligned}$$

or recalling the notations (5.27) and (5.28) and neglecting the boundary terms

$$\sum_k (\tilde{e}^-(t_k^*, 1) I_\phi(t_k^*)) = - \sum_k \int_{t_{k-1}^*}^{t_k^*} (m_f(e_t, \varphi) + a_f(e, \varphi)). \tag{6.15}$$

Summing up, considering equations (6.9) and (6.15) we have thus proved the error representation formula for G .

6.2 Examples

Using the same error estimation algorithm as for the European option we are able to calculate the error in desired quantities for different values of the parameters. The only difference is that the error representation formula (6.3) for the solution G consists of two parts, one from the problem for G and from the problem for F . We calculate the contribution from the two parts separately. This makes it possible to identify regions where a fine mesh is necessary.

In all cases below we have used the boundary conditions $\varphi(0, x) = \delta_1(x, \epsilon)$, where $\epsilon = 5000$, and $\phi(0, s) = \delta_1(s, \epsilon)$, where $\epsilon = 1$. Figures 7 and 8 show dual solutions for the weekly and monthly sampled floating strike lookback put options respectively. We see that the sampling frequency has a significant effect on the dual solution.

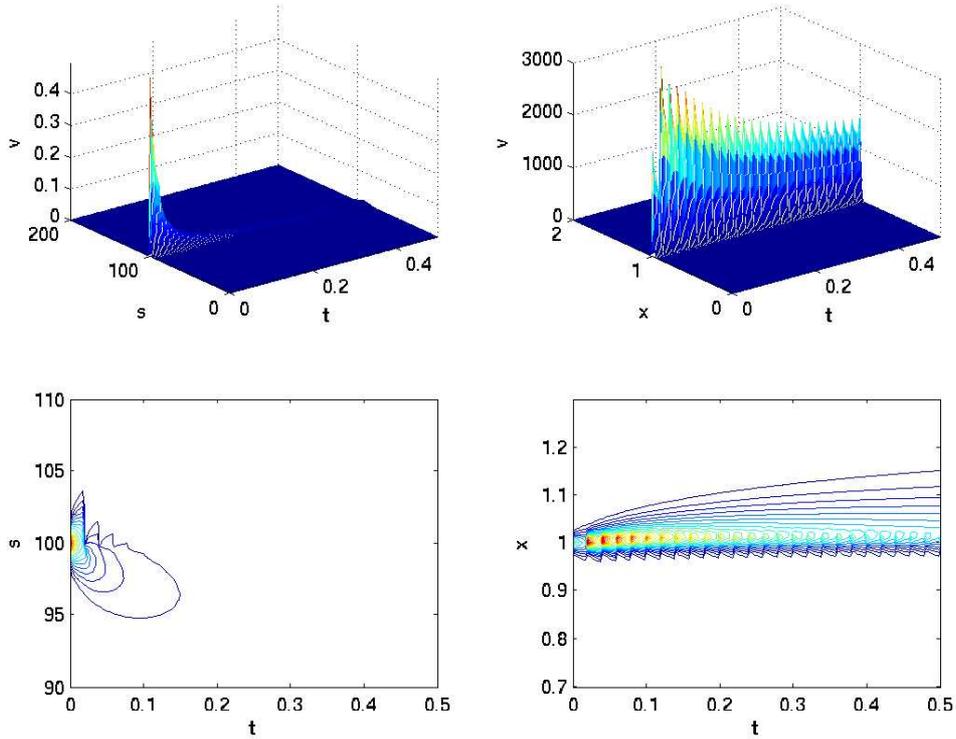


Figure 7: To the left, ϕ , computed using the $cG(2)$ - $dG(1)$ method, with 400 space points and 200 time points. To the right φ , computed using 200 space and time points. Below, contour plots using 30 levels. In both cases $\sigma = 0.2$, $r = 0.05$, $q = 0.0$, and weekly sampling was used.

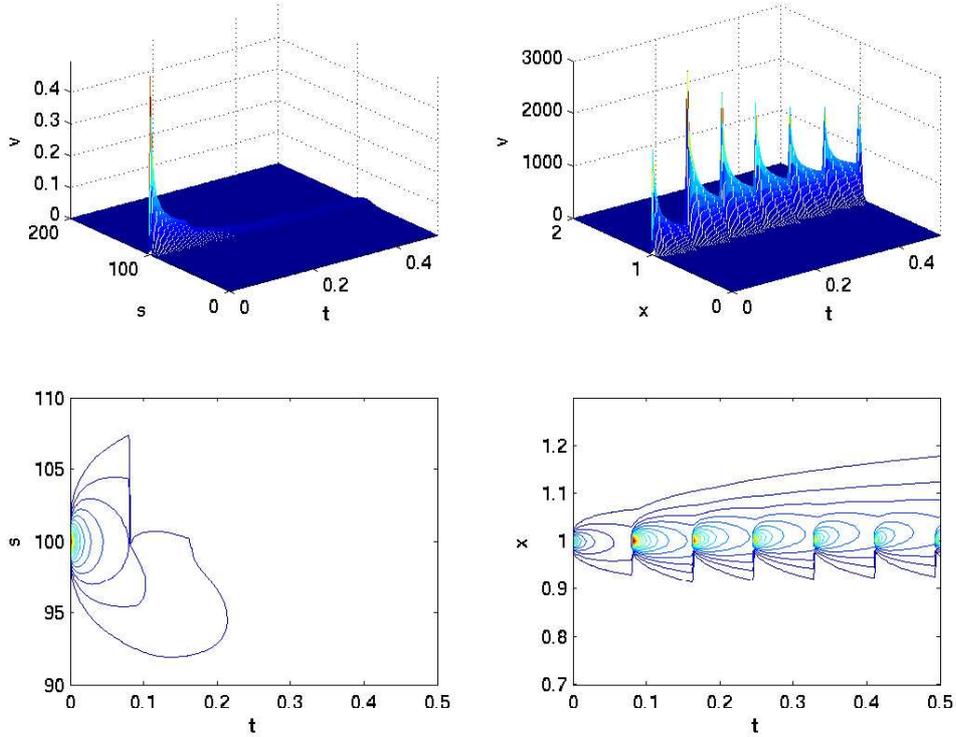


Figure 8: *To the left, ϕ , computed using the cG(2)-dG(1) method, with 400 space points and 200 time points. To the right φ , computed using 200 space and time points. Below, contour plots using 30 levels. In both cases $\sigma = 0.3$, $r = 0.05$, $q = 0.0$, and monthly sampling was used.*

In Figure 9, we see the contributions to the error representation formula (6.3) from each space-time slab. The dual was calculated using the cG(2)-dG(1) method, and the primal using the cG(1)-cG(1) method. The dual mesh was thirty-two times finer in each direction. The value of the functional of the error found by using the error representation formula was in this case 0.0042. We also note that the contribution to the error differs from zero only within a short interval of Ω , just as the dual solution. We now proceed to calculate adaptive meshes.

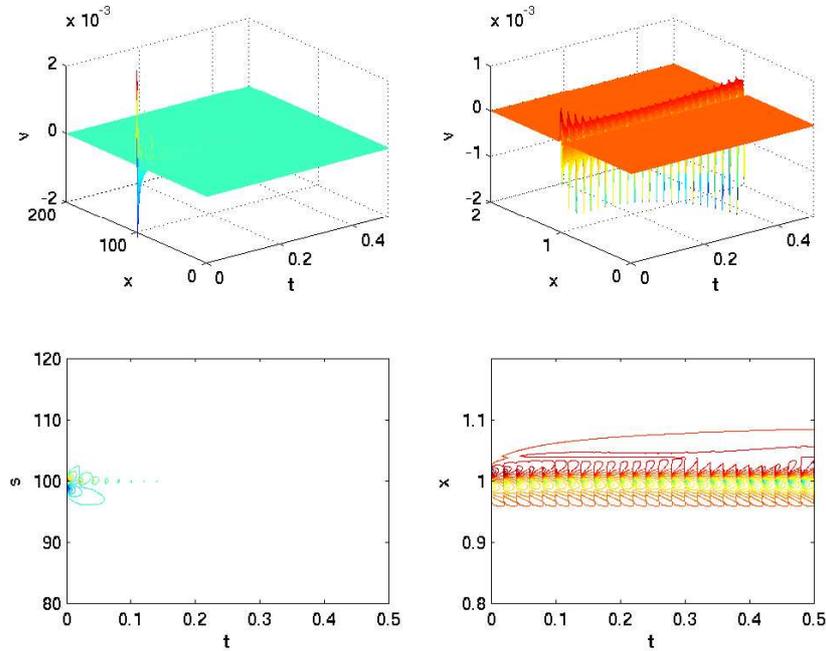


Figure 9: *The contributions to the functional of the error for the weekly sampled fixed strike lookback put when $\sigma = 0.1$, $r = 0.10$ and $q = 0.0$. The dual was computed using the $cG(2)$ - $dG(1)$ method with 800 space and time points, and the primal using the $cG(1)$ - $dG(1)$ method with 25 space and time points.*

7 Adaptive Mesh Refinement for the Fixed Strike Lookback Option

Extending the mesh refinement algorithm used in the case of the European option to this coupled problem is quite straight forward. The error representation formula (6.3) for the solution G consists of two parts, one from the problem for G and from the problem for F . We apply the mesh refinement algorithm to the two parts separately. That is, we refine the two meshes separately.

In Figure 10, we see meshes resulting from using the mesh refinement algorithm in the case of a fixed strike lookback put option with weekly sampling. In this case Q was set to 10%. Two successive refinements were

made on each mesh, starting from a sparse mesh with 25 nodes in time and space. The final meshes has only 30 nodes in each direction, but the functional of the error has decreased by a factor 23. The dual was calculated using a fine mesh with 800 nodes in time and space.

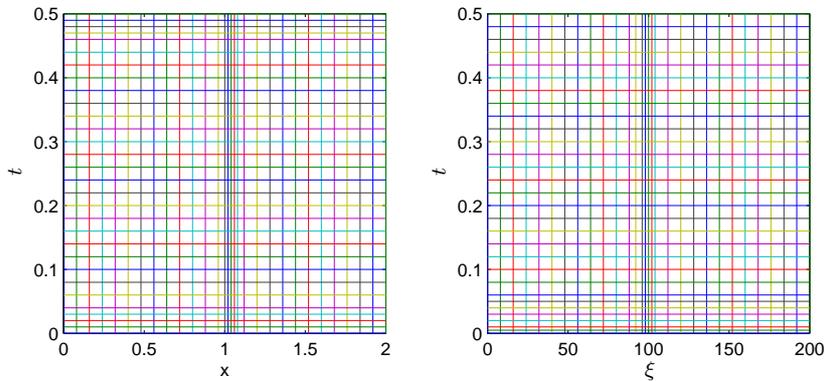


Figure 10: *The resulting meshes using the mesh refinement algorithm, for a fixed strike lookback option with $T = 0.5$, $\sigma = 0.1$, and $r = 0.1$, when $\psi = \delta_1(\xi, \epsilon)$. To the left the mesh for f and to the right the mesh for g . The dual was computed using the $cG(2)$ - $cG(1)$ and the primal using the $cG(1)$ - $dG(1)$ method. Two successive refinements were made.*

8 Results

8.1 The European Option

We begin by validating our method against the known exact solution for the European call option. Recalling the previous calculations in Example 1 in Section 3.3 we know that the error representation formula works and is accurate. The value of the functional of the error found by using the

error representation formula in the test example on the European option was 0.2033, in excellent agreement with the real value of the functional found by using Black-Scholes formula, which was 0.2030. Table 1 compares values of the European call calculated using the cG(1)-cG(1) finite element method mentioned above, with the analytical value derived by Black-Scholes formula. We see that the FE method is very stable and has a maximum relative error of 0.1 percent when 400 time points are used. Figure 11 shows

σ	S(0)	FE(200)	FE(400)	Black-Scholes	Relative error (%)
0.10	90	0.8067	0.8093	0.8101	0.107
	100	5.8478	5.8496	5.8503	0.011
	110	14.9287	14.9297	14.9300	0.002
0.20	90	3.0487	3.0500	3.0504	0.014
	100	8.2767	8.2775	8.2778	0.003
	110	16.0177	16.0184	16.0187	0.002
0.30	90	5.5198	5.5206	5.5209	0.005
	100	10.9058	10.9063	10.9065	0.002
	110	18.0464	18.0468	18.0469	0.0008

Table 1: *The European call calculated using the cG(1)-cG(1) method compared to Black-Scholes analytical value when $r = 0.1$, $q = 0.0$, $T = 0.5$, $K = 100$, and $t = 0$. The number of time and space points is given in parenthesis. The relative error is between the FE(400) solution and the analytical solution.*

the finite element solution calculated using a the adapted mesh in the previous section. The mesh is finer close to time $t = 0$ and close to the strike price, but it is not centered around the strike price. In this way the same accuracy is achieved in less degrees of freedom. The original uniform mesh has 20 nodes in time and space. By using the error representation formula the error was calculated to 0.2 for the uniform mesh. The adapted mesh has only 27 nodes in the spatial direction, but the error has decreased by a factor 70 to 0.0028.

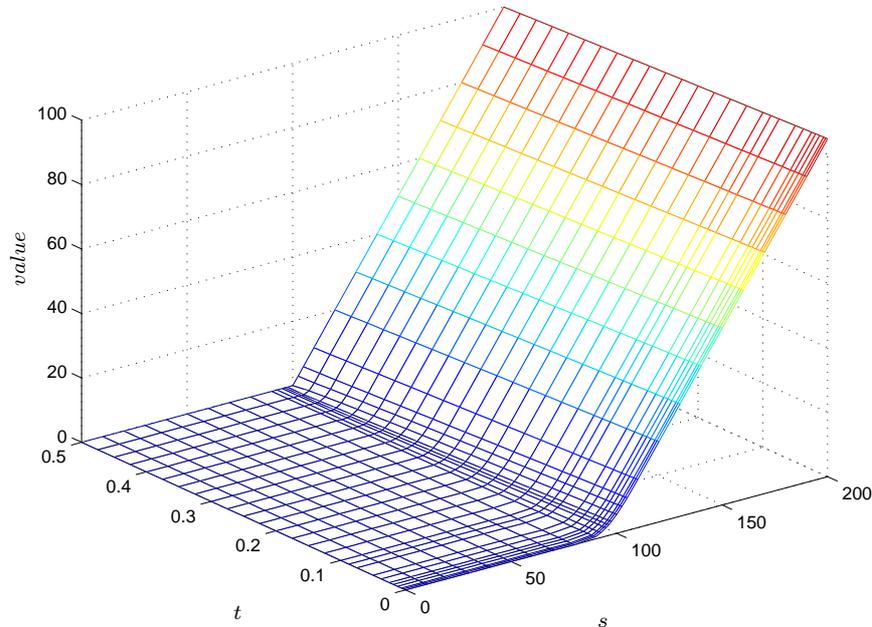


Figure 11: *The finite element solution U , when $\sigma = 0.1$, $q = 0.0$, $K = 100$, and $r = 0.10$. Computed using the $cG(1)$ - $cG(1)$ method on an adapted mesh with 27 time and space points.*

8.2 The Fixed Strike Lookback Option

Table 2 compares values of the discrete fixed strike lookback put option calculated with the finite difference solution in [1], with the finite element solution computed in this paper. As a comparison values of the Monte Carlo solution are given, also these from [1]. The finite element solution developed here shows sufficient precision already for the sparse mesh, whereas the finite difference solution needs a much finer mesh. The reason for this is probably that we choose an individual mesh for each of the two coupled problems.

By using the mesh refinement algorithm an adapted mesh was calculated for the example of the weekly sampled fixed strike lookback put option. The meshes for the two coupled problems were both finer close to the center of Ω . The mesh for f was finer towards time $t = 0$ and $t = T$, whereas the mesh for g was finer close to $t = 0$. In this way the same accuracy is achieved in less degrees of freedom. The original uniform mesh has 25 nodes in time and

space. By using the error representation formula the functional of the error was calculated to 0.89 for the uniform mesh. The adapted mesh has only 30 nodes in each direction, but the functional of the error has decreased by a factor 23 to 0.039.

K	MC	FD(500)	FD(100)	FE(500)	FE(100)
90.0	24.41	24.40	24.27	24.40	24.38
92.5	22.07	22.06	21.93	22.06	22.04
95.0	19.78	19.77	19.64	19.77	19.75
97.5	17.57	17.56	17.43	17.56	17.55
100.0	15.48	15.47	15.34	15.47	15.45
102.5	13.53	13.52	13.39	13.52	13.51
105.0	11.75	11.74	11.62	11.75	11.73
107.5	10.14	10.14	10.03	10.14	10.13
110.0	8.70	8.71	8.62	8.70	8.70

Table 2: *The fixed strike lookback put option when when $r = 0.05$, $q = 0.0$, $T = 1.0$, $t = 0.0$, and $S(0)=100$. MC refers to Monte Carlo solution, FD refers to the finite difference solution in [1], and FE refers to the finite element solution computed in this paper. The number of time and space points is given in parenthesis. Sampling is made at times $0.1, 0.2, \dots, 1.0$.*

Figure 12 shows the value of the monthly sampled lookback put option.

9 Conclusions

The presented a posteriori error estimation formula is verified in the case of the European option were we have access to an analytical solution. The error estimation works well for both European and fixed strike lookback options. The uses of adapted meshes gives superior accuracy and performance than using uniform meshes. By using individual meshes for the two coupled equations in the case of the fixed strike lookback option we achieve better accuracy already for sparse meshes. The technique of using a system of two coupled dual problems was first presented in [8]. Similarly results has been exploited by Målqvist in [14].

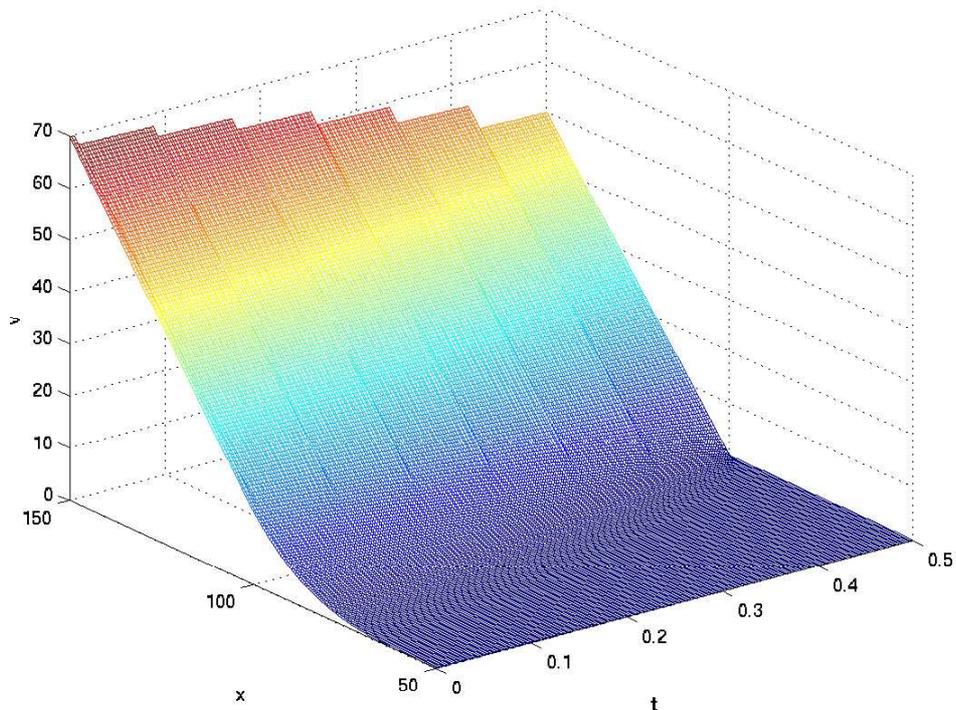


Figure 12: *The monthly sampled lookback put option, computed using the $cG(1)$ - $cG(1)$ method with 200 space and time points. Parameter values are $\sigma = 0.3$, $r = 0.05$, $q = 0.0$, $T = 0.5$, and $t = 0.0$.*

References

- [1] J. ANDREASEN, *The pricing of discretely sampled Asian and lookback options: a change of numeraire approach*, *The Journal of Computational Finance*, 2 (Fall 1998), pp. 5–30.
- [2] S. BABBS, *Binomial valuation of lookback options*. Working paper, Midland Montague, Capital Markets, London, 1992.
- [3] J. BARRAQUAND AND T. PUDET, *Pricing of of American path-dependent contingent claims*, *Math. Fin.*, (1996), pp. 17–51.
- [4] T. CHEUK AND T. VORST, *Lookback options and the observation frequency: a binomial approach*. Working paper, Erasmus University, 1994.

- [5] A. CONZE AND R. VISWANATHAN, *Path-dependent options: The case of lookback options.*, Journal of Finance, (1991), pp. 1893–1907.
- [6] K. ERIKSSON, D. ESTEP, P. HANSBO, AND C. JOHNSON, *Computational Differential Equations*, Studentlitteratur, 1996.
- [7] D. ESTEP, M. LARSON, AND R. WILLIAMS, *Estimating the error of numerical solutions of systems of reaction-diffusion equations*, MEMOIRS of the American Mathematical Society, 146 (2000).
- [8] G. FOUFAS, *Valuing pathdependent options using the finite element method and duality techniques*. Licentiate thesis, 2003.
- [9] G. FOUFAS AND M. G. LARSON, *Valuing european, barrier, and lookback options using the finite element method and duality techniques*. 2007.
- [10] ———, *Valuing asian options using the finite element method and duality techniques*, to appear in JCAM, (2008).
- [11] M. GOLDMAN, H. SOSIN, AND M. GATTO, *Path-dependent options: buy at the low, sell at the high.*, Journal of Finance, (1979), pp. 1111–1128.
- [12] M. GOLDMAN, H. SOSIN, AND L. SHEPP, *On contingent claims that insure ex-post optimal stock market timing*, Journal of Finance, (1979), pp. 401–414.
- [13] J. HULL AND A. WHITE, *Efficient procedures for valuing European and American path-dependent options*, J. Derivatives, (1993), pp. 21–31.
- [14] M. LARSON AND A. MÅLQVIST, *Goal oriented adaptivity for coupled flow and transport problems with applications in oil reservoir simulation*, Comput. Methods Appl. Mech. Eng., (2007), pp. 3546–3561.
- [15] P. WILMOTT, J. DEWYNNE, AND S. HOWISON, *Option pricing*, Oxford Financial Press, 1993.
- [16] R. ZVAN, P. FORSYTH, AND K. VETZAL, *A finite element approach to the pricing of discrete lookbacks with stochastic volatility*, Applied Mathematical Finance, (1999), pp. 87–106.

Paper III

Valuing Asian Options using the Finite Element Method and Duality Techniques

Georgios Foufas* Mats G. Larson†

April 14, 2008

Abstract

The main objective of this paper is to develop an adaptive finite element method for computation of the values, and different sensitivity measures, of the Asian option with both fixed and floating strike. The pricing is based on Black-Scholes PDE-model and a method developed by Večer were the resulting PDE:s are of parabolic type in one spatial dimension and can be applied to both continuous and discrete Asian options. We propose using an adaptive finite element method which is based on a posteriori estimates of the error in desired quantities, which we derive using duality techniques. The a posteriori error estimates are tested and verified, and are used to calculate optimal meshes for each type of option. The use of adapted meshes gives superior accuracy and performance with less degrees of freedom than using uniform meshes. The suggested adaptive finite element method is stable, gives fast and accurate results, and can be applied to other types of options as well.

1 Introduction

Background: The Asian option was invented by Phelim P. Boyle and David Emanuel in 1979, but The Journal of Finance rejected their paper since the asset was not traded at that time (private communication). Asian options are securities with payoffs which depend on the average of the underlying stock price over some time interval. They are commonly traded and are often relatively inexpensive compared to European calls. Asian options were introduced partly to avoid a problem common for European options, where the speculators could drive up the gains from the option by manipulating

*Research Assistant, Department of Mathematics, Chalmers University of Technology, S-412 96 Gothenburg, Sweden, *email*: fofas@math.chalmers.se

†Professor of Applied Mathematics, Corresponding author, Department of Mathematics, Umeå University, S-901 87 Umeå, Sweden, *email*: mats.larson@math.umu.se

the price of the underlying asset near to the maturity date (see Bergman [5], or Wall Street Journal, Jan. 21, 1982, p. 4). The name Asian option probably originates from the Tokyo office of Bankers Trust, where it first was offered (see Nelken [13]).

Previous work: No general analytical price formula is known for the average rate option, on the other hand several approximations that produce closed form expressions have appeared, such as Thompson, [16], who provides tight analytical bounds the price of the Asian option. Geman and Yor computed the Laplace transform of the Asian option price, but numerical inversion remains problematic for low volatility and short maturity cases (see Fu Madan and Wang [9]). Linetski [12], has derived a new integral formula for the continuous sampled Asian option, which also is slowly convergent for low volatility cases. Monte Carlo simulation works well, but sometimes it is computationally expensive.

In general, the price of an Asian option can be found by solving a PDE in two space dimension as noted by Ingersoll [11]. However this PDE often gives oscillatory solutions. Ingersoll also notes that a change of variable gives a one-dimensional PDE for the floating strike Asian option. Rogers and Shi [14], presented a one-dimensional PDE that can model both fixed and floating strike Asian options. They also computed lower and upper bounds for the price of the Asian option, where the lower bound is very accurate. Their PDE is also difficult to solve numerically, since the diffusion term is very small. Zvan, Forsyth and Vetzal [20], suggest a method based on computational fluid dynamics techniques to overcome this difficulty. In [3] Andreasen applied the Rogers-Shi reduction to the discrete Asian option with very good results.

Shreve and Večer [15], shows that the arithmetic Asian option (both with fixed and floating strike) is a special case of an option on a traded account. Options on a traded account generalize the concept of many options (passport, European, American and vacation) and the same pricing techniques can be used to price the Asian option. The resulting PDE:s for the price of Asian options are of parabolic type with one space-dimension and they are easy to solve and give fast and accurate results. Foufas applied the Finite Element (FE) method to this PDE in [8]. Later Večer [19], presented an even simpler two-term one-dimensional PDE for the arithmetic Asian option with general dividends. However, the FE formulation of this equation is almost the same as the FE formulation of the three term PDE introduced by Shreve and Večer. The only difference is one of the coefficients in the FE problem formulation, see Section 3. Recently Topper [17], applied the FE method to Večer's PDE. As pointed out by Topper the FE approach has several advantages compared to other numerical techniques such as Finite Differences

(FD) techniques. For example, using the FE method one receives a solution in the entire domain, not only in isolated nodes as in FD codes. FE codes can also incorporate different kinds of boundary conditions in an easy way. Other important advantages of the FE technique are that it can easily deal with high curvature and irregular shapes of the computational domain. One of the most important advantages in practice is that the sensitivity measures, or the so called greeks, can be calculated more exactly using the FE method.

New contributions: The Asian option is priced using the Black-Scholes PDE-model. The resulting PDE:s are of parabolic type in one spatial dimension. The numerical computation is made using an adaptive finite element method allowing variable resolution in space and time. Whereas Topper uses a commercial FE solver (*PDEase2DTM*) with local mesh refinements we use our own code based on the concept of duality techniques. This technology has to our knowledge not been used before on this type of problems.

In practice one is only interested in the price, and it's derivatives, in one or a few points. Using this criteria, the choice of computational mesh is based on a posteriori estimates of the error in desired quantities, which we derive using duality techniques. In contrast to element based indicators dual techniques allow us to estimate the error in user specified goal quantities expressed by linear functionals which is particularly suitable for financial applications where we are mainly interested in the solution and its derivatives in specific points. For an overview of different a posteriori error estimation techniques we refer to Bangert and Rannacher [4], Ainsworth and Oden [1], and the references therein.

The dual a posteriori error estimation techniques are shown to be very useful and simple. The presented a posteriori error estimation formula is tested and verified in the case of the European option. It is then used to perform mesh refinements in both time and space for the Asian option. This makes it possible to calculate an optimal mesh for each type of option, which dramatically reduces the error without noticeably enhancing the computational effort. The suggested adaptive finite element method is stable and gives fast and accurate results. The technique is general and can be applied to other types of options, such as the floating and fixed strike lookback options studied in a forthcoming paper by the authors.

Outline: In Section 2 we define different kinds of Asian options and present a pricing PDE for arithmetic Asian options. In Section 3 we formulate the adaptive finite element method and derive an a posteriori error estimation. We also test and verify the a posteriori error estimation formula and give some examples. In Section 4 we present an adaptive mesh refinement algorithm based on the a posteriori error estimates. Then the sensitivity

measures, or so called greeks, are presented in Section 5. Finally, in Section 6 we give some numerical results.

2 The Asian Option

2.1 Classification

Different kinds of averages are used, resulting in different types of Asian options, with different values. The method of sampling is also important. A continuous sampling may give easier calculations, but in reality the prices are mostly discretely sampled, and therefore discrete sampling is the most interesting case. The geometric Asian option with time of maturity T and strike price K has the payoff

$$\max\left(\prod_{k=1}^N S(t_k)^{1/N} - K, 0\right), \quad (2.1)$$

where $0 < t_1 < t_2 < \dots < t_N = T$. For this option one can use the Black-Scholes framework to determine a closed-form pricing formula. Note that if $N = 1$ the option is reduced to a European call.

The average rate call with strike price K and time of maturity T has the payoff

$$\max\left(\frac{1}{T} \int_0^T S(t) dt - K, 0\right), \quad (2.2)$$

while the discrete average rate call with strike price K and time of maturity T has the payoff

$$\max\left(\frac{1}{N} \sum_{k=1}^N S(t_k) - K, 0\right), \quad (2.3)$$

where $0 < t_1 < t_2 < \dots < t_N = T$. There are no known closed-form pricing formulas for average rate options, but a variety of numerical techniques have been developed to find the corresponding prices.

The average rate call is cheaper than the European call at the writing date, see Table 1 and Theorem 2.1 in Section 2.4.

There are also variants of the Asian options mentioned above. For a floating strike Asian option the strike K in (2.2) and (2.3) is replaced by the spot price $S(T)$ at maturity. The corresponding options are often called average strike put and discrete average strike put respectively.

2.2 Pricing Arithmetic Asian Options

This article mainly focuses on an article written by Večer [19]. Here we present a short derivation of a pricing PDE for the arithmetic Asian option following the exposition in [19]. For simplicity we consider the case with continuous dividends and continuous sampling. The changes in the case of general dividends and discrete sampling are small. The details can be found in [19].

As noted by Večer the general payoff of the Asian option can be written as

$$(\bar{S}_T - K_1 S_T - K_2)^+ \quad \text{or} \quad (K_2 - K_1 S_T - \bar{S}_T)^+, \quad (2.4)$$

where $\bar{S}_T = \frac{1}{T} \int_0^T S_t dt$. Because of the Asian Put-Call parity

$$e^{-rT} E [(\bar{S}_T - K_1 S_T - K_2)^+] - e^{-rT} E [(K_2 + K_1 S_T - \bar{S}_T)^+] \quad (2.5)$$

$$= e^{-rT} E [\bar{S}_T - K_1 S_T - K_2]$$

$$= \frac{1}{(r - \gamma)T} (e^{-\gamma T} - e^{-rT}) S_0 - K_1 e^{-\gamma T} S_0 - e^{-rT} K_2, \quad (2.6)$$

it is enough to compute the value for the Asian option with the payoff $(\bar{S}_T - K_1 S_T - K_2)^+$, that is for the fixed strike Asian call option if we choose $K_1 = 0$, and for the floating strike Asian put option if choose $K_2 = 0$.

Let the underlying asset evolve under the risk neutral measure according to the equation

$$dS_t = S_t((r - \gamma)dt + \sigma dW_t), \quad (2.7)$$

where r is the interest rate, γ is a continuous dividend yield, and σ is the volatility of the underlying asset. Let also

$$q_t = \frac{1}{(r - \gamma)T} (e^{-\gamma(T-t)} - e^{-r(T-t)}) \quad (2.8)$$

denote the trading strategy, the number of shares held at time t , and let the wealth evolve according to the following self-financing strategy

$$\begin{aligned} dX_t &= q_t dS_t + r(X_t - q_t S_t)dt + q_t \gamma S_t dt \\ &= rX_t dt + q_t (dS_t - rS_t dt + \gamma S_t dt), \end{aligned} \quad (2.9)$$

with the initial wealth

$$X_0 = \frac{1}{(r - \gamma)T} (e^{-\gamma T} - e^{-rT}) S_0 - e^{-rT} K_2 = q_0 S_0 - e^{-rT} K_2. \quad (2.10)$$

We then have that

$$\begin{aligned}
X(T) &= e^{rT} X(0) + \int_0^T q_t e^{r(T-t)} (dS_t - rS_t dt + \gamma S_t dt) \quad (2.11) \\
&= e^{rT} X(0) + q_T S_T - e^{rT} q_0 S_0 + \int_0^T e^{r(T-t)} S_t (q_t \gamma dt - q_t' dt) \\
&= \frac{1}{T} \int_0^T S_t dt - K_2 = \bar{S}_T - K_2.
\end{aligned}$$

Remark 2.1 *By choosing $q_t = e^{\gamma(t-T)}$ we obtain the ordinary European call option, which can be seen by examining equation (2.11). This result is used later on to verify the a posteriori error estimation results, since there exists an analytical solution to the European option problem.*

2.3 A Pricing Partial Differential Equation

Following Večeř we use the change of numeraire technique to reduce the dimensionality of the problem by introducing the process

$$Z_t = \frac{X_t}{e^{\gamma t} S_t}. \quad (2.12)$$

Using Ito's lemma we get that

$$\begin{aligned}
dZ_t &= (Z_t - e^{-\gamma T} q_t) dt - (Z_t - e^{-\gamma T} q_t) \sigma dW_t \quad (2.13) \\
&= -(Z_t - e^{-\gamma T} q_t) \sigma d\tilde{W}_t,
\end{aligned}$$

where $\tilde{W}_t = -\sigma t + W_t$ is a Brownian motion under the numeraire measure. The price of the Asian call option, $V(t, S_t, K_1, K_2)$, can at time $t = 0$ be represented as

$$V(0, S_0, K_1, K_2) = e^{-rT} E [(X_T - K_1 S_T)^+] = S_0 \tilde{E} [(Z_T - K_1)^+]. \quad (2.14)$$

Introducing

$$\bar{u}(0, Z_0) = \tilde{E} [(Z_T - K_1)^+], \quad (2.15)$$

where

$$Z_0 = \frac{X_0}{S_0} = \frac{1}{(r - \gamma)T} (e^{-\gamma T} - e^{-rT}) - e^{-rT} \frac{K_2}{S_0}, \quad (2.16)$$

we can write the price of the option as

$$V(0, S_0, K_1, K_2) = S_0 \bar{u}(0, Z_0). \quad (2.17)$$

It can be shown that \bar{u} is the solution to the following partial differential equation

$$\begin{aligned}
\bar{u}_t + \frac{1}{2} (z - e^{-\gamma t})^2 \sigma^2 \bar{u}_{zz} &= 0, \quad (2.18) \\
\bar{u}(T, z) &= (z - K_1)^+.
\end{aligned}$$

2.4 Comparison of European and Asian Options

At the writing date, the average rate call is cheaper than the European call (cf. the Geman and Yor paper [10]). We here present, to our knowledge, a new and simpler proof of this statement than the one presented in [2].

Theorem 2.1 *If $\rho(t) \geq 0$ and $\int_0^T \rho(t)dt = 1$, then for any $T > 0$,*

$$e^{-rT} E \left[\left(\int_0^T S(\lambda) \rho(\lambda) d\lambda - K \right)^+ \mid \mathcal{F}_0 \right] < e^{-rT} E \left[\left(S(T) - K \right)^+ \mid \mathcal{F}_0 \right].$$

Proof. Note that $E[X \mid \mathcal{F}_0] = E[X]$ so we omit the σ -algebra \mathcal{F}_0 in the following. Note also that

$$E \left[(S(T_0) - K)^+ \right] < E \left[(S(T) - K)^+ \right], \text{ if } T_0 < T,$$

since an American call price is the same as the price of the corresponding European call when the underlying stock does not pay dividends. Now

$$\begin{aligned} E \left[\left(\int_0^T S(\lambda) \rho(\lambda) d\lambda - K \right)^+ \right] &= E \left[\left(\int_0^T (S(\lambda) - K) \rho(\lambda) d\lambda \right)^+ \right] \\ &\leq E \left[\int_0^T (S(\lambda) - K)^+ \rho(\lambda) d\lambda \right] = \int_0^T \rho(\lambda) E \left[(S(\lambda) - K)^+ \right] d\lambda \\ &< \int_0^T \rho(\lambda) E \left[(S(T) - K)^+ \right] d\lambda = \int_0^T \rho(\lambda) d\lambda E \left[(S(T) - K)^+ \right] \\ &= E \left[(S(T) - K)^+ \right], \end{aligned}$$

and the Theorem follows at once. \square

A more detailed comparison of European and Asian options and their sensitivity measures can be found in [2].

	Average rate call			European call		
$K \setminus \sigma$	0.10	0.20	0.30	0.10	0.20	0.30
90	13.73	14.14	15.24	14.63	16.70	19.70
100	5.26	7.04	9.06	6.81	10.45	14.23
110	0.73	2.70	4.86	2.17	6.04	10.02

Table 1: *The European call compared to the average rate call for various strikes K and volatilities σ when $r = 0.05$, $T = 1$ and $t = 0$.*

3 An Adaptive Finite Element Method for the Asian Option

Since there probably does not exist a closed form solution to the PDE (2.18), the price of the Asian option must be obtained numerically. The method used in this article is the finite element method as presented below.

3.1 Variational Formulation

So far we have studied the pricing PDE for Asian options valid for $z \in \mathbb{R}$, but in order to construct a computational mesh we introduce a bounded interval $\Omega = [z_0, z_J] \subset \mathbb{R}^+$ with boundary $\partial\Omega = \{z_0, z_J\}$. We define the usual Hilbert space

$$H^1(\Omega) = \{v : \int_{\Omega} (|\nabla v|^2 + v^2) dz < \infty\}, \quad (3.1)$$

and let \mathcal{W} be the space of functions that are square integrable in time and belongs to $H^1(\Omega)$ in space, that is

$$\mathcal{W} = L^2([0, T], H^1(\Omega)). \quad (3.2)$$

We denote by u the solution to (2.18) on Ω subject to the Dirichlet boundary conditions $u(t, z_0) = 0$ and $u(t, z_J) = z_J$ on $\partial\Omega$. We also use the notation $(u, v)_{\Omega} = \int_{\Omega} uvdz$, and $(u, v)_{\partial\Omega} = u(z_J)v(z_J) - u(z_0)v(z_0)$. Multiplying equation (2.18) by the test function $\{v \in \mathcal{W} : v = 0 \text{ on } \partial\Omega\}$ and integrating on $\Omega \times [0, T]$ we obtain

$$\int_0^T \left((u_t, v)_{\Omega} + \frac{\sigma^2}{2} ((z - e^{-\gamma t} q_t)^2 u_{zz}, v)_{\Omega} \right) dt = 0. \quad (3.3)$$

Using integration by parts we get

$$\begin{aligned} ((z - e^{-\gamma t} q_t)^2 u_{zz}, v)_{\Omega} &= ((z - e^{-\gamma t} q_t)^2 u_z, v)_{\partial\Omega} \\ &\quad - 2((z - e^{-\gamma t} q_t) u_z, v)_{\Omega} - ((z - e^{-\gamma t} q_t)^2 u_z, v_z)_{\Omega}. \end{aligned} \quad (3.4)$$

Thus equation (3.3) becomes

$$\begin{aligned} \int_0^T \left((u_t, v)_{\Omega} - \sigma^2 ((z - e^{-\gamma t} q_t) u_z, v)_{\Omega} \right) dt \\ - \int_0^T \frac{\sigma^2}{2} ((z - e^{-\gamma t} q_t)^2 u_z, v_z)_{\Omega} dt = 0, \end{aligned} \quad (3.5)$$

since $v = 0$ on $\partial\Omega$. Introducing the Dirichlet boundary conditions $u(t, z_0) = 0$ and $u(t, z_J) = z_J$ on $\partial\Omega$ (which is also used by Večer [19]) we get the following problem: find $u \in \mathcal{W}$ such that

$$\begin{cases} \int_0^T ((u_t, v)_\Omega + a_{\Omega(u,v)}) dt = 0, \\ u(T, z) = z^+, \\ u(t, z_0) = 0, \quad u(t, z_J) = z_J, \end{cases} \quad (3.6)$$

for every $\{v \in \mathcal{W} : v = 0 \text{ on } \partial\Omega\}$, where

$$a_{\Omega}(u, v) = -\sigma^2((z - e^{-\gamma t} q_t)u_z, v)_\Omega - \frac{\sigma^2}{2}((z - e^{-\gamma t} q_t)^2 u_z, v_z)_\Omega. \quad (3.7)$$

Remark 3.1 *The variational formulation of the three term PDE developed by Shreve and Večer [15], used to value options on a traded account and Asian options, is in the case of no dividends ($\gamma = 0$) given by the same expression as above but with the following $a_{\Omega}(u, v)$*

$$a_{\Omega}(u, v) = (r + \sigma^2)((q - z)u_z, v)_\Omega - \frac{\sigma^2}{2}((q - z)^2 u_z, v_z)_\Omega. \quad (3.8)$$

The only difference between this expression and (3.7) is the coefficient in front of one of the terms. For a derivation we refer to [8].

This means that the finite element problem presented in this paper is almost exactly the same as the problem received by instead studying the three term PDE derived by Shreve and Večer. The two term PDE appears simpler, but from a variational point of view the two equations are essentially the same.

3.2 Finite Element Approximation

The finite element method is based on solution of the variational problem (3.6) with \mathcal{W} replaced by a finite dimensional function space of piecewise polynomials in space and time. For background on the finite element method see for instance [6].

We now partition $[0, T]$ as $0 = t_0 < t_1 < t_2 < \dots < t_N = T$, denoting each time interval by $I_n = (t_{n-1}, t_n]$ and each time step by $k_n = t_n - t_{n-1}$. Similarly we partition Ω as $z_0 < z_1 < z_2 < \dots < z_J$, denoting each spatial interval by $\kappa_j = [z_{j-1}, z_j)$ and the length of each interval by $h_j = z_j - z_{j-1}$.

In space, we let $\mathcal{V}^p \subset H^1(\Omega)$ denote the space of piecewise continuous functions of order p . On each space-time slab $S_n = I_n \times \Omega$, we define

$$\mathcal{W}_n^q = \{w(t, z) : w(t, z) = \sum_{j=0}^q t^j v_j(z), v_j \in \mathcal{V}^p, (t, z) \in S_n\}. \quad (3.9)$$

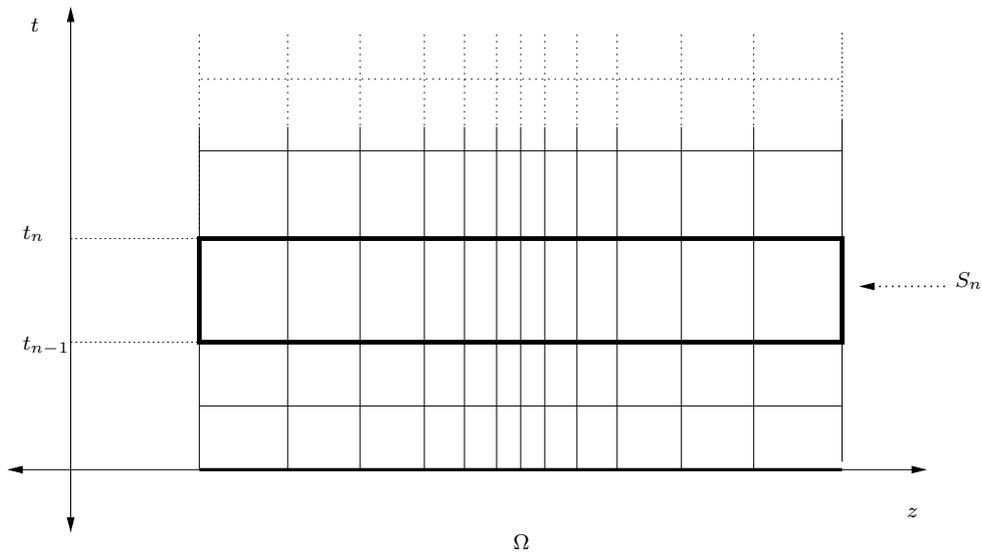


Figure 1: *Space-time discretization.*

Let $\mathcal{W}^q \subset \mathcal{W}$ denote the space of functions defined on $[0, T] \times \Omega$ such that $v|_{S_n} \in \mathcal{W}_n^q$ for $1 \leq n \leq N$. For simplicity, we only give details for the continuous Galerkin method cG(p)-cG(q), (see e.g. [6] or [7]) which is defined by the following discrete version of equation (3.6). Find $U \in \mathcal{W}^q$ such that for $1 \leq n \leq N$

$$\begin{cases} \int_{I_n} ((U_t, v)_\Omega + a_\Omega(U, v)) dt = 0, & \forall \{v \in W_n^0 : v = 0 \text{ on } \partial\Omega\}, \\ U^-(t_n) = U^+(t_n), & n = N - 1, \dots, 1, \\ U^-(t_N) = u_T, \\ U(t_n, z_0) = 0, \quad U(t_n, z_J) = z_J, & n = N - 1, \dots, 1, \end{cases} \quad (3.10)$$

where $U^\pm(t_n) = \lim_{\epsilon \rightarrow 0, \epsilon > 0} U(t_n \pm \epsilon)$. In the cG(1) method the approximation U of u is continuous piecewise linear in time and space, while the test functions v are continuous linear in space and piecewise constant in time. It is also possible to use a discontinuous method in time, we refer to [6], for details on the resulting discontinuous Galerkin method, cG(p)-dG(q).

3.3 A Posteriori Error Estimation

3.3.1 Error Representation Formula

Since we are only interested in the solution, and its derivatives, in one or a few points of Ω at time $t = 0$, we wish to find a mesh tailored for efficient and accurate solution at the points of interest. In order to find such a mesh

we derive a posteriori error estimates of the error in the points of interest using duality techniques (see [6] or [7]).

To represent the error in a linear functional, $(u - U, \psi)$, we introduce the continuous dual problem for equation (2.18). Find $\phi \in \mathcal{W}$ such that

$$\begin{cases} -\phi_t + \sigma^2 \phi + 2\sigma^2(z - e^{-\gamma t} q_t) \phi_z + \frac{\sigma^2}{2}(z - e^{-\gamma t} q_t)^2 \phi_{zz} = 0, \\ \phi(0, z) = \psi. \end{cases} \quad (3.11)$$

For simplicity we consider this equation over the whole space interval, so we don't have to consider any boundary conditions. We extend the finite element solution U outside $\Omega = [z_0, z_J]$ by defining

$$\begin{cases} \bar{U} = 0, & z \leq z_0, \\ \bar{U} = U, & z_0 \leq z \leq z_J, \\ \bar{U} = z, & z \geq z_J. \end{cases} \quad (3.12)$$

Furthermore, we also extend the previous notation and let $(u, v) = (u, v)_{\mathbb{R}}$, and $a(u, v) = a_{\mathbb{R}}(u, v)$ denotes the bilinear forms extended from Ω to \mathbb{R} . Multiplying with the error $e = \bar{u} - \bar{U} \in \mathcal{W}$ and integrating in space and time we get

$$\begin{aligned} & \int_0^T \left(-(\phi_t, e) + \sigma^2(\phi, e) + 2\sigma^2((z - e^{-\gamma t} q_t) \phi_z, e) \right) dt \\ & + \int_0^T \left(\frac{\sigma^2}{2}((z - e^{-\gamma t} q_t)^2 \phi_{zz}, e) \right) dt = 0 \end{aligned} \quad (3.13)$$

Using integration by parts we get

$$\begin{aligned} & -(\phi(T), e(T)) + (\phi(0), e(0)) \\ & + \int_0^T \left((\phi, e_t) + \sigma^2(\phi, e) - 2\sigma^2((z - e^{-\gamma t} q_t) \phi, e_z) \right) dt \\ & + \int_0^T \left(-2\sigma^2(\phi, e) - \frac{\sigma^2}{2}((z - e^{-\gamma t} q_t)^2 \phi_z, e_z) - \sigma^2((z - e^{-\gamma t} q_t) \phi_z, e) \right) dt = 0. \end{aligned} \quad (3.14)$$

Note that integration by parts gives

$$\sigma^2((z - e^{-\gamma t} q_t) \phi_z, e) = -\sigma^2((z - e^{-\gamma t} q_t) \phi, e_z) - \sigma^2(\phi, e), \quad (3.15)$$

using this identity, $\phi(0) = \psi$, and $e(T) = 0$ we get

$$\begin{aligned} & (\psi, e(0)) = \\ & - \int_0^T \left((\phi, e_t) - \sigma^2((z - e^{-\gamma t} q_t) \phi, e_z) - \frac{\sigma^2}{2}((z - e^{-\gamma t} q_t)^2 \phi_z, e_z) \right) dt. \end{aligned} \quad (3.16)$$

Recalling the earlier defined bilinear form (3.7) we can also write

$$(\psi, e(0)) = - \int_0^T \left((e_t, \phi) + a(e, \phi) \right) dt. \quad (3.17)$$

Since $e = u - U$ and u solves equation (3.6) we get the error representation formula

$$\boxed{(\psi, e(0)) = \int_0^T \left((\bar{U}_t, \phi) + a(\bar{U}, \phi) \right) dt} \quad (3.18)$$

If we for example are interested in the error at $z = z_\alpha$, we choose $\psi = \delta_{z_\alpha}(z)$, and get the error representation formula

$$e(0, z_\alpha) = \int_0^T \left((\bar{U}_t, \phi) + a(\bar{U}, \phi) \right) dt. \quad (3.19)$$

If one instead is interested in derivatives of the solution, then a different ψ is chosen, as shown later on.

3.3.2 Estimating the Error

Let $\pi : \mathcal{W} \rightarrow \mathcal{W}^{q-1}$ be the L_2 projection in time, and let P be a suitable interpolation operator into \mathcal{V}^p in space. Thus πP is an interpolation operator such that $\pi P \phi \in \mathcal{W}^{q-1}$. Then using Galerkin orthogonality (3.10), we can replace ϕ by $\phi - \pi P \phi = \phi - P \phi + P \phi - \pi P \phi$. Note that $P \phi = 0$ on $\mathbb{R} \setminus \Omega$. Equation (3.18) can then be written as

$$\begin{aligned} (\psi, e(0)) &= - \int_0^T \left((\bar{U}_t, \phi - P \phi) + a(\bar{U}, \phi - P \phi) \right) dt \\ &\quad - \int_0^T \left((\bar{U}_t, P \phi - \pi P \phi) + a(\bar{U}, P \phi - \pi P \phi) \right) dt \\ &= - \sum_n \sum_j \int_{I_n} \left(R_{\kappa_j}^s(\bar{U}), \phi - P \phi \right) dt \\ &\quad - \sum_n \int_{I_n} \left(R^t(\bar{U}), P \phi - \pi P \phi \right) dt - \sum_n \int_{I_n} \left(R^b(\bar{U}), \phi \right) dt, \end{aligned} \quad (3.20)$$

where

$$\begin{aligned} (R_{\kappa_j}^s(\bar{U}), \phi - P \phi) &= - \frac{\sigma^2}{4} ((z - e^{-\gamma t} q_t)^2 [\bar{U}_z], \phi - P \phi)_{\partial \kappa_j \setminus \partial \Omega} \\ &\quad + (\bar{U}_t + \frac{\sigma^2}{2} (z - e^{-\gamma t} q_t)^2 \bar{U}_{zz}, \phi - P \phi)_{\kappa_j}, \end{aligned} \quad (3.21)$$

is the space residual,

$$(R^t(\bar{U}), P\phi - \pi P\phi) = (\bar{U}_t + \frac{\sigma^2}{2}(z - e^{-\gamma t} q_t)^2 \bar{U}_{zz}, P\phi - \pi P\phi)_{\kappa_j}, \quad (3.22)$$

is the time residual, and

$$(R^b(\bar{U}), \phi) = -\frac{\sigma^2}{2}((z - e^{-\gamma t} q_t)^2 [\bar{U}_z], \phi)_{\partial\Omega}, \quad (3.23)$$

is the boundary residual accounting for the effect of restricting the computation from \mathbb{R} to the finite interval Ω . Note that the residuals are zero outside of Ω since $u = z$ and $u = 0$ satisfies equation (2.18). Here we used the notation $[\bar{U}_z]$ to denote the jump in \bar{U}_z over element interfaces.

Finally, we present an algorithm for calculating the error.

Error estimation algorithm:

- Compute an approximation Φ of ϕ using an enriched finite element space, for instance higher order approximation.
- Compute $P\Phi$.
- Compute $\int_{I_n} (R_{\kappa_j}^s(U), \phi - P\phi) dt$ using quadrature in space and time for each element κ_j and time step.
- Compute $\pi P\Phi$.
- Compute $\int_{I_n} (R^t(U), P\phi - \pi P\phi) dt$ using quadrature in space and time for each time step.

3.3.3 Examples

Using the error estimation algorithm in the previous section we are able to calculate the error in desired quantities for different values of the parameters. This makes it possible to identify regions where a fine mesh is necessary.

Example 1. To estimate the error at $z = z_\alpha$ we let $\psi = \delta_{z_\alpha}(z)$ in (3.11). In order to implement this condition we use the approximation

$$\delta_{z_\alpha}(z) \approx \frac{1}{\epsilon\sqrt{\pi}} e^{-((z-z_\alpha)/\epsilon)^2} := \delta_{z_\alpha}(z, \epsilon), \quad (3.24)$$

where ϵ is a parameter that controls how well the delta function is approximated. In this example we have used $\epsilon = 0.0129$. As seen from Figure 2, the solution to the dual problem differs from zero only within a short interval of Ω . We now check that the error representation formula really works

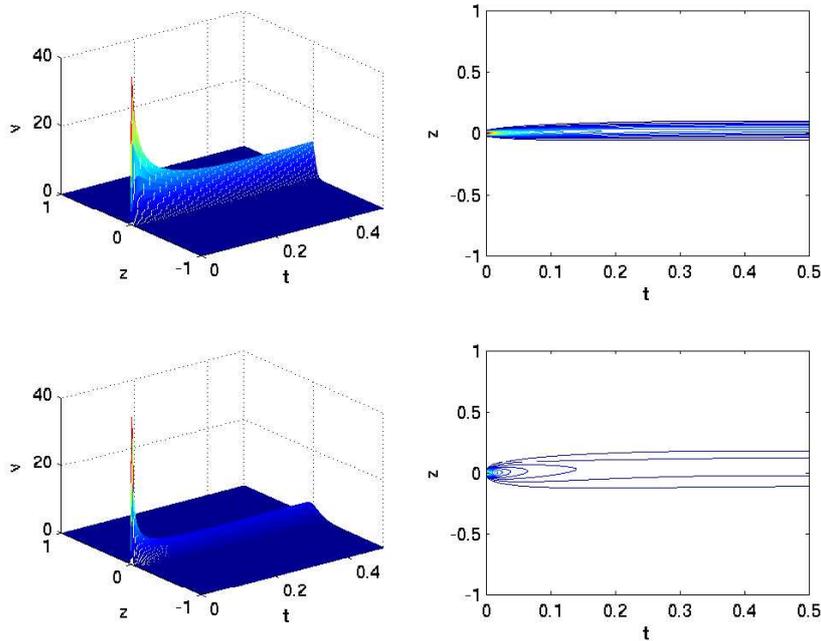


Figure 2: *The dual solution ϕ of an average rate call option. Above on the left, ϕ , for $\psi = \delta_0(z, \epsilon)$ with $\sigma = 0.1$ and $r = 0.1$, and $r=0.1$. Below on the left, ϕ , for $\sigma = 0.3$ and $r = 0.1$. On the right, contour plots using 30 levels. Solutions computed using the cG(2)-dG(1) method with 200 space and time points.*

by testing it in the case of the European option, where we know the exact solution. Recall that we get the value of an European call option by letting $q = e^{\gamma(t-T)} = 1$ in (2.18). By using the error estimation algorithm in the previous section we can get an approximation of the functional of the error, that is an approximation of the right hand side of equation (3.18). This can then be compared to calculating the left hand side of equation (3.18) directly using the real error in the approximate solution, found by using Black-Scholes formula and equation (2.17). The dual solution is calculated on a finer mesh, and using higher order approximations. The primal was

calculated using the cG(1)-cG(1) method with 640 space and time points, and the dual using the cG(2)-dG(1) method with 40 space and time points. The dual mesh was sixteen times finer in each direction. The value of the functional of the error found by using the error representation formula was in this case 0.000134, in excellent agreement with the real value, that is the value of the left hand side of equation (3.18), which was 0.000133.

We now proceed to do the same for the Asian option. By using the error estimation algorithm in the previous section we can get an approximation of the functional of the error, that is an approximation of the right hand side of equation (3.18). The dual solution is calculated on a finer mesh, and using higher order approximations. In Figure 3, we see the contributions to error formula (3.18) from each space-time slab. The primal was calculated using the cG(1)-cG(1) method, and the dual using the cG(2)-dG(1) method. The dual mesh was sixteen times finer in each direction. The value of the functional of the error found by using the error representation formula was in this case 0.0025. We also note that the contribution to the error differs from zero only within a short interval of Ω , just as the dual solution. This means that we may use a more sparse mesh where the contribution to the error is small and thus save computation time. The solution is larger near time $t = 0$, implying that one should use a finer time step there. Obviously the result depends on the value of the volatility σ , and the other parameters, which can be seen from the plot of the dual solution. We will later see how we can use the error representation formula to derive an optimal mesh for each problem.

Example 2. In order to make a good estimation of the derivative of the solution, which is interesting when calculating the greek *delta* (see Section 5), we need to study a different dual problem. We approximate the derivative using the central difference formula

$$\frac{\partial u}{\partial z} \approx \frac{u(z + \mu) - u(z - \mu)}{2\mu} := \frac{\partial_h u}{\partial z}. \quad (3.25)$$

To estimate the error of the derivative of the solution at $z = z_\alpha$, $u_z(z_\alpha)$, we thus choose

$$\begin{aligned} \psi(z) &= \frac{\delta_{z_\alpha}(z - \mu) - \delta_{z_\alpha}(z + \mu)}{2\mu} \\ &\approx \frac{\delta_{z_\alpha}(z - \mu, \epsilon) - \delta_{z_\alpha}(z + \mu, \epsilon)}{2\mu} \end{aligned} \quad (3.26)$$

in (3.11), for an appropriate choice of μ . The error in our estimation of the derivative can be split into two parts

$$\left(\frac{\partial u}{\partial z} - \frac{\partial_h u}{\partial z} \right) = \left(\frac{\partial u}{\partial z} - \frac{\partial_h u}{\partial z} \right) + \left(\frac{\partial_h u}{\partial z} - \frac{\partial_h U}{\partial z} \right). \quad (3.27)$$

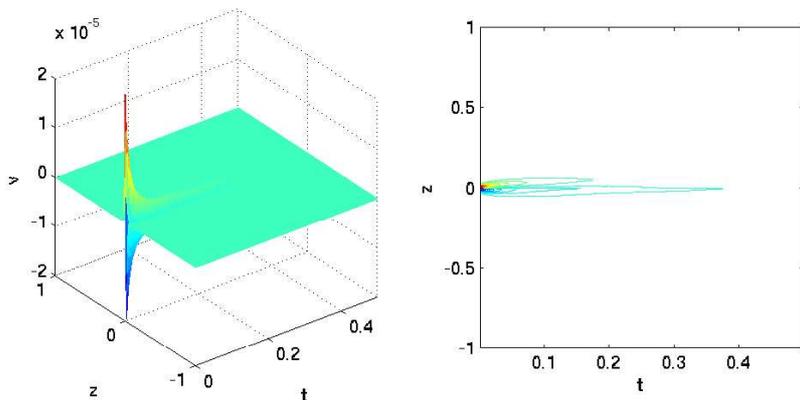


Figure 3: *On the left, the contributions to the error of an average rate call option for $\sigma = 0.1$, and $r = 0.1$, when $\psi = \delta_0(z, \epsilon)$. On the right, contour plot using 30 levels. The dual was computed using the $cG(2)$ - $dG(1)$ method with 640 space and time points, and the primal using the $cG(1)$ - $cG(1)$ method with 40 space and time points.*

The first term corresponds to the error in (3.25), while the second can be estimated using the a posteriori estimate. Figure 4 shows the dual solution for this choice of ψ when $\alpha = 0$, $\mu = 0.02$ and $\epsilon = 0.0129$. Figure 5 shows the contributions to the error estimation formula from each space-time slab. The same numerical methods and meshes were used as in the previous example. We see that this solution also is centrally oriented, implying that the derivative has a local dependence.

4 Adaptive Mesh Refinement

Adaptive mesh refinement may be accomplished in many different ways. Our goal not is to create the best adaptive method, since adaptivity would be to slow to use in reality. Rather we wish to create an optimal mesh

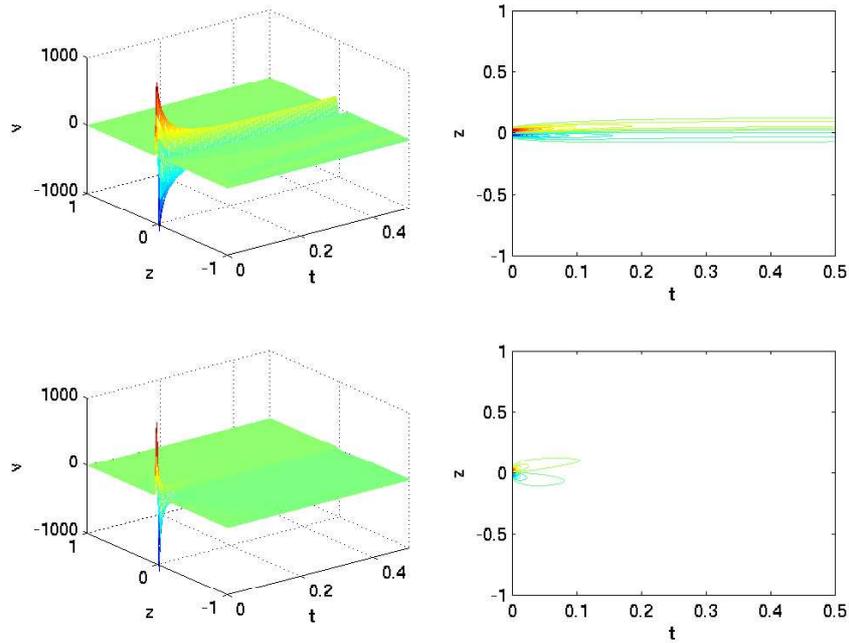


Figure 4: Above on the left, ϕ , for $\sigma = 0.1$, and $r = 0.1$, when ψ is chosen as in example 2. Below on the left, ϕ , for $\sigma = 0.3$. On the right, contour plots using 30 levels. Solutions computed using the $cG(2)$ - $dG(1)$ method with 200 space and time points.

in advance for each case, so that when valuing an option we simply use a suited pre-calculated mesh. This gives superior performance. In this section we show how these meshes are calculated and what typical meshes look like.

Mesh refinement algorithm:

- Compute an approximation U of u using the FE method on a coarse mesh.
- Compute the error in desired quantities by using the a posteriori error estimation algorithm.
- Calculate the time and space averages of the contributions to the error from each space-time slab. This gives us two vectors, one with time

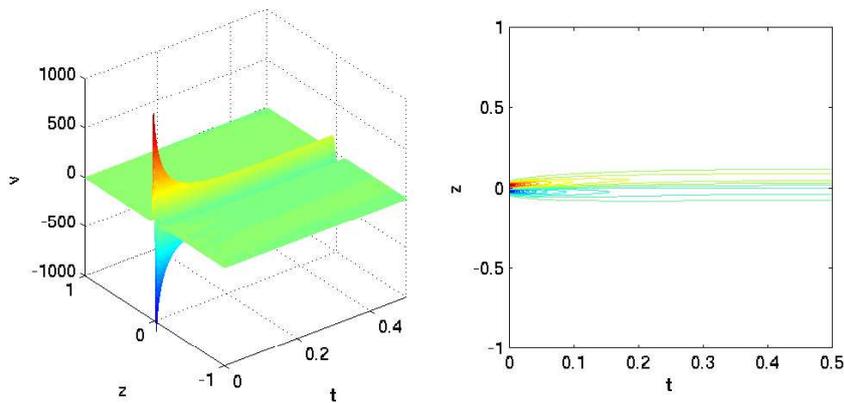


Figure 5: *The contributions to the functional of the error of an average rate call option for $\sigma = 0.1$, and $r = 0.1$, when ψ is chosen as in Example 2. The dual was computed using the $cG(2)$ - $dG(1)$ method with 640 space and time points, and the primal using the $cG(1)$ - $cG(1)$ method with 20 space and time points.*

averages and one with space averages.

- Identify the $Q\%$ largest elements in the space average vector, and refine the corresponding time steps by dividing them in half.
- Identify the $Q\%$ largest elements in the time average vector, and refine the corresponding spatial steps by dividing them in half.
- Compute a new FE approximation U on the refined mesh.
- Repeat until minimum mesh size is reached.

In Figure 6, we see a typical mesh resulting from using the mesh refinement algorithm above. In this case Q was set to 10%. Three successive refinements were made, starting from a sparse mesh with 20 nodes in time and space. The final mesh has only 27 nodes in each direction, but the error has decreased by a factor 25. The dual was calculated using a fine mesh with 640 nodes in time and space.

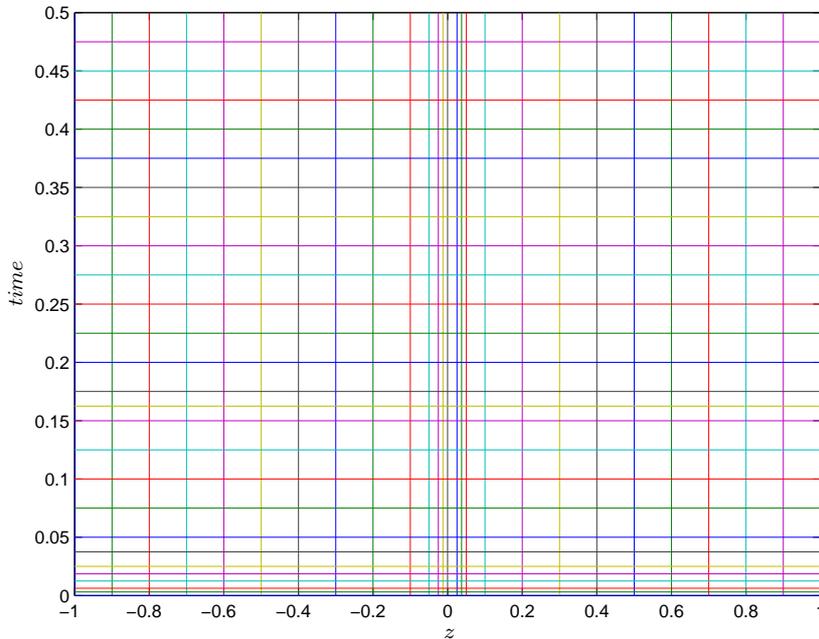


Figure 6: *The resulting mesh using the mesh refinement algorithm, calculated for a average rate call option with $\sigma = 0.1$, and $r = 0.1$, when $\psi = \delta_0(z, \epsilon)$. The dual was computed using the $cG(2)$ - $dG(1)$ and the primal using the $cG(1)$ - $cG(1)$ method. Three successive refinements were made.*

5 The Greeks

In order to hedge our Asian option, we need the sensitivity measures, or the so called greeks. Recalling that $Z_0 = \frac{X_0}{S_0}$ and that $\frac{dX_0}{dS_0} = q_0$ according to equation (2.9), we get by direct calculation

$$\frac{\partial Z_0}{\partial S_0} = \frac{1}{S_0}(q_0 - Z_0). \quad (5.1)$$

Using the chain rule, equation (5.1), and that the price of the Asian option is given in terms of u by the equation

$$V(0, S_0, K_1, K_2) = S_0 u(0, Z_0), \quad (5.2)$$

we get the three greeks at time $t = 0$

$$\Delta = \frac{\partial V}{\partial S_0} = u + (q_0 - Z_0) \frac{\partial u}{\partial Z_0}, \quad (5.3)$$

$$\Gamma = \frac{\partial^2 V}{\partial S_0^2} = \frac{1}{S_0} (q_0 - Z_0)^2 \frac{\partial^2 u}{\partial Z_0^2}, \quad (5.4)$$

$$\Theta = -\frac{\partial V}{\partial t} = -\frac{\partial u}{\partial t}. \quad (5.5)$$

In Figure 7 we see the delta of an average rate call at time $t = 0$ for various strike prices K .

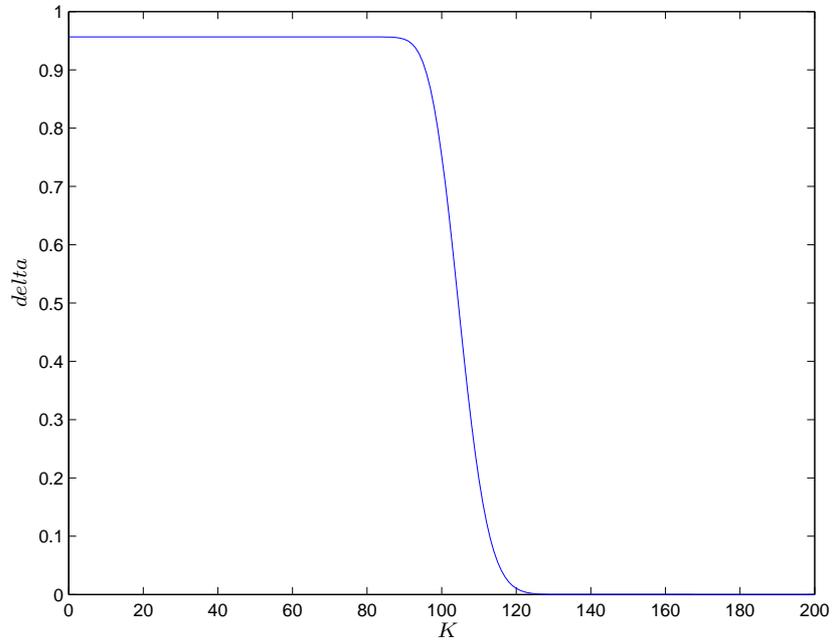


Figure 7: *The delta of an average rate call at time $t = 0$, with $\sigma = 0.3$, $T = 1$, $S_0 = 100$, and $r = 0.1$.*

6 Results

The implementation was done in C++ on a Dell Inspiron PC (700MHz). Many different meshes were used, with both uniform and adapted meshes. In the following we will use $z_0 = -1$ and $z_J = 1$, the accuracy is not improved if a larger interval is used.

We begin by validating our method against the known exact solution for the European call option. Recalling the previous calculations in Example 1 in Section 3.3.3 we know that the error representation formula works and is accurate. The value of the functional of the error found by using the error representation formula in the test example on the European option was 0.000134, in excellent agreement with the real value of the functional found by using Black-Scholes formula, which was 0.000133. Table 2 compares values of the European call calculated using the cG(1)-cG(1) finite element method, with the analytical value derived by Black-Scholes formula. We see that the FE method is very stable and has a maximum relative error of 0.06 percent when 400 time points are used.

σ	K	FE(200)	FE(400)	Black-Scholes	Relative error (%)
0.10	90	14.6207	14.6268	14.6288	0.0137
	100	6.7972	6.8030	6.8050	0.0294
	110	2.1687	2.1726	2.1739	0.0598
0.20	90	16.6983	16.6981	16.6994	0.0078
	100	10.4468	10.4496	10.4506	0.0096
	110	6.0375	6.0395	6.0401	0.0099
0.30	90	19.6932	19.6965	19.6974	0.0046
	100	14.2273	14.2304	14.2313	0.0063
	110	10.0148	10.0189	10.0201	0.0120

Table 2: *The European call calculated using the FE method with 200 and 400 time points compared to Black-Scholes analytical value when $r = 0.05$, $T = 1$ and $t = 0$. The relative error is between the FE(400) solution and the analytical solution.*

In Figure 8 we see the average rate call option value calculated using the adapted mesh from the previous section. The mesh is finer close to the time $t = 0$ and in the center of the spatial interval Ω . In this way higher accuracy is achieved without dramatically increasing the number of space and time points in the mesh. The original uniform mesh has 20 nodes in time and space. By using the error representation formula, the functional of the error was calculated to 0.0025 for the uniform mesh. The adapted mesh has only

27 nodes in each direction, but the functional of the error has decreased by a factor 25 to 0.0001.

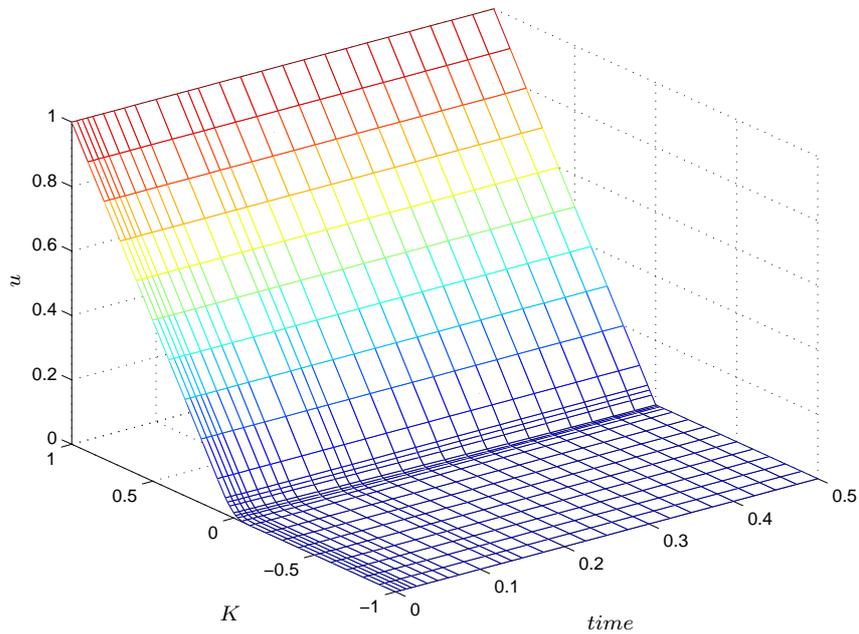


Figure 8: *An average rate call option with $r = 0.10$, $\sigma = 0.10$, $T = 1$ and $t = 0$. Computed on an adapted mesh with 27 space and time points.*

Table 3 compares the results of the method developed in this paper with the results of Večer [18], Zvan, Forsyth, and Vetzal [20], and Rogers, and Shi [14]. To be consistent with their results a uniform mesh with same number of time and space points (200 space points and 400 time points) was used in the computation of the finite element results in Table 3. The Monte Carlo results were obtained from Večer [18], and the lower and upper bounds are from Rogers and Shi [14]. The last column gives the value of the greek delta. As seen from the table all methods are accurate and always give answers within analytical bounds. The most important difference between them is the computation time required to receive the results. It takes approximately 0.05 seconds of CPU time to calculate the price using this uniform mesh with 200 space points and 400 time points. Using an adapted mesh instead we could achieve the same accuracy but with a coarser mesh, and thus speed

up the calculation significantly.

σ	K	Foufas	Večer	Zvan et al.	Monte Carlo	Lower	Upper	Δ
0.05	95	11.112	11.112	11.094	11.094	11.094	11.114	0.929
	100	6.810	6.810	6.793	6.795	6.794	6.810	0.925
	105	2.754	2.750	2.748	2.745	2.744	2.761	0.764
0.10	90	15.416	15.416	15.399	15.399	15.399	15.445	0.928
	100	7.042	7.036	7.030	7.028	7.028	7.066	0.842
	110	1.422	1.421	1.410	1.418	1.413	1.451	0.355
0.20	90	15.659	15.659	15.643	15.642	15.641	15.748	0.879
	100	8.427	8.424	8.409	8.409	8.408	8.515	0.703
	110	3.570	3.568	3.554	3.556	3.554	3.661	0.422
0.30	90	16.533	16.533	16.514	16.516	16.512	16.732	0.806
	100	10.231	10.230	10.210	10.210	10.208	10.429	0.644
	110	5.750	5.748	5.729	5.731	5.728	5.948	0.451

Table 3: Comparison of results of different methods for the average rate call with $r = 0.15$, $S_0 = 100$, $T = 1$ and $t = 0$. The Monte Carlo results are from Večer [18] and the lower and upper bounds are from Rogers and Shi [14]. Δ refers to the value of the greek delta, calculated using the FEM method.

References

- [1] M. AINSWORTH AND J. T. ODEN, *A posteriori error estimation in finite element analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2000.
- [2] B. ALZIARY, J. DECAMPS, AND P. KOEHL, *A P.D.E. approach to Asian options: analytical and numerical evidence*, *Journal of Banking and Finance*, 21 (1997), pp. 613–640.
- [3] J. ANDREASEN, *The pricing of discretely sampled Asian and lookback options: a change of numeraire approach*, *The Journal of Computational Finance*, 2 (Fall 1998), pp. 5–30.
- [4] W. BANGERTH AND R. RANNACHER, *Adaptive finite element methods for differential equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2003.
- [5] Y. BERGMAN, *Pricing path contingent claims*, *Research in Finance*, 5 (1985), pp. 229–241.
- [6] K. ERIKSSON, D. ESTEP, P. HANSBO, AND C. JOHNSON, *Computational Differential Equations*, Studentlitteratur, 1996.
- [7] D. ESTEP, M. LARSON, AND R. WILLIAMS, *Estimating the error of numerical solutions of systems of reaction-diffusion equations*, MEMOIRS of the American Mathematical Society, 146 (2000).
- [8] G. FOUFAS, *PDE-methods for Asian options*, Master’s thesis, Chalmers University of Technology, 2000.
- [9] M. FU, D. MADAN, AND T. WANG, *Pricing continuous Asian options: a comparison of Monte Carlo and Laplace transform inversion methods*, *The Journal of Computational Finance*, 2 (Winter 1998/99).
- [10] H. GEMAN AND M. YOR, *Bessel processes, Asian options, and perpetuities*, *Math. Finance*, 3 (1993), pp. 349–375.
- [11] J. INGERSOLL, *Theory of Financial Decision Making*, Oxford, 1987.
- [12] V. LINETSKY, *Exact pricing of Asian options: An application of spectral theory*. working paper, 2002.
- [13] I. NELKEN, *The Handbook of Exotic Options*, Irwin Professional Book team, 1996.

- [14] L. ROGERS AND Z. SHI, *The value of an Asian option*, Journal of Applied Probability, 32 (1995), pp. 1077–1088.
- [15] S. SHREVE AND J. VEČEŘ, *Options on a traded account: Vacation calls, vacation puts and passport options*, Finance and Stochastics, (2000).
- [16] G. THOMPSON, *Fast narrow bounds on the value of Asian options*. Working paper, 1999.
- [17] J. TOPPER, *Financial Engineering with finite elements*, Wiley Finance, 2005.
- [18] J. VEČEŘ, *A new PDE approach for pricing arithmetic average Asian options*, The Journal of Computational Finance, (2000).
- [19] ———, *Unified Asian pricing*, RISK, (2002), pp. 113–116.
- [20] R. ZVAN, P. FORSYTH, AND K. VETZAL, *Robust numerical methods for PDE models of Asian options*, The Journal of Computational Finance, 1 (Winter 1997/98).

Paper IV

A Note on the Connection Between the Greeks and A Posteriori Error Analysis

Georgios Foufas* Mats G. Larson†

April 14, 2008

Abstract

In this paper we present a new connection between some of the sensitivity measures, also known as the Greeks, and a finite element based a posteriori error analysis. This is not only a nice feature of the a posteriori error analysis but it also gives us an alternative way of calculating these Greeks. The presented error estimation formula splits the error in parts originating from how good the numerical approximation is and in parts originating from how well the parameters are approximated. The study is based on the finite element method applied to the European option problem, but the technique is general and can be applied to other option valuation problems as well.

1 Introduction

The valuation of different types of derivative contracts is very important in modern financial theory and practice. Not only the option price itself is important to calculate in a fast a stable manner, but also certain sensitivity measures, or the so called Greeks. The reason is that these Greeks are used when hedging the options.

The finite element method is widely used in other fields as a tool for finding approximate solutions to partial differential equations (PDE) as well as of integral equations. It was developed in the 1950's and 1960's by engineers, and was mainly used in structural mechanics, see e.g. [11] for an overview. The finite element method also has a strong mathematical foundation in functional analysis, see [1]. The mathematical foundation provides the tools to derive analytical error estimates which can be used in a constructive way

*Research Assistant, Department of Mathematics, Chalmers University of Technology, S-412 96 Gothenburg, Sweden, *email*: fofas@math.chalmers.se

†Professor of Applied Mathematics, Corresponding author, Department of Mathematics, Umeå University, S-901 87 Umeå, Sweden, *email*: mats.larson@math.umu.se

to improve the approximative solution. In finance it has not been used that frequently yet compared to other methods such as the finite difference method. Recently Topper [8] wrote an excellent book applying the finite element method to different option pricing problems. As noted by Topper the finite element method is well suited for calculating the Greeks since it gives a polynomial approximation in the spatial variables. The derivatives of polynomials can easily be calculated analytically and as a result very fast. In [4], [5], and [6] the authors apply an adaptive finite element method to different option pricing problems. The adaptive finite element method is based on piecewise polynomial approximations in space and time. The a posteriori estimates for the error in point wise values of the solution and it's derivatives are calculated using duality techniques. The estimates are used to determine suitable local resolution in space and time. In this paper we extend the previously developed a posteriori error analysis to include also changes in the parameters in the partial differential equation. As a bonus we receive a new way of calculating some of the Greeks.

The remainder of the paper is organized as follows: in Section 2 we give a very short mathematical background and present the model problem, the Black-Scholes equation. In Section 3 we introduce the Greeks and discuss how to calculate them. Then in Section 4 we formulate the finite element method and apply it to the European option. In Section 5 we derive a new a posteriori error estimate for the ordinary European option with a connection to some of the Greeks, and present a new way of calculating these Greeks. Finally in Section 6 we give some conclusions.

2 Mathematical Background

As a model problem we choose to study the European option and the Black-Scholes model. This equation can be solved analytically. At the same time it can be used with minor changes to value other exotic options such as the barrier option and the lookback option. This makes it suitable as a model problem for demonstrating our idea. We might as well have chosen to demonstrate the technique on another option with a different pricing PDE, such as the Asian option, also studied by the authors [6].

We consider a continuous time trading economy on a bounded time horizon $[0, T]$. Probability is represented by the probability space $(\Omega_T, \mathcal{F}_T, P)$, where $\Omega_T = C[0, T]$, P is the corresponding Wiener measure, and $\mathcal{F}_T = \sigma(W(t); t \leq T)$. For simplicity we consider the standard Black-Scholes setting with a risk free asset and a dividend paying stock. Let $B(t)$ denote the price of a risk free asset at time t governed by the equation $B(t) = B(0)e^{rt}$, where r is the constant interest rate. Further we denote by $S(t)$ the value

of an asset at time t . We assume the existence of an equivalent martingale measure Q , under which the discounted stock price $e^{-r(T-t)}S_t$ is an \mathcal{F}_t -martingale. The existence of the risk neutral measure Q assures that the market is free of arbitrage possibilities. Under Q the stock price follows the stochastic differential equation

$$dS(t) = (r - \nu)S(t)dt + S(t)\sigma dW(t), \quad (2.1)$$

where r is the constant interest rate, ν is the constant continuous dividend yield, σ is the volatility, and $W(t)$ is a Q Brownian motion process. Here σ is assumed to be a positive real number. The solution of (2.1) is

$$S(t) = S(0)e^{(r-\nu-\frac{\sigma^2}{2})t+\sigma W(t)}. \quad (2.2)$$

The value of the ordinary European option, $u(t, S(t)) = u(t, s)$, is given as the solution to Black-Scholes equation

$$u_t(t, s) + \frac{\sigma^2 s^2}{2} u_{ss}(t, s) + (r - \nu)su_s(t, s) - ru(t, s) = 0, \quad t < T, \quad (2.3)$$

which is valid for $s = S(t) \in \mathbb{R}^+$.

3 The Greeks

The different sensitivity measures of options and other derivatives are commonly referred to as Greeks, because they are often denoted by Greek letters. The Greeks are very important tools in risk management, especially since they are used for hedging purposes. Each Greek measures a different type of risk associated with an option position (with the exception of theta). A portfolio of options can be adjusted according to these Greeks (hedged) to achieve a desired exposure. Therefore financial market models possessing the property of easy computation of the Greeks are desirable. In the Black-Scholes model the Greeks are very easy to calculate and this is one reason for the model's continued popularity in the market.

Let V denote the value of an option or an other derivative, and let s denote the price of the underlying asset. The *delta* measures the sensitivity to changes in the price of the underlying asset and is calculated as the derivative of V regarding to the underlying's price s ,

$$\Delta = \frac{\partial V}{\partial s}. \quad (3.1)$$

The rate of change of *delta* is called *gamma*, defined as

$$\Gamma = \frac{\partial^2 V}{\partial s^2}, \quad (3.2)$$

and the *speed* is the third derivative of V with regarding to s ,

$$\frac{\partial^3 V}{\partial s^3}. \quad (3.3)$$

The decay of value in time of a portfolio is represented by the *theta*, where

$$\Theta = -\frac{\partial V}{\partial t}. \quad (3.4)$$

Sensitivity to volatility is called *vega* and is defined by

$$vega = \frac{\partial V}{\partial \sigma}, \quad (3.5)$$

and sensitivity to interest rate is called *rho*, defied as

$$\rho = \frac{\partial V}{\partial r}. \quad (3.6)$$

3.1 Calculation of the Greeks

As noted by Topper [8] the finite element method is well suited for calculating the Greeks since it gives a polynomial approximation in the spatial variables. The derivatives of polynomials can easily be calculated analytically and as a result very fast. For this to work the finite element shape functions must of course be at least of the same order as the order of derivative we wish to calculate. For example, to calculate the Greek γ we need at least quadratic shape functions. One can improve the estimates by taking the Greeks at so called Moan Points, which are points were the derivatives of the finite element approximation have higher accuracy, see [7]. However, when calculating the Greeks *vega* and *rho* we need to use a different approach since they involve derivatives with respect to the parameters σ and r , and not the stock price s . For these we apply a sensitivity analysis approach described in Section 5.2. As noted by Wilmott the Greeks *vega* and *rho* should be used with care since they for some options lack financial meaning, see [9].

4 A Finite Element Method for the European Option

In order to construct a computational mesh we introduce a bounded interval $\Omega = [s_{min}, s_{max}] \subset \mathbb{R}^+$ with boundary $\partial\Omega = \{s_{min}, s_{max}\}$. We define the usual Hilbert space

$$H^1(\Omega) = \{v : \int_{\Omega} (|\nabla v|^2 + v^2) ds < \infty\}, \quad (4.1)$$

and let \mathcal{W} be the space of functions that are square integrable in time and belongs to $H^1(\Omega)$ in space, that is

$$\mathcal{W} = L^2([0, T], H^1(\Omega)). \quad (4.2)$$

We also use the notation $(u, v) = \int_{\Omega} uv ds$, and $(u, v)_{\partial\Omega} = u(s_{max})v(s_{max}) - u(s_{min})v(s_{min})$.

4.1 Variational Formulation

Multiplying the Black-Scholes equation (2.3) by the test function $v \in \mathcal{W}$ and integrating on $\Omega \times [0, T]$ we obtain

$$\int_0^T \left((u_t, v) + (r - \nu)(su_s, v) + \frac{\sigma^2}{2} (s^2 u_{ss}, v) - r(u, v) \right) dt = 0. \quad (4.3)$$

Using integration by parts we get

$$(s^2 u_{ss}, v) = (s^2 u_s, v)_{\partial\Omega} - 2(su_s, v) - (s^2 u_s, v_s). \quad (4.4)$$

Thus equation (4.3) becomes

$$\int_0^T \left((u_t, v) + (r - \nu - \sigma^2)(su_s, v) - \frac{\sigma^2}{2} (s^2 u_s, v_s) + \frac{\sigma^2}{2} (s^2 u_s, v)_{\partial\Omega} - r(u, v) \right) dt = 0. \quad (4.5)$$

The boundary conditions for the European call option are $u(t, 0) = 0$ and $u(t, s) \sim se^{-\nu(T-t)}$ as $s \rightarrow \infty$, and for the corresponding put $u(t, 0) = Ke^{-r(T-t)}$ and $u(t, s) \sim 0$ as $s \rightarrow \infty$, see for example Wilmott, [10]. For simplicity of implementation we use the artificial boundary condition $u_{ss} = 0$ on $\partial\Omega$ for both the put and the call instead. This boundary condition works for all contracts if the payoff is at most linear in the underlying (see [10]) and does not affect the accuracy of the solution. Using equation (2.3) we can rewrite the boundary condition as

$$u_s = \frac{r}{s(r - \nu)} u - \frac{1}{s(r - \nu)} u_t, \quad (4.6)$$

and enforce it weakly by inserting identity (4.6) into equation (4.5). We thus want to solve the problem: find $u \in \mathcal{W}$ such that

$$\begin{cases} \int_0^T (m(u_t, v) + a(u, v)) dt = 0, \\ u(T, s) = \begin{cases} \max(s - K, 0), & \text{for a call,} \\ \max(K - s, 0), & \text{for a put,} \end{cases} \end{cases} \quad (4.7)$$

for every $v \in \mathcal{W}$, where

$$m(u_t, v) = (u_t, v) - \frac{\sigma^2}{2(r - \nu)}(su_t, v)_{\partial\Omega}, \quad (4.8)$$

and

$$\begin{aligned} a(u, v) &= (r - \nu - \sigma^2)(su_s, v) - \frac{\sigma^2}{2}(s^2u_s, v_s) \\ &\quad + \frac{\sigma^2 r}{2(r - \nu)}(su, v)_{\partial\Omega} - r(u, v). \end{aligned} \quad (4.9)$$

4.2 Finite Element Approximation

The finite element method is based on solution of the variational problem (4.7) with \mathcal{W} replaced by a finite dimensional function space of piecewise polynomials in space and time. For background on the finite element method see for instance [2].

We now partition $[0, T]$ as $0 = t_0 < t_1 < t_2 < \dots < t_N = T$, denoting each time interval by $I_n = (t_{n-1}, t_n]$ and each time step by $k_n = t_n - t_{n-1}$. Similarly we partition Ω as $s_{min} = s_0 < s_1 < s_2 < \dots < s_J = s_{max}$, denoting each spatial interval by $\kappa_j = [s_{j-1}, s_j)$ and the length of each interval by $h_j = s_j - s_{j-1}$.

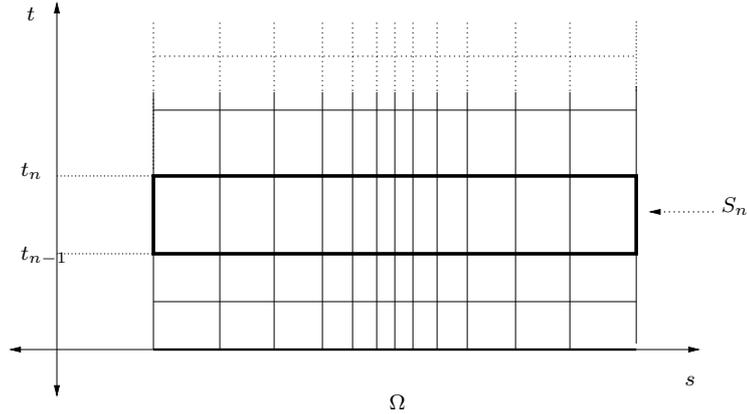


Figure 1: *Space-time discretization.*

In space, we let $\mathcal{V}^p \subset H^1(\Omega)$ denote the space of piecewise continuous

functions of order p . On each space-time slab $S_n = I_n \times \Omega$, we define

$$\mathcal{W}_n^q = \{w(t, s) : w(t, s) = \sum_{j=0}^q t^j v_j(s), v_j \in \mathcal{V}^p, (t, s) \in S_n\}. \quad (4.10)$$

Let $\mathcal{W}^q \subset \mathcal{W}$ denote the space of functions defined on $[0, T] \times \Omega$ such that $v|_{S_n} \in \mathcal{W}_n^q$ for $1 \leq n \leq N$. For simplicity, we only give details for the continuous Galerkin method cG(p)-cG(q), (see e.g. [2] or [3]) which is defined by the following discrete version of equation (4.7). Find $U \in \mathcal{W}^q$ such that for $1 \leq n \leq N$

$$\begin{cases} \int_{I_n} (m(U_t, v) + a(U, v)) dt = 0 & \text{for all } v \in \mathcal{W}_n^{q-1}, \\ U^-(t_n) = U^+(t_n), & n = N - 1, \dots, 1, \\ U^-(t_N) = u_T, \end{cases} \quad (4.11)$$

where $U^\pm(t_n) = \lim_{\epsilon \rightarrow 0, \epsilon > 0} U(t_n \pm \epsilon)$. In the cG(1) method the approximation U of u is continuous piecewise linear in time and space, while the test functions v are continuous linear in space and piecewise constant in time. It is also possible to use a discontinuous method in time, we refer to [2], for details on the resulting discontinuous Galerkin method, cG(p)-dG(q).

5 A Posteriori Error Estimation

5.1 Error Representation Formula

Since we are only interested in the solution, and its derivatives, in one or a few points of Ω at time $t = 0$, we wish to find a mesh tailored for efficient and accurate solution at the points of interest. In order to find such a mesh we derive a posteriori error estimates of the error in the points of interest using duality techniques (see [2] or [3]).

To represent the error in a linear functional, $(u - U, \psi)$, we introduce the continuous dual problem for the Black-Scholes equation (2.3). Let now $\bar{\sigma}$ and \bar{r} represent slightly perturbed versions of σ and r respectively. Find $\phi \in \mathcal{W}$ such that

$$\begin{cases} -\phi_t + (\bar{\sigma}^2 + \nu - 2\bar{r})\phi - (\bar{r} - \nu - 2\bar{\sigma}^2)s\phi_s + \frac{\bar{\sigma}^2}{2}s^2\phi_{ss} = 0, \\ \phi(0, s) = \psi. \end{cases} \quad (5.1)$$

For simplicity we consider this equation over the whole space interval neglecting boundary conditions. Multiplying with the error $e = u - U \in \mathcal{W}$

and integrating in space and time we get

$$\int_0^T \left(-(\phi_t, e) + (\bar{\sigma}^2 + \nu - 2\bar{r})(\phi, e) - (\bar{r} - \nu - 2\bar{\sigma}^2)(s\phi_s, e) + \frac{\bar{\sigma}^2}{2}(s^2\phi_{ss}, e) \right) dt = 0. \quad (5.2)$$

The functions ϕ and ϕ_s are in principle zero close to $s = s_{min}$ and $s = s_{max}$ if the domain is large enough. Using integration by parts and neglecting the boundary terms we get

$$\begin{aligned} & -(\phi(T, s), e(T, s)) + (\phi(0, s), e(0, s)) \\ & + \int_0^T \left((\phi, e_t) + (\bar{\sigma}^2 + \nu - 2\bar{r})(\phi, e) + (\bar{r} - \nu - 2\bar{\sigma}^2)(s\phi, e_s) \right) dt \\ & + \int_0^T \left((\bar{r} - \nu - 2\bar{\sigma}^2)(\phi, e) - \frac{\bar{\sigma}^2}{2}(s^2\phi_s, e_s) - \bar{\sigma}^2(s\phi_s, e) \right) dt = 0. \end{aligned} \quad (5.3)$$

Note that integration by parts gives

$$-\bar{\sigma}^2(s\phi_s, e) = \bar{\sigma}^2(s\phi, e_s) + \bar{\sigma}^2(\phi, e), \quad (5.4)$$

using this identity, $\phi(0, s) = \psi$, and $e(T) = 0$, we get

$$\begin{aligned} & (\psi, e(0, s)) \\ & = - \int_0^T \left((\phi, e_t) - \bar{r}(\phi, e) + (\bar{r} - \nu - \bar{\sigma}^2)(s\phi, e_s) - \frac{\bar{\sigma}^2}{2}(s^2\phi_s, e_s) \right) dt. \end{aligned} \quad (5.5)$$

Since $e = u - U$ we can rewrite (5.5) as

$$\begin{aligned} & (\psi, e(0, s)) \\ & = - \int_0^T \left((\phi, u_t) - r(\phi, u) + (r - \nu - \sigma^2)(s\phi, u_s) - \frac{\sigma^2}{2}(s^2\phi_s, u_s) \right) dt \\ & + \int_0^T \left((\phi, U_t) - \bar{r}(\phi, U) + (\bar{r} - \nu - \bar{\sigma}^2)(s\phi, U_s) - \frac{\bar{\sigma}^2}{2}(s^2\phi_s, U_s) \right) dt \\ & - \int_0^T \left(-(\bar{r} - r)(\phi, u) + ((\bar{r} - r) - (\bar{\sigma}^2 - \sigma^2))(s\phi, u_s) \right) dt \\ & - \int_0^T \left(-\left(\frac{\bar{\sigma}^2}{2} - \frac{\sigma^2}{2}\right)(s^2\phi_s, u_s) \right) dt \end{aligned} \quad (5.6)$$

Since u solves equation (4.7) we get

$$\begin{aligned}
(\psi, e(0, s)) & \tag{5.7} \\
&= \int_0^T \left((\phi, U_t) - \bar{r}(\phi, U) + (\bar{r} - \nu - \bar{\sigma}^2)(s\phi, U_s) - \frac{\bar{\sigma}^2}{2}(s^2\phi_s, U_s) \right) dt \\
&\quad + (\bar{r} - r) \int_0^T ((\phi, u) - (s\phi, u_s)) dt \\
&\quad + (\bar{\sigma} - \sigma) \left(\frac{(\bar{\sigma} + \sigma)}{2} \right) \int_0^T (2(s\phi, u_s) + (s^2\phi_s, u_s)) dt.
\end{aligned}$$

Recalling the earlier defined bilinear forms (4.8) and (4.9), and that we neglect the boundary terms we can also write

$$\begin{aligned}
(\psi, e(0, s)) &= \int_0^T \left(m(U_t, \phi) + a(U, \phi) \right) dt \tag{5.8} \\
&\quad + (\bar{r} - r) \int_0^T ((\phi, u) - (s\phi, u_s)) dt \\
&\quad + (\bar{\sigma} - \sigma) \left(\frac{(\bar{\sigma} + \sigma)}{2} \right) \int_0^T (2(s\phi, u_s) + (s^2\phi_s, u_s)) dt.
\end{aligned}$$

In order to simplify the notations we will from now on use the notation $e(0)$ to denote $e(0, s)$. In the same way $\phi(0)$, $u(0)$, $\rho(0)$ and $vega(0)$ will then naturally stand for $\phi(0, s)$, $u(0, s)$, $\rho(0, s)$ and $vega(0, s)$ respectively. We now want to show that equation (5.8) can be written as the following error representation formula

$$\begin{aligned}
(\psi, e(0, s)) &= \int_0^T \left(m(U_t, \phi) + a(U, \phi) \right) dt \tag{5.9} \\
&\quad + (\bar{r} - r)(\psi, \rho(0)) + (\bar{\sigma} - \sigma)(\psi, vega(0)),
\end{aligned}$$

that is we want to show that

$$(\psi, \rho(0)) = \int_0^T ((\phi, u) + (s\phi, u_s)) dt, \tag{5.10}$$

and that

$$(\psi, vega(0)) = \left(\frac{(\bar{\sigma} + \sigma)}{2} \right) \int_0^T (2(s\phi, u_s) + (s^2\phi_s, u_s)) dt, \tag{5.11}$$

where ρ and $vega$ actually are the ordinary Greeks known as $\rho = \frac{\partial u}{\partial r}$ and $vega = \frac{\partial u}{\partial \sigma}$. This will give us the opportunity to calculate these two Greeks in a new way using dual information. The proof is given below.

5.2 Calculating the Greeks using Duality

We begin the proof by re-writing our governing pricing partial differential equation (2.3) in the form

$$\frac{\partial}{\partial t} + \mathcal{L}u = 0, \quad (5.12)$$

where

$$\mathcal{L} = \frac{\sigma^2 s^2}{2} \frac{\partial^2}{\partial s^2} + (r - \nu)s \frac{\partial}{\partial s} - r(\cdot), \quad (5.13)$$

and u is the solution. Differentiating equation (5.12) with respect to r we get

$$\frac{\partial}{\partial t} \frac{\partial u}{\partial r} + \mathcal{L} \frac{\partial u}{\partial r} + \frac{\partial \mathcal{L}}{\partial r} u = 0. \quad (5.14)$$

Multiplying both sides with a testfunction ϕ , integrating in time and space we get

$$\int_0^T \left(\frac{\partial}{\partial t} \frac{\partial u}{\partial r} + \mathcal{L} \frac{\partial u}{\partial r}, \phi \right) dt = \int_0^T \left(-\frac{\partial \mathcal{L}}{\partial r} u, \phi \right) dt. \quad (5.15)$$

The dual operator to \mathcal{L} denoted \mathcal{L}^* is given by (see Section 5)

$$\mathcal{L}^* = \frac{\sigma^2 s^2}{2} \frac{\partial^2}{\partial s^2} - (r - \nu - 2\sigma^2)s \frac{\partial}{\partial s} - (\sigma^2 + \nu - 2r)(\cdot). \quad (5.16)$$

Using integration by parts we can rewrite the left hand side of equation (5.15) as

$$\begin{aligned} & \left[\left(\frac{\partial u}{\partial r}, \phi \right) \right]_0^T - \int_0^T \left(\frac{\partial u}{\partial r}, \frac{\partial \phi}{\partial t} \right) dt + \int_0^T \left(\frac{\partial u}{\partial r}, \mathcal{L}^* \phi \right) dt \\ &= \left(\frac{\partial u}{\partial r}(T), \phi(T) \right) - \left(\frac{\partial u}{\partial r}(0), \phi(0) \right) + \int_0^T \left(\frac{\partial u}{\partial r}, -\frac{\partial \phi}{\partial t} + \mathcal{L}^* \phi \right) dt \\ &= - \left(\frac{\partial u}{\partial r}(0), \phi(0) \right) + \int_0^T \left(\frac{\partial u}{\partial r}, -\frac{\partial \phi}{\partial t} + \mathcal{L}^* \phi \right) dt, \end{aligned} \quad (5.17)$$

where the last equality follows since $\frac{\partial u}{\partial r}(T) = 0$. Recalling the dual problem (5.1)

$$\begin{cases} -\frac{\partial \phi}{\partial t} + \mathcal{L}^* \phi = 0, \\ \phi(0) = \psi, \end{cases} \quad (5.18)$$

we can now rewrite equation (5.15) as

$$\begin{aligned} \int_0^T \left(\frac{\partial u}{\partial r}(0), \psi \right) dt &= - \int_0^T \left(\frac{\partial \mathcal{L}}{\partial r} u, \phi \right) dt \\ &= - \int_0^T \left(s \frac{\partial u}{\partial s} - u, \phi \right) dt. \end{aligned} \quad (5.19)$$

Differentiating equation (5.12) with respect to σ instead of r we can in the same way as above show that

$$\int_0^T \left(\frac{\partial u}{\partial \sigma}(0), \psi \right) dt = - \int_0^T \left(\frac{\partial \mathcal{L}}{\partial \sigma} u, \phi \right) dt = - \int_0^T \left(\sigma s^2 \frac{\partial^2 u}{\partial s^2}, \phi \right) dt. \quad (5.20)$$

Using integration by parts, neglecting boundary terms, we can rewrite equation (5.20) as

$$\int_0^T \left(\frac{\partial u}{\partial \sigma}(0), \psi \right) dt = \sigma \int_0^T \left(2s \frac{\partial u}{\partial s}, \phi \right) dt + \sigma \int_0^T \left(s^2 \frac{\partial u}{\partial s}, \frac{\partial \phi}{\partial s} \right) dt. \quad (5.21)$$

Summing up, we have shown that

$$(\psi, \rho(0)) = \left(\psi, \frac{\partial u}{\partial r}(0) \right) = \int_0^T ((\phi, u) + (s\phi, u_s)) dt, \quad (5.22)$$

and

$$(\psi, \text{vega}(0)) = \left(\psi, \frac{\partial u}{\partial \sigma}(0) \right) = \sigma \int_0^T (2(s\phi, u_s) + (s^2\phi_s, u_s)) dt, \quad (5.23)$$

that is (5.10) and (5.11) holds since $\frac{\bar{\sigma} + \sigma}{2} \simeq \sigma$.

5.3 Estimating the Error

If we for example are interested in the error at $s = s_\alpha$, we choose $\psi = \delta_{s_\alpha}(s)$, and get the error representation formula

$$\begin{aligned} e(0, s_\alpha) & \quad (5.24) \\ &= \int_0^T \left(m(U_t, \phi) + a(U, \phi) \right) dt + (\bar{r} - r)\rho(0, s_\alpha) + (\bar{\sigma} - \sigma)\text{vega}(0, s_\alpha) \end{aligned}$$

If one instead is interested in derivatives of the solution, then a different ψ is chosen. More details about this and how to calculate the error, error estimation algorithms, and mesh refinement can be found in [4].

6 Conclusions

We have presented a new connection between the a posteriori error analysis and the Greeks ρ and *vega*. The presented error representation formula, equation (5.9), splits the error in three parts, where the first part corresponds to the discretization error and the second and third parts corresponds to how well the interest rate and the volatility is estimated correspondingly.

Looking at the second term on the right hand side in detail we see that it very naturally includes the derivative of the interest rate, that is the Greek ρ , and what kind of data we are interested in, that is ψ . The same reasoning holds for the third term on the right hand side. This gives us the opportunity to calculate these two Greeks in a new way using dual information. This new way of calculating these Greeks needs to be implemented and compared to the traditional way.

References

- [1] S. BRENNER AND L. SCOTT, *The mathematical theory of finite element methods*, Springer Verlag, 1994.
- [2] K. ERIKSSON, D. ESTEP, P. HANSBO, AND C. JOHNSON, *Computational Differential Equations*, Studentlitteratur, 1996.
- [3] D. ESTEP, M. LARSON, AND R. WILLIAMS, *Estimating the error of numerical solutions of systems of reaction-diffusion equations*, MEMOIRS of the American Mathematical Society, 146 (2000).
- [4] G. FOUFAS AND M. G. LARSON, *Valuing european, barrier, and lookback options using the finite element method and duality techniques*. 2007.
- [5] ———, *Valuing fixed strike lookback options using the finite element method and duality techniques*. 2007.
- [6] ———, *Valuing asian options using the finite element method and duality techniques*, to appear in JCAM, (2008).
- [7] T. MOAN, *On the local distribution of errors by finite element approximations. in: Theory and practice in finite element structural analysis*, in Tokyo Seminar on Finite Element Analysis, 1973, pp. 43–60.
- [8] J. TOPPER, *Financial Engineering with finite elements*, Wiley Finance, 2005.
- [9] P. WILMOTT, *Paul Wilmott on quantitative finance.*, Wiley, 2000.
- [10] P. WILMOTT, J. DEWYNNE, AND S. HOWISON, *Option pricing*, Oxford Financial Press, 1993.
- [11] C. ZIENKIEWICZ, *The birth of the finite element method and of computational mechanics*, Internat. J. Numer. Methods Engrg., 60 (2004), pp. 3–10.

Paper V

A Posteriori Error Analysis of Weighted POD

Georgios Foufas and Mats G. Larson

April 14, 2008

Abstract

In this paper we develop an a posteriori error analysis for different model reduction techniques, such as the POD method and extensions of it. We also present a new model reduction technique, the Weighted POD method. Model reduction is the problem of obtaining a lower-dimensional approximation to a high-dimensional dynamical system. Here we use the finite element method and adopt different SVD based model reduction techniques used in fluid and solid dynamics, which enables us to reduce the size of the problem, which radically improves the performance. The a posteriori error estimates are derived using duality techniques.

1 Introduction

Model reduction is the problem of obtaining a lower-dimensional approximation to a high-dimensional dynamical system. There are two main sets of methods, singular value decomposition (SVD) based methods, and moment-matching methods. Moment matching methods have no global error bounds, and do not automatically preserve stability, whereas SVD based methods have error bounds and preserve stability. For a good survey of model reduction methods, see for example [2].

Here we use the finite element (FE) method and adopt different SVD based model reduction techniques used in fluid and solid dynamics, which enables us to reduce the size of the problem, which radically improves the performance. The standard FE basis is in some sense non-optimal, the question is what to use instead. As so elegantly described by [15], “The principal idea of dimensional model reduction is to find a small number of generalized co-ordinates in which to express the dynamics, ideally with some bounds on the truncation error”. In the context of FE models this can be realized by using several linear combinations of the FE basis functions

(modes or generalized coordinates) instead of the individual basis functions. Many different generalized coordinates functions have been proposed, see for example the solid dynamic articles [1], [19], [18], [24], [11], [10], [4], and [5], listed in chronological order.

In particular we use the so called proper orthogonal decomposition method, also known as POD, and extensions thereof, the so called balanced truncation method, [21], and the new Weighted POD method which is our extension of the POD method, presented in this paper. POD is closely related to the use of *empirical eigenvector* as a set of generalized coordinates. POD, also known as principal component analysis, or the Karhunen-Loève expansion, has been used a long time for developing low dimensional models in fluid dynamics, see for example Lumley 1970, [16], Sirovich 1987, [23], or Holmes *et al.* 1996, [9].

The truncation error introduced by using a small number of modes needs to be investigated. Kline, [13], has given some insight to the linear vibration problems. A posteriori error estimates of linear vibration problems vibration are provided by Cabos, [3]. Joo and Wilson described an application of the Ritz vectors in finite element mesh adaptation for dynamic problems, [12]. The understanding of the truncation error for nonlinear problems is rather limited compared to the linear case.

For control problems involving general sets of ordinary differential equations, some related error measures have been developed. In the linear case these error bounds were derived by Glover, [8], and Enns, [6], and for nonlinear systems error bounds were derived by Wood *et al.*, [25], and Scherpen, [22].

Recent results on model reduction of finite elements methods can be found in [15], the same authors have also developed model reduction for general Lagrangian systems in [14]. In [7] Fofas and Larson apply the POD method and extensions of it to option pricing problems.

Outline: In Section 2 we present the model problem and derive an a posteriori error estimation, and also study the special case of simple eigenfunction expansion. In Section 3 we present the so called weighted POD method, an extension of the POD method. We also discuss the different special cases POD and balanced truncation. Finally, in Section 4 we state some conclusions.

2 Model Reduction

The idea is, given a set of data that lies in the vector space \mathcal{W} , to find a subspace \mathcal{W}_r of fixed dimension r such that the error in the projection onto the subspace is minimized. The subspace can be determined in a number of

ways, for example by using the weighted POD method described later on.

2.1 Model Problem

We choose to analyze the following simple model problem to demonstrate the technique and show the basic ideas, find $u = u(x, t)$ such that

$$\begin{cases} \dot{u}(x, t) - \mathcal{L}u(x, t) = f, & x \in \Omega, t \in [0, T], \\ u(x, 0) = u_0, \\ \nabla u|_{\partial\Omega} = 0, \end{cases} \quad (2.1)$$

where $\Omega \subset \mathbb{R}$, $\partial\Omega$ is the boundary of Ω , and \mathcal{L} is a simple operator, for example $\mathcal{L} = \Delta$. Multiplying this equation, with $\mathcal{L} = \Delta$, by the test function $v \in \mathcal{W}$ and integrating on $\Omega \times [0, T]$ we obtain

$$\int_0^T ((\dot{u}, v) - (\Delta u, v)) dt = \int_0^T (f, v) dt, \quad (2.2)$$

where we use the notation $(u, v) = \int_{\Omega} uv ds$. Using integration by parts we get

$$\int_0^T ((\dot{u}, v) + (\nabla u, \nabla v)) dt = \int_0^T (f, v) dt. \quad (2.3)$$

2.2 A Posteriori Error Estimation

Let U denote the approximate solution to the model problem (2.3) calculated as usual with the complete basis, and let U^* denote the approximate solution calculated with the model reduction technique. The error can then be splitted into two parts, $e = (u - U) + (U - U^*)$, where the first part is the usual discretization error, and the second part is the error made by using the model reduction technique. To represent the error in a linear functional, (e, ψ) , we introduce the continuous dual problem for the model problem (2.1). Find $\phi \in \mathcal{W}$ such that

$$\begin{cases} -\phi_t - \mathcal{L}^{-1}\phi = 0, \\ \phi(x, T) = \psi, \end{cases} \quad (2.4)$$

where \mathcal{L}^{-1} is the inverse operator of \mathcal{L} . For simplicity we here use $\mathcal{L} = \Delta$. Multiplying equation (2.4) with the error $e \in \mathcal{W}_r$ and integrating in space and time we get

$$\int_0^T ((-\phi_t, e) - (\Delta\phi, e)) dt = 0. \quad (2.5)$$

Integrating by parts and using the boundary conditions we get

$$\int_0^T (-(\phi(T), e(T)) + (\phi(0), e(0)) + (\phi, e_t) + (\nabla\phi, \nabla e)) dt = 0. \quad (2.6)$$

Using the conditions $u(0) - U(0) = 0$, $\phi(T) = \psi$ we have

$$(\psi, e(T)) = (\phi(0), U(0) - U^*(0)) + \int_0^T \left((\phi, e_t) + (\nabla\phi, \nabla e) \right) dt \quad (2.7)$$

Since $e = (u - U) + (U - U^*)$ and u solves equation (2.3) we get the error representation formula

$$\begin{aligned} (\psi, e(T)) &= (\phi(0), U(0) - U^*(0)) + \int_0^T (\phi, f) dt \\ &\quad - \int_0^T \left((\phi, U_t^*) + (\nabla\phi, \nabla U^*) \right) dt \\ &= (\phi(0), U(0) - U^*(0)) - \int_0^T (R(U^*), \phi) dt, \end{aligned} \quad (2.8)$$

where the last equality follows by intergration by parts and $R(U^*) = U^* - \Delta U^* - f$ is the residual.

Let $\pi : \mathcal{W} \rightarrow \mathcal{W}^{q-1}$ be the L_2 projection in time, and let P be a suitable interpolation operator into \mathcal{V}^p in space, and let P_r be a suitable projection operator onto the subspace W_r . Thus πP is an interpolation operator such that $\pi P\phi \in \mathcal{W}^{q-1}$. Then using Galerkin orthogonality, we can replace ϕ by $\phi - \pi P_r P\phi = (\phi - P\phi) + (P\phi - P_r P\phi) + (I - \pi)P_r P\phi$. Equation (2.8) can then be written as

$$\begin{aligned} (\psi, e(T)) &= (\phi(0), U(0) - U^*(0)) + \int_0^T (R(U^*), \phi - P\phi) dt \\ &\quad + \int_0^T (R(U^*), P\phi - P_r P\phi) dt \\ &\quad + \int_0^T (R(U^*), (I - \pi)P_r P\phi) dt. \end{aligned} \quad (2.9)$$

2.3 Eigenfunction Expansion

In the case when the operator \mathcal{L} is symmetric and $f = 0$, we can apply a simple form of model reduction, namely to use a truncated eigenvector expansion of the operator \mathcal{L} as basis. In this case the error representation formula is simplified.

Let ξ_j be the eigenvectors, and λ_j the eigenvalues, of the operator $\mathcal{L} = \Delta$, that is

$$\Delta \xi_j = \lambda_j \xi_j. \quad (2.10)$$

Note that we can write $U^* = \sum_{i=1}^M \alpha_i \xi_i$, where α_i are the coefficients determined by the numerical method, and M is the number of basis functions in the truncated basis. We can then write the residual $R(U^*)$ as

$R(U^*) = \sum_{i=1}^M (\dot{\alpha}_i \xi_i + \alpha_i A \xi_i) = \sum_{i=1}^M (\dot{\alpha}_i + \alpha_i \lambda_i) \xi_i$, where A is the matrix corresponding to the Laplace operator. Since $(P\phi - P_r P\phi) \notin \text{span}\{\xi_1, \dots, \xi_M\}$ we can rewrite the error representation formula (2.9) as

$$\begin{aligned} (\psi, e(T)) &= (\phi(0), U(0) - U^*(0)) + \int_0^T (R(U^*), \phi - P\phi) dt \\ &+ \int_0^T (R(U^*), (I - \pi)P_r P\phi) dt. \end{aligned} \quad (2.11)$$

3 Weighted POD

We here present the weighted POD method, an extension of the so called POD method. For a detailed analysis of the POD method we refer to Holmes et. al. [9] and the references therein.

Suppose we have set of scalar fields $\{U_k\}$, each being a function $U = U(x), x \in \Omega$. In the POD method one then assumes that each U belong to the linear, infinite-dimensional Hilbert space $L^2([0, 1])$, of square integrable functions with inner product

$$(f, g) = \int_0^1 f(x)g(x) dx. \quad (3.1)$$

In the weighted POD method we instead use the assumption that each U belong to a vector space \mathcal{W} , with a weighted inner product $(f, g)_{\mathcal{W}}$, with the only restriction that it must be a bilinear positive definite functional, possibly involving derivatives of the functions f and g and a weight-function. It may for example be the Hilbert space $H_1(\Omega)$, with inner product

$$(f, g)_{H_1} = \int_{\Omega} f(x)g(x) dx + \int_{\Omega} \nabla f(x)\nabla g(x) dx. \quad (3.2)$$

Following the exposition in Holmes et. al., we now want to find a basis $\{\varphi_j(x)\}_{j=1}^{\infty}$ for \mathcal{W} that is optimal for our data set in the sense that representations of the form

$$U_N(x) = \sum_{j=1}^N a_j \varphi_j(x) \quad (3.3)$$

describe typical members of $\{U_k\}$ better than any other representation of the same dimension in any other basis. Typical refers in this sense to an average operation. Denote the average operation $\langle \cdot \rangle$, which is assumed to commute with the spatial integration in the inner product. Mathematically the statement of optimality is that we should choose φ to maximize the

averaged projection of U onto φ , suitably normalized

$$\max_{\varphi \in \mathcal{W}} \frac{\langle |(U, \varphi)_{\mathcal{W}}|^2 \rangle}{\|\varphi\|_{\mathcal{W}}^2}, \quad (3.4)$$

where $|\cdot|$ denotes the modulus and $\|\cdot\|_{\mathcal{W}}$ is the norm defined by

$$\|f\|_{\mathcal{W}} = (f, f)_{\mathcal{W}}^{\frac{1}{2}} \quad (3.5)$$

The solution of (3.4) only gives one single function, whereas we are interested in finding a set of functions, which together provide the desired basis.

As described by Holmes et. al. we instead try to extremize $\langle |(U, \varphi)_{\mathcal{W}}|^2 \rangle$ subject to the condition $\|\varphi\|_{\mathcal{W}}^2 = 1$. The corresponding functional for this variational problem is

$$J[\varphi] = \langle |(U, \varphi)_{\mathcal{W}}|^2 \rangle - \lambda(\|\varphi\|_{\mathcal{W}}^2 - 1). \quad (3.6)$$

A necessary condition for extrema is that for all variations $\varphi + \delta\vartheta$, $\delta \in \mathbb{R}$

$$\frac{d}{d\delta} J[\varphi + \delta\vartheta]_{\delta=0} = 0. \quad (3.7)$$

Using (3.6) we have

$$\begin{aligned} \frac{d}{d\delta} J[\varphi + \delta\vartheta]_{\delta=0} & \quad (3.8) \\ &= \frac{d}{d\delta} [\langle (U, \varphi + \delta\vartheta)_{\mathcal{W}}(\varphi + \delta\vartheta, U)_{\mathcal{W}} \rangle - \lambda(\varphi + \delta\vartheta, \varphi + \delta\vartheta)_{\mathcal{W}}]_{\delta=0} \\ &= 2[\langle (U, \vartheta)_{\mathcal{W}}(\varphi, U)_{\mathcal{W}} \rangle - \lambda(\varphi, \vartheta)_{\mathcal{W}}] = 0. \end{aligned}$$

We thus receive the eigenvalue problem

$$\langle (U, \vartheta)_{\mathcal{W}}(\varphi, U)_{\mathcal{W}} \rangle = \lambda(\varphi, \vartheta)_{\mathcal{W}}. \quad (3.9)$$

The optimal basis is given by the eigenfunctions φ , called weighted POD modes. Discretising in space we expand φ and ϑ in a finite element basis, that is

$$\varphi = \sum_{j=1}^N \xi_j N_j(x), \quad (3.10)$$

and

$$\vartheta = \sum_{j=1}^N \mu_j N_j(x), \quad (3.11)$$

where N_j is the standard finite element basis function. Inserting into equation (3.9), we get an $N \times N$ eigenvalue problem

$$A\xi = \lambda M\xi, \quad (3.12)$$

where element (k, l) of the matrix A is

$$A_{kl} = \langle (U, N_k)_{\mathcal{W}}(N_l, U)_{\mathcal{W}} \rangle, \quad (3.13)$$

and element (k, l) of the matrix M is

$$M_{kl} = (N_l, N_k)_{\mathcal{W}}. \quad (3.14)$$

3.1 Method of Snapshots

Data, U , is often given as snapshots $U(t_j)$ at discrete times t_1, \dots, t_M . In the method of snapshots, developed by Sirovic [23], we use the snapshots $U(t_j)$ as our basis functions, that is

$$\varphi = \sum_{j=1}^M \xi_j U(t_j), \quad (3.15)$$

and

$$\vartheta = \sum_{j=1}^M \mu_j U(t_j). \quad (3.16)$$

Inserting into equation (3.9), we now get an $M \times M$ eigenvalue problem instead

$$A\xi = \lambda M\xi, \quad (3.17)$$

where the matrixes M and A have the elements

$$M_{kl} = (U(t_l), U(t_k))_{\mathcal{W}}, \quad (3.18)$$

and

$$A_{kl} = \frac{1}{M} \sum_{i=1}^M (U(t_i), U(t_k))_{\mathcal{W}}(U(t_l), U(t_i))_{\mathcal{W}}, \quad (3.19)$$

respectively. Note that we have approximated the time integral in (3.13) with an average sum. Discretising the snapshots $U(t_j)$ in the finite element basis

$$U(t_j) = \sum_{n=1}^N U_j^n N_n(x), \quad (3.20)$$

we see that the elements of the matrix M can be written as

$$M_{kl} = U_l^* W U_k, \quad (3.21)$$

where $U_j = [U_j^1 \dots U_j^N]$ and W is the matrix with elements

$$W_{kl} = (N_k, N_l)_W. \quad (3.22)$$

In the same way we write the elements of the matrix A as

$$A_{kl} = \frac{1}{M} \sum_{i=1}^M U_i^* W U_k U_l^* W U_i = \frac{1}{M} U_k^* W \sum_{i=1}^M (U_i U_i^*) W U_l, \quad (3.23)$$

where the last equality follows since W is symmetric. Collecting the snapshot vectors U_j as columns in a matrix X , that is

$$X = [U_1 \dots U_m], \quad (3.24)$$

we note that

$$\sum_{i=1}^M (U_i U_i^*) = X X^*, \quad (3.25)$$

and the matrixes M and A can be written as

$$M = X^* W X, \quad (3.26)$$

and that

$$A = X^* W X X^* W X. \quad (3.27)$$

The eigenvalue problem (3.12) may then be rewritten in matrixform as

$$X^* W X X^* W X \xi = \lambda X^* W X \xi, \quad (3.28)$$

which can be simplified to

$$X^* W X \xi = \lambda \xi. \quad (3.29)$$

Instead of solving equation (3.29) one can solve the corresponding singular value decomposition (SVD) problem

$$X^* W X \xi = C \Sigma D^* = [C_1 C_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} D_1^* \\ D_2^* \end{bmatrix} = C_1 \Sigma_1 D_1^*, \quad (3.30)$$

where $\Sigma_1 \in \mathbb{R}^{r \times r}$ is an invertible diagonal matrix containing the so called Hankel singular values, r is the rank of $X^* W X$, and $C_1^* C_1 = D_1^* D_1 = I_r$.

3.2 Basic POD

The original snapshot POD method is just a special case of the weighted POD method. If we choose the weight-function w equal to one, and choose \mathcal{W} to be the usual Hilbert space L^2 , we get the original snapshot POD method.

In this case the matrix W reduces to the usual finite element mass matrix

$$m_{kl} = (\varphi_k, \varphi_l), \quad (3.31)$$

and the eigenvalue problem (3.29) now reads

$$X^* m X \xi = \lambda \xi. \quad (3.32)$$

This is almost the original snapshot POD eigenvalue problem, studied by for example Rowley, [21]. The only difference is the lack of the mass matrix m in the basic POD method

$$X^* X \xi = \lambda \xi. \quad (3.33)$$

As noted in Rowley [21], POD modes can also be calculated by solving a SVD of the snapshot-matrix X instead of solving the eigenvalue problem (3.32). The SVD problem has better roundoff properties although it requires more computation. This is the technique we use later on in the examples.

3.3 Dual Information

By slightly extending the definition of the inner product, we can incorporate dual information in the choice of the base. Just as we collected the primal snapshots in a matrix X , we collect the snapshots of the dual data vectors ϕ_j as columns in a matrix Y , that is

$$Y = [\phi_1 \dots \phi_m]. \quad (3.34)$$

We note that

$$\sum_{i=1}^M (\phi_i \phi_i^*) = Y Y^*. \quad (3.35)$$

3.3.1 Balanced Truncation

Balanced truncation is a model reduction method based on dual information developed by Moore [17]. Rowley [20] developed a snapshot version of it where one solves the SVD of the matrix $Y^* X$

$$Y^* X = U \Sigma V^* = [U_1 U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} = U_1 \Sigma_1 V_1^* \quad (3.36)$$

where $\Sigma_1 \in \mathbb{R}^{r \times r}$ is invertible, r is the rank of Y^*X , and $U_1^*U_1 = V_1^*V_1 = I_r$.

We will show that by defining the inner product

$$(f, g)_{\mathcal{W}} = \int_{\Omega} f(x)YY^*g(x)dx, \quad (3.37)$$

we actually receive the balanced truncation method. Going through the calculations in Section 3.1 we see that we still get the same SVD problem to solve,

$$X^*WX\xi = C_1\Sigma D_1^*, \quad (3.38)$$

but with a different W ,

$$W = YY^*. \quad (3.39)$$

We note that according to equation (3.36)

$$(Y^*X)^* = X^*Y = (U_1\Sigma_1V_1^*)^* = V_1^*\Sigma_1U_1^*, \quad (3.40)$$

which gives that

$$X^*YY^*X = V_1^*\Sigma_1U_1^*U_1\Sigma_1V_1 = V_1^*\Sigma_1^2V_1, \quad (3.41)$$

since $U_1^*U_1 = I_r$. That is, we get the same problem as equation (3.38) but with a different scaling of the Hankel singular values.

4 Conclusions

We have presented an extension of the balanced truncation method or the closely related balanced POD method. The weighted POD method share the same benefits as the previous methods, but the thought is that it can be adapted by choice of inner product to different cases. This obviously needs more attention and study. We also provide an a posteriori error analysis which to our knowledge has not been presented before. As mentioned earlier the authors have applied the methods presented in this article to option pricing problems, see [7]. The methods are there tested and compared numerically on European and Asian options.

References

- [1] B. ALMROTH, P. STERN, AND F. BROGAN, *Automatic choice of global shape functions in structural analysis*, AIAA Journal, 16 (1978), pp. 525–528.
- [2] A. ANTOULAS, D. SORENSSEN, AND S. GUGERCIN, *A survey of model reduction methods for large-scale systems*, Contemp. Math., 280 (2001), pp. 193–219.
- [3] C. CABOS, *Error bound for dynamic responses in forced vibration problems*, SIAM Journal of Scientific Computing, 15 (1994), pp. 1–15.
- [4] A. CHAN AND K. HSIAO, *Nonlinear analysis using a reduced number of variables*, Computer Methods in Applied Mechanics and Engineering, (1985), pp. 899–913.
- [5] C. CHANG AND J. ENGBLOM, *Nonlinear dynamical response of impulsively loaded structures: a reduced basis approach*, AIAA Journal, 29 (1991), pp. 613–618.
- [6] D. ENNS, *Model reduction for control system design*, PhD thesis, Stanford University, 1984.
- [7] G. FOUFAS AND M. G. LARSON, *Model reduction in option pricing using weighted POD*. 2007.
- [8] K. GLOVER, *All optimal hankel-norm approximations of linear multi-variable systems and their l -infinity error bounds*, International Journal of Control, (1984), pp. 1115–1193.
- [9] P. HOLMES, J. LUMLEY, AND G. BERKOOZ, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press: Cambridge, 1996.
- [10] S. IDELSOHN AND A. CARDONA, *a load-dependent basis for reduced nonlinear structural dynamics*, Computer and Structures, 1-3 (1985), pp. 203–210.
- [11] ———, *A reduction method for nonlinear structural dynamics analysis*, Computer Methods in Applied Mechanics and Engineering, (1985), pp. 253–279.
- [12] K.-J. JOO AND E. WILSON, *An adaptive finite element technique for structural dynamic analysis*, Computers and Structures, 30 (1988).

- [13] K. KLINE, *Dynamic analysis using a reduced basis of exact models and Ritz vectors*, AIAA Journal, 24 (1986), pp. 2022–2029.
- [14] S. LALL, P. KRYSL, AND J. MARSDEN, *Structure-preserving model reduction of mechanical systems*. In preparation, 2000.
- [15] ———, *Dimensional model reduction in non-linear finite element dynamics of solids and structures*, International Journal for Numerical Methods in Engineering, (2001), pp. 479–504.
- [16] J. LUMLEY, *Stochastic Tools in Turbulence*, Academic Press, 1970.
- [17] B. MOORE, *Principial component analysis in linear systems: Controllability, observability, and model reduction*, IEEE Trans. Automat. Contr., 26 (1981).
- [18] A. NOOR, *Recent advances in reduction methods for nonlinear problems*, Computers and Structures, (1981), pp. 31–33.
- [19] A. NOOR AND J. PETERS, *Reduced basis technique for nonlinear analysis of structures*, AIAA Journal, 18 (1980), pp. 455–462.
- [20] C. ROWLEY, *Balanced model reduction for large linear systems using snapshots*, Submitted to the 43rd IEEE Conference on Decision and Control, (2004).
- [21] ———, *Model reduction for fluids, using balanced proper orthogonal decomposition*, Int. J. on bifurcation and Chaos, (2005).
- [22] J. SCHERPEN, *H-infinity balancing for nonlinear systems*, International Journal of Robust and Nonlinear Control, (1996), pp. 645–688.
- [23] L. SIROVICH, *Turbulence and the dynamics of coherent structures, parts i-iii*, Q. Appl. Math., XLV (1987), pp. 561–590.
- [24] E. WILSON, M. YUAN, AND J. DICKENS, *Dynamic analysis by direct superposition of Ritz vectors*, Earthquake Engineering and Structural Dynamics, (1982), pp. 813–821.
- [25] G. WOOD, P. GODDARD, AND K. GLOVER, *Approximation of linear parameter-varying systems*, In the Proceedings of the IEEE Conference on Decision and Control, 1996.

Paper VI

Model Reduction in Option Pricing using Weighted POD

Georgios Foufas and Mats G. Larson

April 14, 2008

Abstract

The main objective of this paper is to apply different model reduction techniques, such as the POD method and a newly developed extension of it, Weighted POD, to the problem of pricing exotic options. Model reduction is a method that seeks to construct a lower-dimensional approximation to a high-dimensional dynamical system. In this paper we use the finite element method and adopt SVD based model reduction techniques used in fluid and solid dynamics, which enables us to substantially reduce the size of the problem, leading to a radical improvement of the performance. The techniques are tested and compared on European and Asian options.

1 Introduction

The market of different types of derivative contracts has grown very fast in recent years. The importance of calculating prices in a fast and stable manner has become more and more eminent. Vanilla contracts allow for fast valuation, and sometimes even analytical formulas exist, but most exotic contracts and multidimensional contracts are more time consuming to value. Demands for fast solutions have led many to use different kinds of analytical formulas based on some limiting assumption. Other methods such as the finite difference method or the finite element (FE) method can also be made fast through different kinds of special implementations. The computationally most expensive phase for these two methods usually is the repeated solving of linear system of equations. Still, for high dimensional contracts (at least higher than four to five) there are no real alternatives to Monte Carlo, or Quasi Monte Carlo, simulations where much of the research is made today.

Model reduction is the problem of obtaining a lower-dimensional approximation to a high-dimensional dynamical system. There are two main sets

of methods, SVD based methods, and moment matching methods. Moment matching methods have been used in finance, but to our knowledge we are the first to try SVD based methods. SVD based methods have error bounds and preserve stability, but moment matching methods have no global error bounds, and do not automatically preserve stability. For a good survey of model reduction methods, see for example [2].

Here we use the FE method and adopt different singular value decomposition (SVD) based model reduction techniques used in fluid and solid dynamics, which enables us to reduce the size of the problem, which radically improves the performance. The standard FE basis is in some sense non-optimal, the question is what to use instead. As so elegantly described by [20], “The principal idea of dimensional model reduction is to find a small number of generalized co-ordinates in which to express the dynamics, ideally with some bounds on the truncation error”. In the context of FE models this can be realized by using several linear combinations of the FE basis functions (modes or generalized coordinates) instead of the individual basis functions. Many different generalized coordinates functions have been proposed, see for example the solid dynamic articles [1], [23], [22], [30], [16], [15], [4], and [5], listed in chronological order.

In particular we use the so called proper orthogonal decomposition method, also known as POD, and extensions thereof, the so called balanced truncation method [24], and the new Weighted POD method [10]. We apply these methods to the European and the Asian option pricing problem. For the European option there exists an analytical solution, but for the Asian option one has to rely on numerical techniques. The European option is included in the study only as a reference to the study of the Asian option. POD is closely related to the use of *empirical eigenvector* as a set of generalized coordinates. POD, also known as principal component analysis, or the Karhunen-Loève expansion, has been used a long time for developing low dimensional models in fluid dynamics, see for example Lumley 1970 [21], Sirovich 1987 [26], or Holmes *et al.* 1996 [14].

The truncation error introduced by using a small number of modes needs to be investigated. Kline [18], has given some insight to the linear vibration problems. A posteriori error estimates of linear vibration problems vibration are provided by Cabos [3]. Joo and Wilson described an application of the Ritz vectors in finite element mesh adaptation for dynamic problems, [17]. The understanding of the truncation error for nonlinear problems is rather limited compared to the linear case.

For control problems involving general sets of ordinary differential equations, some related error measures have been developed. In the linear case these error bounds were derived by Glover [13], and Enns [6], and for non-

linear systems they were derived by Wood *et al.* [31], and Scherpen [25].

Recent results on model reduction of finite elements methods can be found in [20], the same authors have also developed model reduction for general Lagrangian systems in [19].

Outline: In Section 2 we present the FE method and apply it to the European option problem. Then in Section 3 we apply the FE method to the Asian option problem. In Section 4 we present the different model reduction techniques, the POD method, the balanced truncation method, and the weighted POD method. Then in Section 5 we present some numerical examples, compare the different techniques, and present a sensitivity analysis. Finally, in Section 6 we state some conclusions.

2 An Adaptive Finite Element Method for the European Option

In this section we present the finite element method and apply it to the basic European option. For a more detailed analysis with a posteriori error estimates see [11].

2.1 Mathematical Background

We consider a continuous time trading economy on a bounded time horizon $[0, T]$. Probability is represented by the probability space $(\Omega_T, \mathcal{F}_T, P)$, where $\Omega_T = C[0, T]$, P is the corresponding Wiener measure, and $\mathcal{F}_T = \sigma(W(t); t \leq T)$. For simplicity we consider the standard Black-Scholes setting with a risk free asset and a dividend paying stock. Let $B(t)$ denote the price of a risk free asset at time t governed by the equation $B(t) = B(0)e^{rt}$, where r is the constant interest rate. Further we denote by $S(t)$ the value of an asset at time t . We assume the existence of an equivalent martingale measure Q , under which the discounted stock price $e^{-r(T-t)}S_t$ is an \mathcal{F}_t -martingale. The existence of the risk neutral measure Q assures that the market is free of arbitrage possibilities. Under Q the stock price follows the stochastic differential equation

$$dS(t) = (r - \nu)S(t)dt + S(t)\sigma dW(t), \quad (2.1)$$

where r is the constant interest rate, ν is the constant continuous dividend yield, σ is the volatility, and $W(t)$ is a Q Brownian motion process. Here σ is assumed to be a positive real number. The solution of (2.1) is

$$S(t) = S(0)e^{(r-\nu-\frac{\sigma^2}{2})t+\sigma W(t)}. \quad (2.2)$$

2.2 The Black-Scholes PDE

The value of the ordinary European option, $u(t, S(t)) = u(t, s)$, is given as the solution to Black-Scholes equation

$$u_t(t, s) + \frac{\sigma^2 s^2}{2} u_{ss}(t, s) + (r - \nu) s u_s(t, s) - r u(t, s) = 0, \quad t < T, \quad (2.3)$$

which is valid for $s = S(t) \in \mathbb{R}^+$. In order to construct a computational mesh we introduce a bounded interval $\Omega = [s_{min}, s_{max}] \subset \mathbb{R}^+$ with boundary $\partial\Omega = \{s_{min}, s_{max}\}$. We define the usual Hilbert space

$$H^1(\Omega) = \{v : \int_{\Omega} (|\nabla v|^2 + v^2) ds < \infty\}, \quad (2.4)$$

and let \mathcal{W} be the space of functions that are square integrable in time and belongs to $H^1(\Omega)$ in space, that is

$$\mathcal{W} = L^2([0, T], H^1(\Omega)). \quad (2.5)$$

We also use the notation $(u, v) = \int_{\Omega} u v ds$, and $(u, v)_{\partial\Omega} = u(s_{max})v(s_{max}) - u(s_{min})v(s_{min})$.

2.3 Variational Formulation

Multiplying the Black-Scholes equation (2.3) by the test function $v \in \mathcal{W}$ and integrating on $\Omega \times [0, T]$ we obtain

$$\int_0^T \left((u_t, v) + (r - \nu)(s u_s, v) + \frac{\sigma^2}{2} (s^2 u_{ss}, v) - r(u, v) \right) dt = 0. \quad (2.6)$$

Using integration by parts we get

$$(s^2 u_{ss}, v) = (s^2 u_s, v)_{\partial\Omega} - 2(s u_s, v) - (s^2 u_s, v_s). \quad (2.7)$$

Thus equation (2.6) becomes

$$\begin{aligned} \int_0^T \left((u_t, v) + (r - \nu - \sigma^2)(s u_s, v) \right. \\ \left. - \frac{\sigma^2}{2} (s^2 u_s, v_s) + \frac{\sigma^2}{2} (s^2 u_s, v)_{\partial\Omega} - r(u, v) \right) dt = 0. \end{aligned} \quad (2.8)$$

The boundary conditions for the European call option are $u(t, 0) = 0$ and $u(t, s) \sim s e^{-\nu(T-t)}$ as $s \rightarrow \infty$, and for the corresponding put $u(t, 0) = K e^{-r(T-t)}$ and $u(t, s) \sim 0$ as $s \rightarrow \infty$, see for example Wilmott, [29]. For simplicity of implementation we use the artificial boundary condition $u_{ss} = 0$

on $\partial\Omega$ for both the put and the call instead. This boundary condition works for all contracts if the payoff is at most linear in the underlying (see [29]) and does not affect the accuracy of the solution. Using equation (2.3) we can rewrite the boundary condition as

$$u_s = \frac{r}{s(r-\nu)}u - \frac{1}{s(r-\nu)}u_t, \quad (2.9)$$

and enforce it weakly by inserting identity (2.9) into equation (2.8). We thus want to solve the problem: find $u \in \mathcal{W}$ such that

$$\begin{cases} \int_0^T (m(u_t, v) + a(u, v)) dt = 0, \\ u(T, s) = \max(s - K, 0), \end{cases} \quad (2.10)$$

for every $v \in \mathcal{W}$, where

$$m(u_t, v) = (u_t, v) - \frac{\sigma^2}{2(r-\nu)}(su_t, v)_{\partial\Omega}, \quad (2.11)$$

and

$$\begin{aligned} a(u, v) &= (r - \nu - \sigma^2)(su_s, v) - \frac{\sigma^2}{2}(s^2u_s, v_s) \\ &\quad + \frac{\sigma^2 r}{2(r-\nu)}(su, v)_{\partial\Omega} - r(u, v). \end{aligned} \quad (2.12)$$

2.4 Finite Element Approximation

The finite element method is based on solution of the variational problem (2.10) with \mathcal{W} replaced by a finite dimensional function space of piecewise polynomials in space and time. For background on the finite element method see for instance [7].

We now partition $[0, T]$ as $0 = t_0 < t_1 < t_2 < \dots < t_N = T$, denoting each time interval by $I_n = (t_{n-1}, t_n)$ and each time step by $k_n = t_n - t_{n-1}$. Similarly we partition Ω as $s_{min} = s_0 < s_1 < s_2 < \dots < s_J = s_{max}$, denoting each spatial interval by $\kappa_j = [s_{j-1}, s_j)$ and the length of each interval by $h_j = s_j - s_{j-1}$.

In space, we let $\mathcal{V}^p \subset H^1(\Omega)$ denote the space of piecewise continuous functions of order p . On each space-time slab $S_n = I_n \times \Omega$, we define

$$\mathcal{W}_n^q = \{w(t, s) : w(t, s) = \sum_{j=0}^q t^j v_j(s), v_j \in \mathcal{V}^p, (t, s) \in S_n\}. \quad (2.13)$$

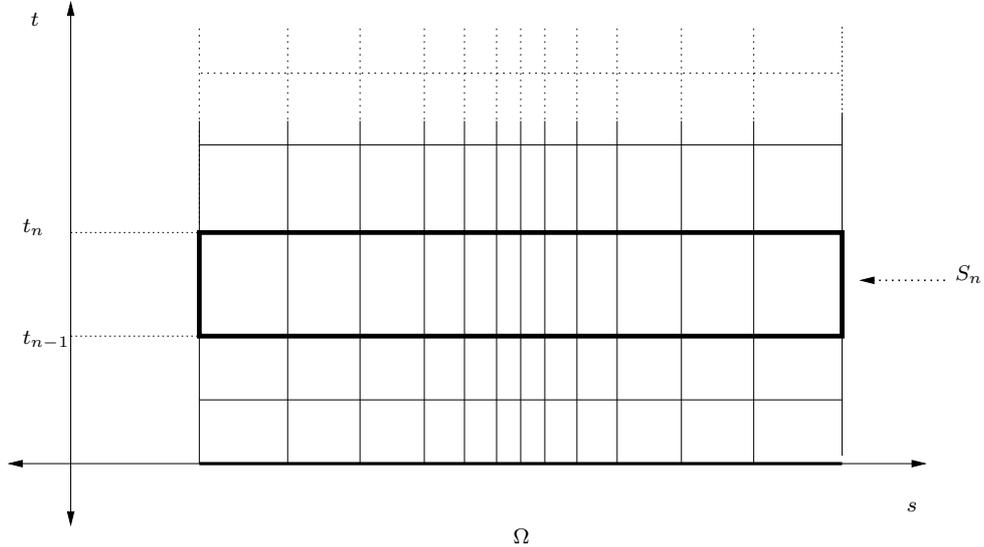


Figure 1: *Space-time discretization.*

Let $\mathcal{W}^q \subset \mathcal{W}$ denote the space of functions defined on $[0, T] \times \Omega$ such that $v|_{S_n} \in \mathcal{W}_n^q$ for $1 \leq n \leq N$. For simplicity, we only give details for the continuous Galerkin method cG(p)-cG(q), (see e.g. [7] or [8]) which is defined by the following discrete version of equation (2.10). Find $U \in \mathcal{W}^q$ such that for $1 \leq n \leq N$

$$\begin{cases} \int_{I_n} (m(U_t, v) + a(U, v)) dt = 0 & \text{for all } v \in \mathcal{W}_n^{q-1}, \\ U^-(t_n) = U^+(t_n), & n = N-1, \dots, 1, \\ U^-(t_N) = u_T, \end{cases} \quad (2.14)$$

where $U^\pm(t_n) = \lim_{\epsilon \rightarrow 0, \epsilon > 0} U(t_n \pm \epsilon)$. In the cG(1) method the approximation U of u is continuous piecewise linear in time and space, while the test functions v are continuous linear in space and piecewise constant in time. It is also possible to use a discontinuous method in time, we refer to [7], for details on the resulting discontinuous Galerkin method, cG(p)-dG(q).

2.5 Matrix Equations

We now derive the matrix equations for the case $p = q = 1$. Using the notation $U_n = U(t_n)$ and computing the time integral in equation (2.14)

yields the scheme: for $1 \leq n \leq N$

$$m(U_n - U_{n-1}, v) + k_n a\left(\frac{U_n + U_{n-1}}{2}, v\right) = 0 \quad \text{for all } v \in \mathcal{W}_n^0, \quad (2.15)$$

which is the classical Crank-Nicolson method.

Let $\{\varphi_j\}_{j=0}^J$ be the standard nodal basis of \mathcal{P}_1 (see Figure 2). Then $U_n \in \mathcal{P}_1$ can be written as

$$U_n(s) = \sum_{j=0}^J \xi_{nj} \varphi_j(s), \quad 1 \leq n \leq N, \quad (2.16)$$

and the test function v can be written as

$$v(s) = \sum_{i=0}^J \gamma_{ni} \varphi_i(s), \quad 1 \leq n \leq N, \quad (2.17)$$

for reals $\xi_{n0}, \dots, \xi_{nJ}, \gamma_{n0}, \dots, \gamma_{nJ}$.

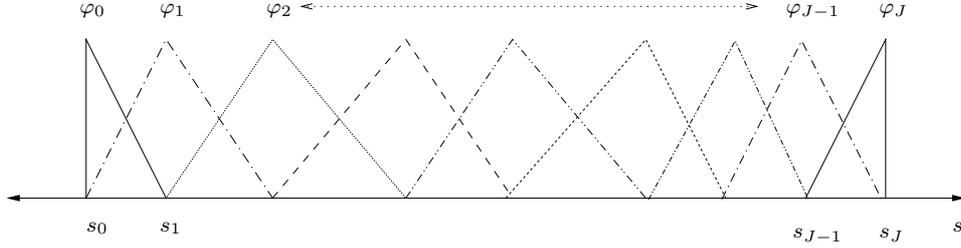


Figure 2: *The hat-functions φ in the cG(1) method.*

Let now ξ_n be the vector of all $\xi_{n,j}$, $j = 0, 1, \dots, J$. If the expressions above for U and v are inserted into equation (2.15) we receive the matrix equation

$$(\xi_n - \xi_{n-1})M + (\xi_n + \xi_{n-1})\frac{k_n A}{2} = 0, \quad 1 \leq n \leq N, \quad (2.18)$$

where

$$M = (\varphi_j, \varphi_i) - \frac{\sigma^2}{2(r - \nu)} (s\varphi_j, \varphi_i)_{\partial\Omega}, \quad 0 \leq i, j \leq J, \quad (2.19)$$

and

$$A = (r - \nu - \sigma^2)A_0 - \frac{\sigma^2}{2}A_1 - rA_2, \quad (2.20)$$

where

$$\begin{aligned} A_0 &= (s\varphi_{j,s}, \varphi_i), \quad A_1 = (s^2\varphi_{j,s}, \varphi_{i,s}) - \frac{r}{r-\nu}(s\varphi_j, \varphi_i)_{\partial\Omega}, \\ A_2 &= (\varphi_j, \varphi_i), \quad 0 \leq i, j \leq J. \end{aligned} \quad (2.21)$$

Rearranging the terms in equation (2.18) we get the matrix equation we need to solve successively backwards in time in order to obtain U_0 given U_N

$$\xi_{n-1} \left(M - \frac{k_n A}{2} \right) = \xi_n \left(M + \frac{k_n A}{2} \right), \quad 1 \leq n \leq N. \quad (2.22)$$

3 Asian Options

Here we give a short presentation of the pricing of the Asian option using the finite element method presented in the previous section. In this paper we use a method developed by Večer [28]. For a more thorough study including adaptivity and a posteriori error estimation see [12] or [9].

3.1 A Pricing Partial Differential Equation

The price of the Asian option, $V(t, S_t, X_t)$, can be represented as

$$V(t, S_t, X_t) = e^{-r(T-t)} E [V(T, S_T, X_T) | \mathcal{F}_t], \quad (3.1)$$

which is the usual expression for the value of a derivate. It can be shown (see [9]) that (3.1) is the solution to the following partial differential equation

$$-rv + V_t + rsV_s + qrsV_x + \frac{\sigma^2 s^2}{2} (V_{ss} + 2qV_{sx} + q^2V_{xx}) = 0. \quad (3.2)$$

We can use the change of variable

$$Z_t = \frac{X_t}{S_t}, \quad (3.3)$$

to reduce the dimensionality of (3.2) (see [28] or [9])

$$\begin{cases} \frac{\partial \bar{u}}{\partial t} + r(q_t - z) \frac{\partial \bar{u}}{\partial z} + \frac{\sigma^2}{2} (q_t - z)^2 \frac{\partial^2 \bar{u}}{\partial z^2} = 0, \\ \bar{u}(T, z) = z^+, \end{cases} \quad (3.4)$$

where $q_t = \mu([t, T])$. The price of the Asian option is then given in terms of \bar{u} by the equation

$$V(t, S_t, X_t) = S_t \bar{u}\left(t, \frac{X_t}{S_t}\right). \quad (3.5)$$

3.2 Variational Formulation

So far we have studied the pricing PDE for Asian options valid for $z \in \mathbb{R}$, but in order to construct a computational mesh we introduce a bounded interval $\Omega = [z_0, z_J] \subset \mathbb{R}^+$ with boundary $\partial\Omega = \{z_0, z_J\}$. Let H_1 and \mathcal{W} be defined just as in the European case. We denote by u the solution to (3.4) on Ω subject to the Dirichlet boundary conditions $u(t, z_0) = 0$ and $u(t, z_J) = z_J$ on $\partial\Omega$. We also use the notation $(u, v)_\Omega = \int_\Omega uv dz$, and $(u, v)_{\partial\Omega} = u(z_J)v(z_J) - u(z_0)v(z_0)$. Multiplying equation (3.4) by the test function $\{v \in \mathcal{W} : v = 0 \text{ on } \partial\Omega\}$ and integrating on $\Omega \times [0, T]$ we obtain

$$\int_0^T \left((u_t, v)_\Omega + r((q-z)u_z, v)_\Omega + \frac{\sigma^2}{2} ((q-z)^2 u_{zz}, v)_\Omega \right) dt = 0. \quad (3.6)$$

Using integration by parts we get

$$\begin{aligned} ((q-z)^2 u_{zz}, v)_\Omega &= ((q-z)^2 u_z, v)_{\partial\Omega} + 2((q-z)u_z, v)_\Omega \\ &\quad - ((q-z)^2 u_z, v_z)_\Omega. \end{aligned} \quad (3.7)$$

Thus equation (3.6) becomes

$$\int_0^T \left((u_t, v)_\Omega + (r + \sigma^2)((q-z)u_z, v)_\Omega - \frac{\sigma^2}{2} ((q-z)^2 u_z, v_z)_\Omega \right) dt = 0. \quad (3.8)$$

since $v = 0$ on $\partial\Omega$. Introducing the Dirichlet boundary conditions $u(t, z_0) = 0$ and $u(t, z_J) = z_J$ on $\partial\Omega$ (which is also used by Večeř, [28]) we get the following problem: find $u \in \mathcal{W}$ such that

$$\begin{cases} \int_0^T ((u_t, v)_\Omega + a_\Omega(u, v)) dt = 0, \\ u(T, z) = z^+, \\ u(t, z_0) = 0, \quad u(t, z_J) = z_J, \end{cases} \quad (3.9)$$

for every $\{v \in \mathcal{W} : v = 0 \text{ on } \partial\Omega\}$, where

$$a_\Omega(u, v) = (r + \sigma^2)((q-z)u_z, v)_\Omega - \frac{\sigma^2}{2} ((q-z)^2 u_z, v_z)_\Omega. \quad (3.10)$$

3.3 Finite Element Approximation

Applying the same finite element method as in the case of the European option we get the following FE problem: find $U \in \mathcal{W}^q$ such that for $1 \leq n \leq N$

$$\begin{cases} \int_{I_n} ((U_t, v)_\Omega + a_\Omega(U, v)) dt = 0 \quad \text{for all } \{v \in W_n^0 : v = 0 \text{ on } \partial\Omega\} \\ U^-(t_n) = U^+(t_n), \quad n = N-1, \dots, 1 \\ U^-(t_N) = u_T \\ U(t_n, z_0) = 0, \quad U(t_n, z_J) = z_J, \quad n = N-1, \dots, 1, \end{cases} \quad (3.11)$$

where a_Ω is given by (3.10), and $U^\pm(t_n) = \lim_{\epsilon \rightarrow 0, \epsilon > 0} U(t_n \pm \epsilon)$.

4 Model Reduction Methods

4.1 POD

The idea is, given a set of data that lies in the vector space \mathcal{W} , to find a subspace \mathcal{W}_r of fixed dimension r such that the error in the projection onto the subspace is minimized. Suppose we have set of data $x(t) = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n$, with $0 \leq t \leq T$, where each scalar field $\{x_k\}$ is a function $x_k = x_k(z), z \in \Omega$. In the POD method one then assumes that each x_k belong to the linear, infinite-dimensional Hilbert space $L^2(\Omega)$, of square integrable functions with inner product

$$(f, g) = \int_{\Omega} f(z)g(z) dz. \quad (4.1)$$

Following the exposition in Rowley [24] we seek a projection $P_r : \mathbb{R}^n \rightarrow \mathbb{R}^n$ of fixed rank r , that minimizes the total error

$$\int_0^T \|x(t) - P_r x(t)\|^2 dt. \quad (4.2)$$

Now introduce the $n \times n$ matrix

$$R = \int_0^T x(t)x(t)^* dt, \quad (4.3)$$

where x^* denotes the transpose of x , and calculate the eigenvalues and eigenvectors of R given by

$$R\eta_k = \lambda_k \eta_k, \quad \lambda_1 \geq \dots \geq \lambda_n \geq 0. \quad (4.4)$$

Since the matrix R is symmetric, positive semidefinite, all the eigenvalues λ_k are real and non-negative, the eigenvectors η_k may be chosen orthonormal. The main result of POD is that the optimal subspace of dimension r is spanned by $\{\eta_1, \eta_2, \dots, \eta_r\}$, and the optimal projection P_r is given by

$$P_r = \sum_{k=1}^r \eta_k \eta_k^*. \quad (4.5)$$

The vectors η_k is then used as the new basis and are called POD modes.

These POD modes can then be used to form reduced order methods by applying Galerkin projection. Writing the dynamics of our system as

$$\dot{x} = f(x(t)), \quad (4.6)$$

we define a new variable $x_r(t) \in \text{span}\{\eta_1, \dots, \eta_r\}$ by $\dot{x}_r(t) = P_r f(x_r(t))$. Let now

$$x_r(t) = \sum_{j=1}^r a_j(t) \eta_j. \quad (4.7)$$

Substituting this into equation (4.6) and multiplying by η_k^* we obtain

$$\dot{a}_k(t) = \eta_k^* f(x_r(t)), \quad k = 1, \dots, r, \quad (4.8)$$

a set of ODEs that describe the dynamics of $x_r(t)$.

4.2 Method of Snapshots for POD

In our case data is typically given at discrete times, even though we now the solution everywhere in time. According to Sirovich, [26], one can transform the $n \times n$ eigenvalue problem (4.4) into an $m \times m$ eigenvalue problem by exchanging the integral in (4.3) with a sum, using quadrature

$$R = \sum_{j=1}^m x(t_j) x(t_j)^* \delta_j, \quad (4.9)$$

where δ_j are the quadrature weights. Assembling the data into an $n \times m$ matrix

$$X = [x(t_1) \sqrt{\delta_1} \dots x(t_m) \sqrt{\delta_m}] \quad (4.10)$$

we can write the sum (4.9) as $R = XX^*$. One then solves the eigenvalue problem

$$X^* X u_k = \lambda_k u_k, \quad u_k \in \mathbb{R}^m. \quad (4.11)$$

For a more detailed derivation of this equation we refer to the excellent book [14]. The eigenvectors u_k may be chosen to be orthonormal, and the POD modes are given by $\varphi_k = X u_k / \sqrt{\lambda_k}$. The $m \times m$ eigenvalue problem (4.11) is more efficient than the $n \times n$ eigenvalue problem (4.4) when the number of snapshots m is smaller than the number of states n . As noted by Rowley [24], the POD modes are optimal at approximating a given data set, but they are not necessarily the best modes for describing the dynamics that generated a particular dataset, since low-energy features may be critically important to the dynamics. Sometimes, adding more POD modes can actually make dynamical models worse, see [27]. This is part of the motivation behind balanced truncation, described in the next section. For a more detailed analysis of POD and balanced truncation we refer to Rowley and the references therein.

As noted in Rowley, POD modes can also be calculated by solving a singular value decomposition (SVD) of the snapshot matrix X instead of

solving the eigenvalue problem (4.11). The SVD problem has better roundoff properties although it requires more computation. This is the technique we use later on in the examples.

4.3 Balanced Truncation

Here we give a brief presentation of the balanced truncation method. A more detailed analysis and an extension of the method (balanced POD) suitable for large systems can be found in Rowley [24].

Balanced truncation may be viewed as POD with respect to a particular inner product, or as biorthogonal decomposition, as noted by Rowley. Balanced truncation uses dual information in the choice of the basis. Just as we collected the primal snapshots in a matrix X , we collect the snapshots of the dual data vectors ϕ_j as columns in a matrix Y , that is

$$Y = [\phi_1 \dots \phi_m]. \quad (4.12)$$

The balancing modes are then computed by forming SVD of the matrix Y^*X

$$Y^*X = U\Sigma V^* = [U_1 U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} = U_1 \Sigma_1 V_1^* \quad (4.13)$$

where $\Sigma_1 \in \mathbb{R}^{r \times r}$ is invertible, r is the rank of Y^*X , and $U_1^* U_1 = V_1^* V_1 = I_r$. The matrix Σ_1 contains the so called Hankel singular values and the columns of T_1 form the balancing transformation, where

$$T_1 = X V_1 \Sigma_1^{-1/2}. \quad (4.14)$$

4.4 Weighted POD

Weighted POD is an extension of POD and balanced truncation. In [10] Foufas and Larson and develop the weighted POD method and present an a posteriori error analysis for the POD method and the weighted POD method. Here we just state the method and later on we test it on option valuation problems.

In the weighted POD method we use the assumption that each U belong to a vector space \mathcal{W} , with a weighted inner product $(f, g)_{\mathcal{W}}$, with the only restriction that it must be a bilinear positive definite functional, possibly involving derivatives of the functions f and g and a weight function. It may for example be the Hilbert space $H_1(\Omega)$, with inner product

$$(f, g)_{H_1} = \int_{\Omega} f(x)g(x) dx + \int_{\Omega} \nabla f(x)\nabla g(x) dx. \quad (4.15)$$

Using the same technique as Holmes et. al. [14] it is shown in [10] that the weighted POD problem reduces to solving the following SVD

$$X^*WX\xi = C\Sigma D^* = [C_1C_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} D_1^* \\ D_2^* \end{bmatrix} = C_1\Sigma_1D_1^*, \quad (4.16)$$

where W is the matrix with elements

$$W_{kl} = (\varphi_k, \varphi_l)_W, \quad (4.17)$$

$\Sigma_1 \in \mathbb{R}^{r \times r}$ is an invertible diagonal matrix containing the Hankel singular values, r is the rank of X^*WX , and $C_1^*C_1 = D_1^*D_1 = I_r$. Here φ are the standard finite element basis functions.

As mentioned earlier weighted POD is an extension of POD and balanced truncation. Actually POD and balanced truncation is received by choosing $W = m$ and $W = YY^*$ respectively, where m is the mass matrix.

5 Numerical Examples

Here we apply the different model reduction techniques presented above to the European option and the Asian option. As mentioned in the introduction, the European option is used as a reference only. We also present a sensitivity analysis.

5.1 The European Option

Figure 3 shows the finite element solution for an European option when $\sigma = 0.3$ and $r = 0.05$, $\nu = 0$, and $T = 1$. The solution is computed using the cG(2)-dG(1) method with 200 space and time points.

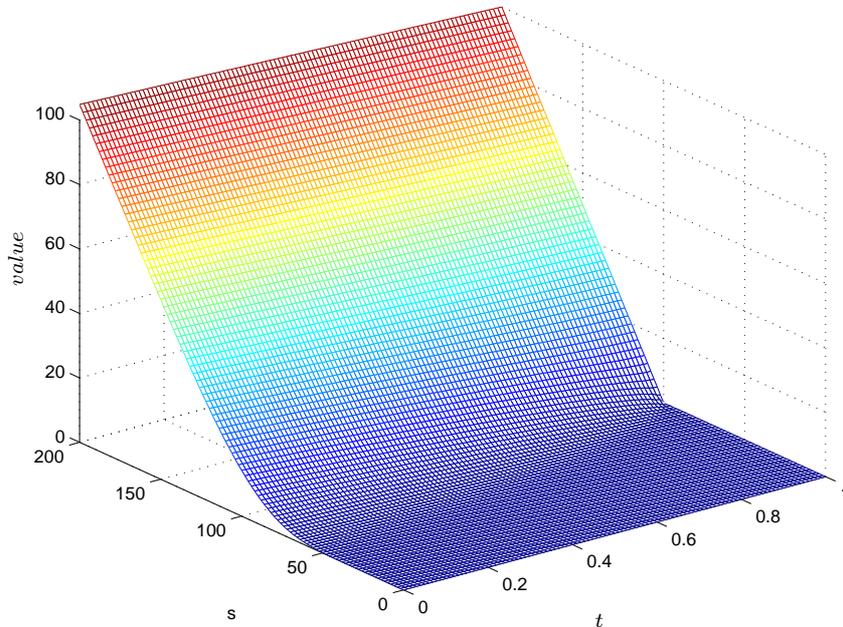


Figure 3: The price of an European option when $\sigma = 0.3$ and $r = 0.05$, $\nu = 0$, and $T = 1$. Computed using the $cG(2)$ - $dG(1)$ method with 200 space and time points.

In figure 4 we see the relative error of the reduced order methods when the finite element solution is regarded as the true solution, that is the relative error e_{rel} is defined as

$$e_{rel} = \frac{\|U - U_{red}\|}{\|U\|}, \quad (5.1)$$

where U is the ordinary finite element solution, U_{red} is the reduced order finite element solution, and $\|a\|$ denotes the l^2 norm of a . We see that all methods perform well and already after including only a few basis functions we have a very good solution. We also notice that the balanced truncation method and the weighted POD method (with the weight-matrix $W = MYY^*M^*$), which both takes dual information into account in the choice of the basis, performs slightly better when point wise errors are concerned and slightly worse when we look at the error in the entire domain, which makes perfect sense.

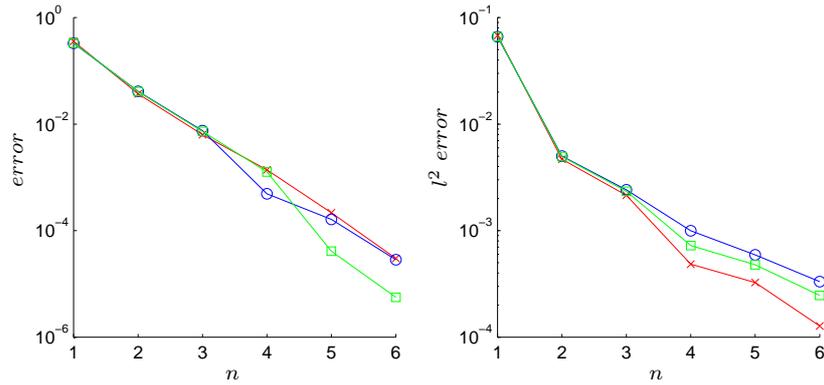


Figure 4: The relative error of an European option for the POD (\times) solution, the balanced truncation solution (\circ), and the weighted POD solution (\square) with the weight-matrix $W = MYY^*M^*$, when the finite element solution is the regarded as the true solution. On the left, the error in the point of interest, that is $S = 100$, and on the right, the l^2 error, where n is the number of basis functions. The finite element solutions were calculated using the cG(2)-dG(1) method with 100 space and time points, when $\sigma = 0.3$, $r = 0.05$, $\nu = 0$, and $T = 1$.

5.2 The Asian Option

Figure 5 shows the finite element solution for an Asian option when $\sigma = 0.3$ and $r = 0.05$, $\nu = 0$, and $T = 1$. The solution is computed using the cG(2)-dG(1) method with 100 space and time points.

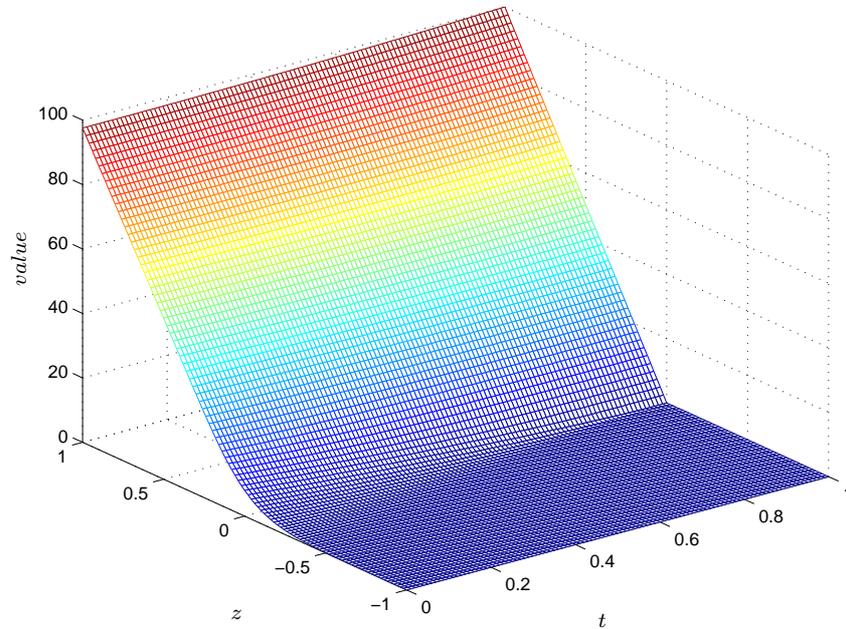


Figure 5: *The price of an Asian option when $\sigma = 0.3$ and $r = 0.05$, $\nu = 0$, and $T = 1$. Computed using the $cG(2)$ - $dG(1)$ method with 100 space and time points.*

In figure 6 we see the relative error e_{rel} , defined by equation (5.1), of the reduced order methods when the finite element solution is regarded as the true solution. Just as for the European option, we see that all methods perform well and already after including only a few basis functions we have a very good solution. The balanced truncation method and the weighted POD method (with the weight-matrix $W = MYY^*M^*$) performs slightly better when point wise errors are concerned and slightly worse when we look at the error in the entire domain.

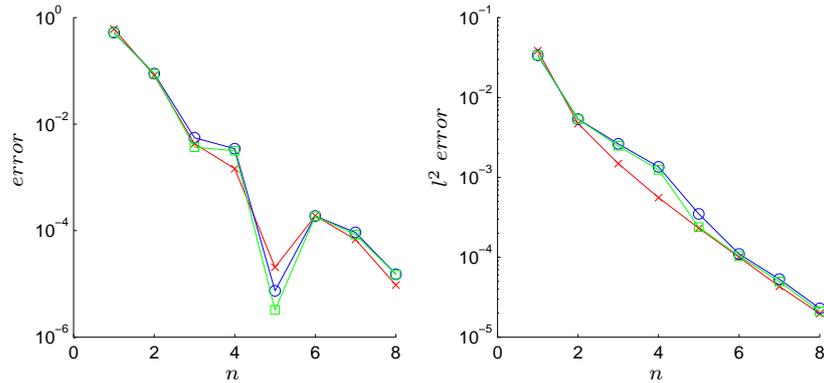


Figure 6: The relative error of an Asian option for the POD (\times) solution, and the balanced truncation solution (\circ), and the weighted POD solution (\square) with the weight-matrix $W = MYY^*M^*$, when the finite element solution is the regarded as the true solution. On the left, the error in the point of interest, that is $S = 100$, and on the right, the l^2 error, where n is the number of basis functions. The finite element solutions were calculated using the $cG(2)$ - $dG(1)$ method with 100 space and time points, when $\sigma = 0.3$, $r = 0.05$, $\nu = 0$, and $T = 1$.

5.3 Sensitivity Analysis

As pointed out previously the SVD based model reduction techniques seem to be working very well. When dealing with option pricing problems there is always a demand for performance. We do not want to calculate a new basis each time we price an option. The question is how stable the reduced basis is for changes in the parameters. If the basis functions are not so sensitive towards changes in the parameters then we do not need basis functions for that many different settings of the parameters. The strategy is to in advance calculate basis functions for a number of different combinations of

parameter settings. With the computer storage capacities existing today it is no problem at all to save a great number of basis functions.

In figures 7 and 8 we see how sensitive the balanced truncation method is for changes in the parameter σ for the case of the European option and the Asian option. That is if we calculate a base using a certain value for σ , and then use this base for calculating the price of the same option but for a slightly different σ , $\sigma*b$, where $b = 0.9, 0.95, 0.99, 1.0, 1.01, 1.05, 1.1$. The figures show the relative error between finite element solution and the balanced truncation solution calculated for different values of the parameter σ but with the same base.

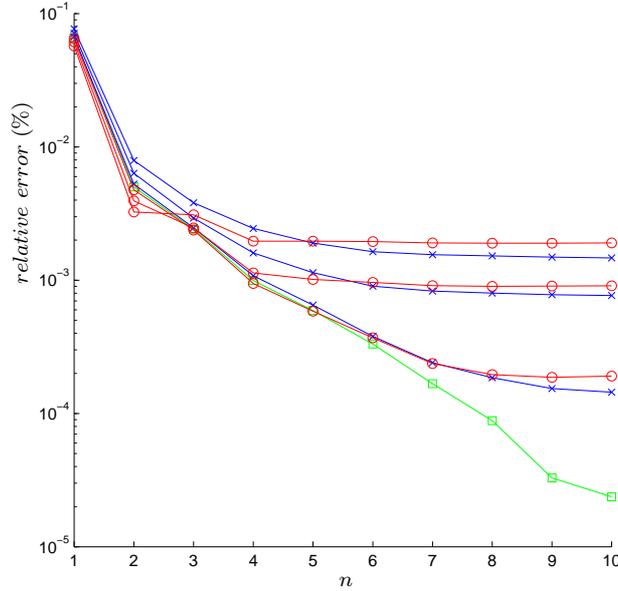


Figure 7: Relative errors for the European option between the finite element solution and the balanced truncation solution calculated for different values of the parameter σ but using the same modes, where n is the number of basis functions. The modes were calculated using $\sigma = 0.3$. The relative errors when $b = 0.9, 0.95, 0.99$ are plotted using (\times) at the data points, and the relative errors when $b = 1.1, 1.01, 1.05$ are plotted using (\circ) . As a reference the error when $b = 1.0$ is plotted using (\square) . The finite element solutions were calculated using the $cG(2)$ - $dG(1)$ method with 100 space and time points, when $\sigma = 0.3$, $r = 0.05$, $\nu = 0$, and $T = 1$.

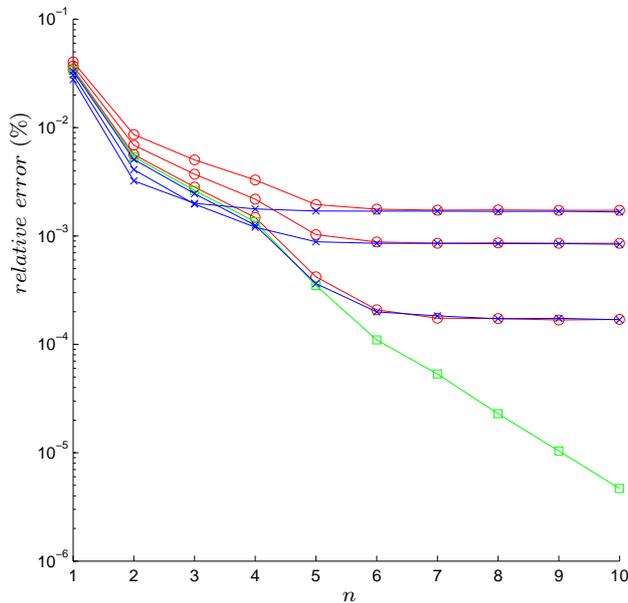


Figure 8: *Relative errors for the Asian option between the finite element solution and the balanced truncation solution calculated for different values of the parameter sigma but using the same modes, where n is the number of basis functions. The modes were calculate using $\sigma = 0.3$. The relative errors when $b = 0.9, 0.95, 0.99$ are plotted using (\times) at the data points, and the relative errors when $b = 1.1, 1.01, 1.05$ are plotted using (\circ) . As a reference the error when $b = 1.0$ is plotted using (\square) . The finite element solutions were calculated using the $cG(2)$ - $dG(1)$ method with 100 space and time points, when $\sigma = 0.3$, $r = 0.05$, $\nu = 0$, and $T = 1$.*

6 Conclusions

All methods work well on both European options and Asian options. Already after including just a few modes in the basis we receive very good accuracy. The balanced truncation method and the weighted POD method which both uses dual information performs slightly better when pointwise errors are concerned and slightly worse otherwise. For more complex problems one would expect even more from the dual methods. POD modes are very effective at describing a particular dataset, but they are not necessary the best modes for describing the dynamics that generate a particular dataset,

since sometimes low energy modes may be more dynamically important than high energy modes, see [27] or [24]. For the option pricing problems studied here one can draw the conclusion that all the methods perform well and could allow a significant speed up of the existing pricing procedures. For multidimensional problems one would expect that the method is even more efficient. Of course more testing needs to be done, but the first tests indeed look very promising. The sensitivity analysis indicates that the calculated modes are sensitive toward changes in the parameter σ but not that much that it makes the methods unusable.

References

- [1] B. ALMROTH, P. STERN, AND F. BROGAN, *Automatic choice of global shape functions in structural analysis*, AIAA Journal, 16 (1978), pp. 525–528.
- [2] A. ANTOULAS, D. SORENSEN, AND S. GUGERCIN, *A survey of model reduction methods for large-scale systems*, Contemp. Math., 280 (2001), pp. 193–219.
- [3] C. CABOS, *Error bound for dynamic responses in forced vibration problems*, SIAM Journal of Scientific Computing, 15 (1994), pp. 1–15.
- [4] A. CHAN AND K. HSIAO, *Nonlinear analysis using a reduced number of variables*, Computer Methods in Applied Mechanics and Engineering, (1985), pp. 899–913.
- [5] C. CHANG AND J. ENGBLOM, *Nonlinear dynamical response of impulsively loaded structures: a reduced basis approach*, AIAA Journal, 29 (1991), pp. 613–618.
- [6] D. ENNS, *Model reduction for control system design*, PhD thesis, Stanford University, 1984.
- [7] K. ERIKSSON, D. ESTEP, P. HANSBO, AND C. JOHNSON, *Computational Differential Equations*, Studentlitteratur, 1996.
- [8] D. ESTEP, M. LARSON, AND R. WILLIAMS, *Estimating the error of numerical solutions of systems of reaction-diffusion equations*, MEMOIRS of the American Mathematical Society, 146 (2000).
- [9] G. FOUFAS, *PDE-methods for Asian options*, Master’s thesis, Chalmers Univeristy of Technology, 2000.

- [10] G. FOUFAS AND M. G. LARSON, *A posteriori error analysis of weighted POD*. 2007.
- [11] ———, *Valuing european, barrier, and lookback options using the finite element method and duality techniques*. 2007.
- [12] ———, *Valuing asian options using the finite element method and duality techniques*, to appear in *JCAM*, (2008).
- [13] K. GLOVER, *All optimal hankel-norm approximations of linear multi-variable systems and their l-infinity error bounds*, *International Journal of Control*, (1984), pp. 1115–1193.
- [14] P. HOLMES, J. LUMLEY, AND G. BERKOOZ, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press: Cambridge, 1996.
- [15] S. IDELSOHN AND A. CARDONA, *a load-dependent basis for reduced nonlinear structural dynamics*, *Computer and Structures*, 1-3 (1985), pp. 203–210.
- [16] ———, *A reduction method for nonlinear structural dynamics analysis*, *Computer Methods in Applied Mechanics and Engineering*, (1985), pp. 253–279.
- [17] K.-J. JOO AND E. WILSON, *An adaptive finite element technique for structural dynamic analysis*, *Computers and Structures*, 30 (1988).
- [18] K. KLINE, *Dynamic analysis using a reduced basis of exact models and Ritz vectors*, *AIAA Journal*, 24 (1986), pp. 2022–2029.
- [19] S. LALL, P. KRYSL, AND J. MARSDEN, *Structure-preserving model reduction of mechanical systems*. In preparation, 2000.
- [20] ———, *Dimensional model reduction in non-linear finite element dynamics of solids and structures*, *International Journal for Numerical Methods in Engineering*, (2001), pp. 479–504.
- [21] J. LUMLEY, *Stochastic Tools in Turbulence*, Academic Press, 1970.
- [22] A. NOOR, *Recent advances in reduction methods for nonlinear problems*, *Computers and Structures*, (1981), pp. 31–33.
- [23] A. NOOR AND J. PETERS, *Reduced basis technique for nonlinear analysis of structures*, *AIAA Journal*, 18 (1980), pp. 455–462.

- [24] C. ROWLEY, *Model reduction for fluids, using balanced proper orthogonal decomposition*, Int. J. on bifurcation and Chaos, (2005).
- [25] J. SCHERPEN, *H-infinity balancing for nonlinear systems*, International Journal of Robust and Nonlinear Control, (1996), pp. 645–688.
- [26] L. SIROVICH, *Turbulence and the dynamics of coherent structures, parts i-iii*, Q. Appl. Math., XLV (1987), pp. 561–590.
- [27] T. SMITH, *Low-dimensional models of plane Couette flow using the proper orthogonal decomposition*, PhD thesis, Princeton University, 2003.
- [28] J. VEČEŘ, *A new PDE approach for pricing arithmetic average Asian options*, The Journal of Computational Finance, (2000).
- [29] P. WILMOTT, J. DEWYNNE, AND S. HOWISON, *Option pricing*, Oxford Financial Press, 1993.
- [30] E. WILSON, M. YUAN, AND J. DICKENS, *Dynamic analysis by direct superposition of Ritz vectors*, Earthquake Engineering and Structural Dynamics, (1982), pp. 813–821.
- [31] G. WOOD, P. GODDARD, AND K. GLOVER, *Approximation of linear parameter-varying systems*, In the Proceedings of the IEEE Conference on Decision and Control, 1996.

Paper VII

Option Manager: A Software Package for Calculating and Visualizing Exotic Option Prices and Greeks

Georgios Foufas*

April 14, 2008

Abstract

In this report we present a software project that gives the user the ability to easily calculate different exotic option prices and the corresponding Greeks in a graphical user interface (GUI). The software *Option Manager* is implemented in C++ with a GUI developed in *Matlab's Guide*. The program features the ability to show option prices and Greeks graphically as evolutions in time or as a space-time plot for a specific time. The valuation is done using the finite element method, and features dual techniques as well. The program is also equipped with the availability to calculate error estimations and show them graphically. This gives the user not just a tool for calculating prices and Greeks in an easy understandable way, but at the same time it aids to the understanding with visualization of the prices, Greeks, and error plots.

1 Introduction

The valuation of different types of derivative contracts is very important in modern financial theory and practice. Not only the option price itself is important to calculate, but also the sensitivity measures, or the so called Greeks, are important to be able to calculate in a fast and stable way, since they are used when hedging the options.

The performance demands on the valuation process is usually very high. Many different methods have been applied to attack these problems. The demand for performance have led some to use approximations that produce closed form expressions. Others rely on numerical methods such as binomial

*Research Assistant, Department of Mathematics, Chalmers University of Technology, S-412 96 Gothenburg , Sweden, *email*: fofas@math.chalmers.se

and trinomial tree methods. Tree methods are easy to understand and can be applied to many types of problems, at the same time they sometimes work less well and they lack error analysis. Another frequently used method is Finite Difference (FD) method. We use another method, the so called Finite Element (FE) method.

The FE method is widely used in other fields as a tool for finding approximate solutions to partial differential equations (PDE) as well as of integral equations. It was developed in the 1950's and 1960's by engineers, and was mainly used in structural mechanics, see e.g. [6] for an overview. The FE method also has a strong mathematical foundation in functional analysis, see [1]. The mathematical foundation provides the tools to derive analytical error estimates which can be used to improve the approximative solution. The FE method has several advantages compared to the FD method, for example, using the FE method one receives a solution in the entire domain, not only in isolated nodes as in FD codes. FE codes can also incorporate different kinds of boundary conditions in an easy way. Other important advantages of the FE technique are that it can easily deal with high curvature and irregular shapes of the computational domain. One of the most important advantages in practice is that the sensitivity measures, or the so called Greeks, can be calculated more exactly using the FE method. In finance it has not been used that frequently compared to other methods such as the FD method. Recently Topper [5] wrote an excellent book applying the FE method to different option pricing problems. Topper uses a commercial software whereas we have developed our own software, see [3], [2], or [4].

A good program should be flexible and easy to use. At the same time the performance demands set limitations to what kind of programming languages that could be used. Flexibility of the code is achieved by programming in the object oriented language C++. We believe that C++ is very suitable language for developing an object oriented fast code that easily can be renewed to add new features without rewriting the whole program. At the same time it is hard to write a graphical user interface (GUI) in C++ that should be able to plot option prices in two and three dimensions. Therefore we decided to write the valuation code in C++ and develop the GUI in *Matlab's* easy to use *Guide*, a software package where one can develop GUI's in an easy way with the advantage that all the excellent visualization features in *Matlab* are available. The different tools needed in a finite element software are implemented as classes, which makes them easy to renew and combine. These classes include grids, vector, matrices, solvers, error estimators and so on.

Outline: In Section 2 we give a specification to what the software can do and how it is organized. Then in Section 3 we explain how to use the

software and show some examples. Finally in Section 4 we provide some future plans for development of the software.

2 Specification and Organisation

Option Manager is able to calculate option prices and Greeks for different exotic options. At the moment these options consists of European options, barrier options, floating and fixed strike lookback options, and Asian options. The user can easily choose the type of option, different variants, and sampling frequency in drop-down menus. The parameters and grid specification choices are entered in boxes. Then all one has to do is to press the *run* button. Figure 1 shows a screen shot of the software's GUI.

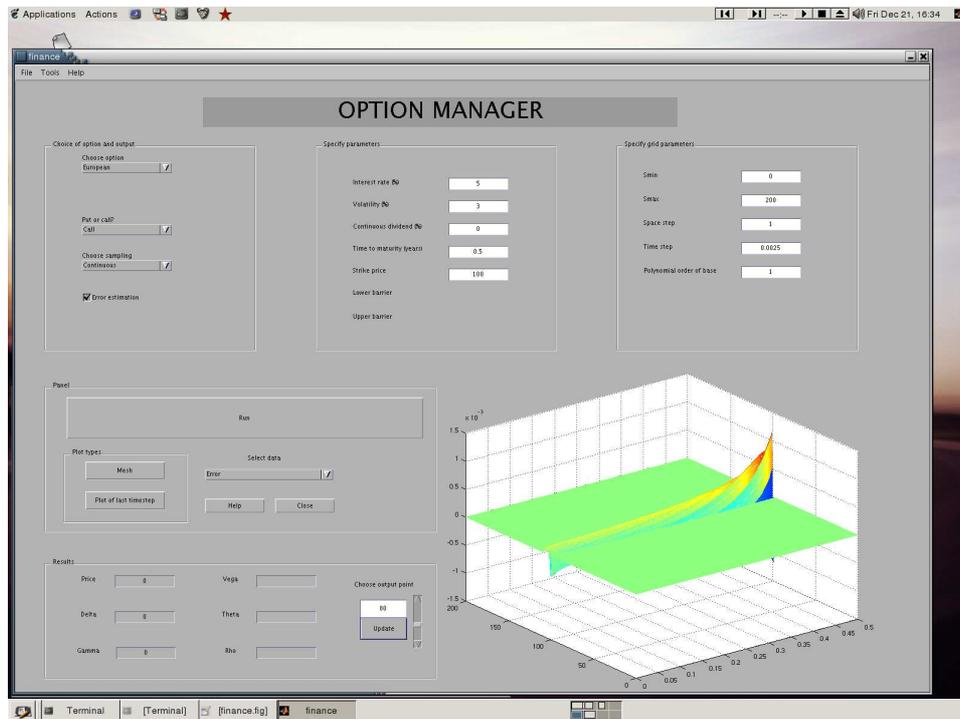


Figure 1: *Screen shot of Option Manager.*

If one ticks the box error estimation the software also calculates the solution to the corresponding dual problem to the chosen problem. The dual problem is used to calculate the error. For more information about a posteriori error estimation, and dual techniques, concerning option pricing

problems we refer to [4], [2], and [3].

The option prices, dual solutions, errors, and the Greeks, can be viewed as a 2D-plot at a certain time or as time-space evolution which makes it possible to view the solution over the whole space-time domain. This gives the user not just a tool for calculating prices and Greeks in an easy understandable way, but at the same time it aids to the understanding with visualization of the prices, Greeks, and error plots. Also for a user not familiar with the finite element method this helps in understanding how the dual techniques works and what role the dual solution plays. The prices and the Greeks are also presented numerically in boxes at the bottom left corner at a point of interest in space chosen by the user at time $t = 0$.

3 A Brief Users Guide

The option manager graphical user interface is divided into five different panels and one plot window, each with a different function, see Figure 1. These five panels or groups are named *Choice of option and output*, *Specify parameters*, *Specify grid parameters*, *Run and Visualization*, and *Results*.

3.1 Choice of Option and Output

In this panel one chooses which option that is to be valued by using the drop-down menus. The choice of option might trigger the appearance of another drop-down menu. Here the user also enters which type of option that is to be valued, put or call, and which sampling that is to be used. If the box error estimation is ticked then the dual solution and the a posteriori error is also calculated. In Figure 2 we a screenshot of the first panel.

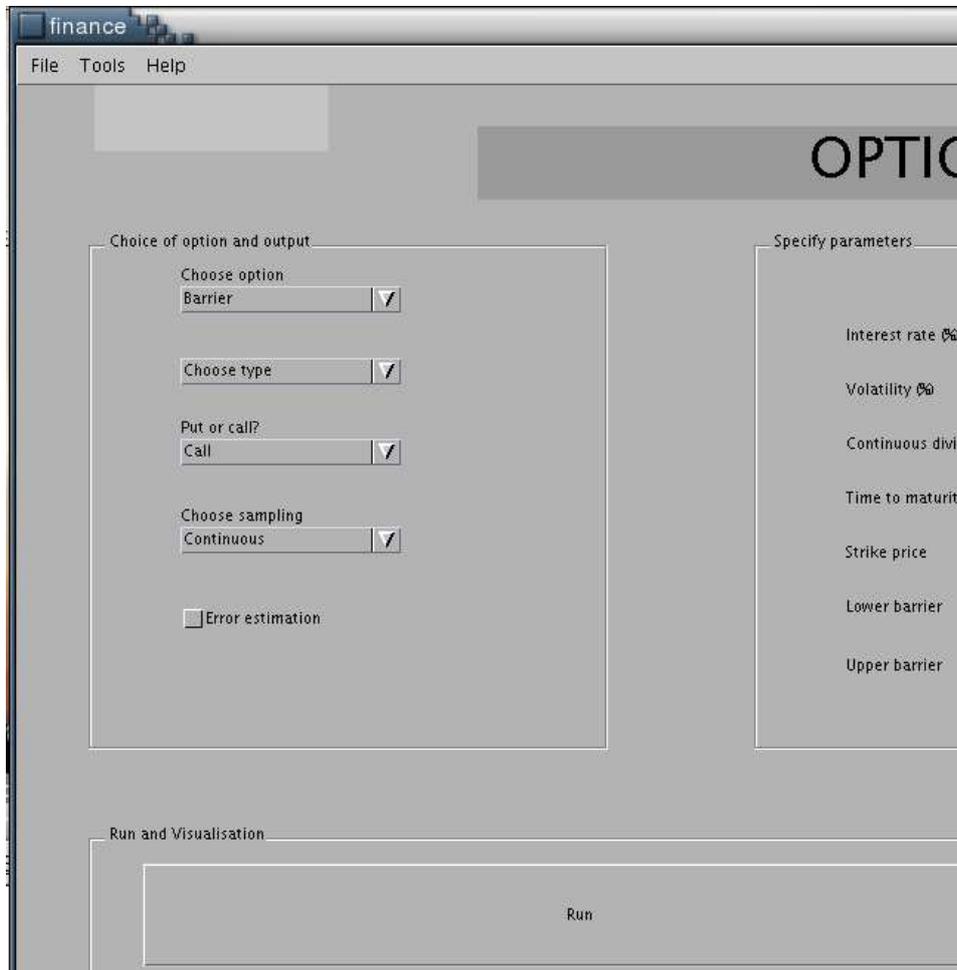


Figure 2: Screen shot of *Option Manager*, the first panel in the upper left corner.

3.2 Choice of Parameters

In this panel the option parameters are set, such as the interest rate, the volatility, the dividend yield, the strike price, and if necessary the lower and/or upper barrier level, by entering the values for these into the boxes. By pressing the *tab* button one easily switches to the next box after entering a value. In Figure 3 we see a screenshot showing the panel *Choice of parameters* and *Specify grid parameters*.

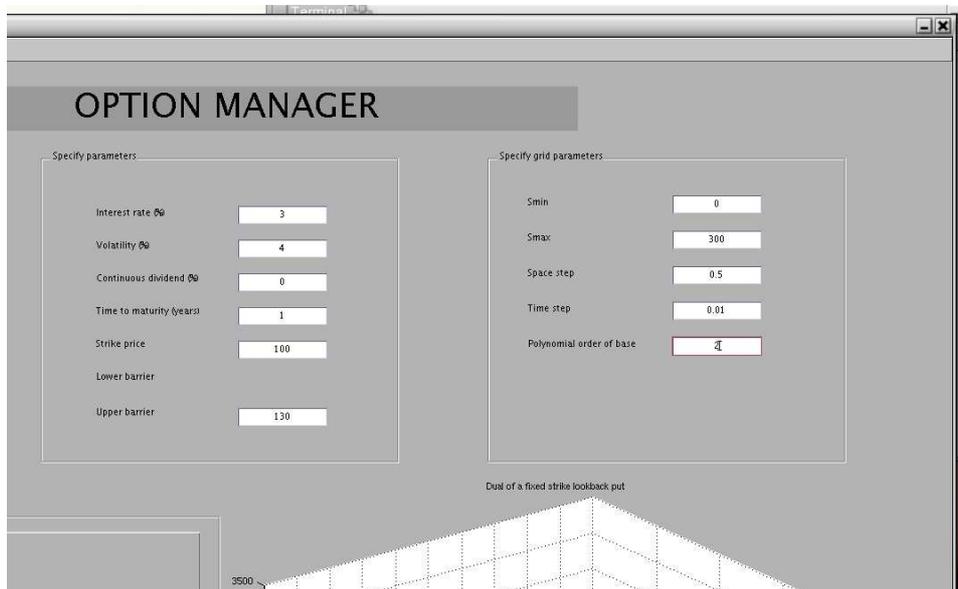


Figure 3: Screen shot of Option Manager, the second and third panel in the upper right corner.

3.3 Specify Grid Parameters

The grid parameters that are used in the calculation are specified in this panel by simply entering them into the boxes. One can also upload a data-sheet with parameters by entering the *File* menu in the upper left corner of the window. All parameters are then set automatically according to the data-sheet.

3.4 Run and Visualization

When all the previous panels have been attended to one is ready to start the calculation by pressing the *Run* button. It might take a while for the solution to appear in the *Results* panel and in the graphical window depending on whether or not error estimation was chosen in the first panel. In the drop down menu *Select data* one chooses what type of data that is to be displayed graphically in the window to the right. The choices are *Primal solution*, *Dual solution*, *Error*, *Gamma*, and *Delta*. The choice is finalized by choosing which type of visualization that is wanted by pressing either of the buttons *Mesh* or *Plot of last time step*. The *Mesh* button gives a space time plot over the evolution of the chosen data. The results are presented in the window in

the bottom right corner of the GUI. In this panel there are also the two other buttons, *Help* and *Close*. By pressing the *Help* button a help menu opens, and by pressing the *Close* button the program is terminated. In Figure 4 we see a screenshot with an up and out barrier call option.

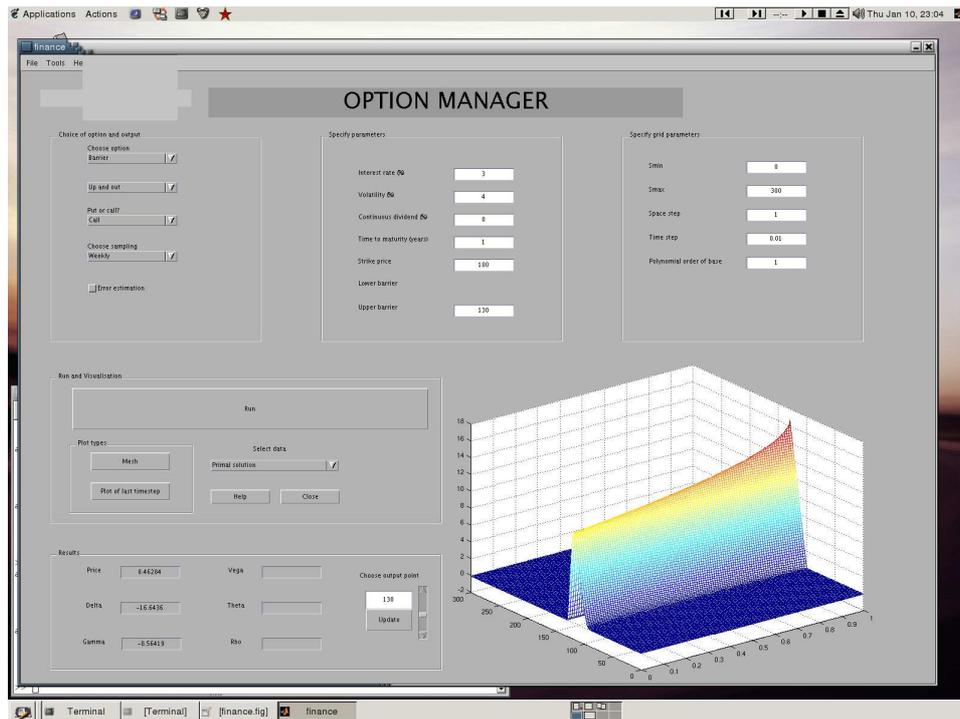


Figure 4: *Screen shot of Option Manager. In the plot window we see the value of an up and out barrier call option.*

3.5 Results

This panel contains the results in number form. The option price and the Greeks are presented in the boxes at a point of interest in space. The point is chosen by either entering a numeric value in the box and then pressing the *Update* button or by using the slider to the right.

4 Future Developments

There are some obvious improvements of the graphical user interface and the valuation program. More options needs to be included, also multidimen-

sional contracts. The visualization feature then also needs to be updated. For example it would be nice to be able to rotate the figure and to use the other Matlab tools such as magnify, changes of axes etc. The use of adapted meshes needs to be incorporated. The program should have some adapted meshes saved for each option. Also one should be able to compute an adapted mesh with the program.

References

- [1] S. BRENNER AND L. SCOTT, *The mathematical theory of finite element methods*, Springer Verlag, 1994.
- [2] G. FOUFAS AND M. G. LARSON, *Valuing european, barrier, and lookback options using the finite element method and duality techniques*. 2007.
- [3] ———, *Valuing fixed strike lookback options using the finite element method and duality techniques*. 2007.
- [4] ———, *Valuing asian options using the finite element method and duality techniques*, to appear in JCAM, (2008).
- [5] J. TOPPER, *Financial Engineering with finite elements*, Wiley Finance, 2005.
- [6] C. ZIENKIEWICZ, *The birth of the finite element method and of computational mechanics*, Internat. J. Numer. Methods Engrg., 60 (2004), pp. 3–10.