Valuation of Derivative Securities using Stochastic Analytic and Numerical Methods

David Heath

A thesis submitted for the degree of Doctor of Philosophy
of the Australian National University
November 1995
Declaration

I declare that, except where otherwise stated in the text, the work presented in this thesis is my own.

8/11/95
Acknowledgement

I gratefully acknowledge and thank the following people and institutions: My supervisor Dr. sc. E. Platen for his technical guidance and expertise, for sharing his wide knowledge and skills in applied stochastic analysis, and for his judgement and sound advice on various strategic issues. Bankers Trust, Australia and in particular Dr. G. Kentwell for contributions made to the Centre for Financial Mathematics at the Australian National University. Professors C. C. Heyde, R. Elliott, F. Delbaen and J. van der Hoek and Drs. A. Brace and M. Musiela for their helpful discussions with me. Ms. K. Platen who patiently and efficiently prepared the LaTex version of this document and my wife Tricia, children Rachael and Patrick, and parents John and Kitty for their love, support and encouragement during the period of this research.
Abstract

This thesis details methods and procedures to compute prices and hedging strategies for derivative securities in financial mathematics using stochastic analytic, numerical and variance reduction techniques.

Results are obtained on explicit hedge ratio representations for non-smooth payoff functionals and multidimensional diffusion processes with stopping boundaries. These methods are used to determine hedge ratios for the maximum of several assets and lookback options. A number of powerful variance reduction techniques are described. These include the use of measure transformations, discrete versions of importance sampling estimators, control variates based on Ito integral representations, stratified sampling and quasi Monte Carlo. For many of these techniques explicit formulas for the variance of the resulting estimators are obtained.

Pricing and hedging procedures are developed for a class of foreign exchange barrier options under stochastic volatility. These procedures are applied to the calculation of down-and-out call options for the Heston model. A general methodology for pricing discount bonds and options on discount bonds for multifactor term structure models is established. This approach is used for both European and American style securities for a version of the two-factor Fong and Vasicek model, extended to include time dependent parameters. For American pricing an exact representation of the early exercise premium is derived for a class of one-factor models. This enables both American option prices and the corresponding two-dimensional critical exercise boundary to be computed for the extended Fong and Vasicek model.
Contents

Declaration ................................................................. III
Acknowledgement ....................................................... V
Abstract ........................................................................ VII
Preface ........................................................................ XI
Survey ........................................................................... XV

Chapter 1: Calculation of Hedge Ratios for Non-smooth Payoff Functionals .......... 1
1.1 Contingent Claim Pricing Fundamentals .............................................. 2
1.2 Explicit Hedge Ratios for One-Dimensional Diffusions ...................... 5
1.3 Explicit Hedge Ratios for Multidimensional Diffusions .................. 10
1.4 Hedge Ratios for Non-Smooth Payoffs .......................................... 14
1.5 Absolutely Continuous Payoff Functions ...................................... 22
1.6 Maximum of Several Assets ......................................................... 26
1.7 Explicit Hedge Ratios for Lookback Options .................................... 31

Chapter 2: Variance Reduction Techniques ............................................ 39
2.1 Measure Transformations ........................................................... 40
2.2 An Alternative Variance Reduced Estimator .................................. 44
2.3 Discrete Time Variance Reduced Estimators .................................. 50
2.4 Control Variates and Integral Representations ................................ 59
2.5 Other Variance Reduction Methods ........................................... 66

Chapter 3: Valuation of Barrier Options under Stochastic Volatility .......... 73
3.1 The Black-Scholes Framework ...................................................... 74
3.2 A Model With Stochastic Volatility .............................................. 77
3.3 Numerical Procedures for Barrier Options .................................... 80
3.4 Simulation Results for Barrier Options ......................................... 86
Chapter 4: Valuation of European Bond Options
for Two-Factor Interest Rate Models ........................................ 93

4.1 Two-Factor Stochastic Volatility Models .................................. 94
4.2 Pricing of Contingent Claims .................................................... 97
4.3 Stochastic Numerical and Variance Reduction Methods .............. 101
4.4 Application to the Extended Fong and Vasicek Model ............... 105
4.5 Experimental Results for European Pricing ............................... 109

Chapter 5: Valuation of American Bond Options
for Two-Factor Interest Rate Models ........................................... 115

5.1 American Options for Two-Factor Bond Models ......................... 116
5.2 Representation of the Early Exercise Premium ........................... 118
5.3 Explicit Representation of the Early Exercise Premium
for the Extended Vasicek Model ............................................... 123
5.4 Computation of Terms for the Extended Vasicek Model ............... 128
5.5 American Pricing for the Extended Fong and Vasicek Model ........... 131
5.6 Computational Results for American Pricing ............................. 132

Bibliography .............................................................................. 137
Preface

Most financial assets evolve in an uncertain manner over time and, as a result, the general theory of stochastic processes is viewed by many as providing the natural mathematical framework for the analysis, valuation and management of these securities. This theory, both in its discrete and continuous time forms, provides a powerful and unifying set of analytic tools which forms the basis of a growing number of successful applications to financial markets. These applications have their origins in the seminal work of Black & Scholes (1973), Merton (1973), and the arbitrage-free pricing methodology developed by Harrison & Kreps (1979), Harrison & Pliska (1981) and Duffie (1988).

The theory of stochastic processes when applied to derivative security valuations requires, in essence, three main problem areas to be addressed. Firstly, there is the challenge of finding suitable stochastic models that can represent the underlying securities. Secondly, there is the problem of parameter estimation and fitting the model to actual market data. This task is usually integrated with procedures for using the model and is sometimes referred to as model calibration. Finally, there is the problem of computing prices and hedging strategies based on the models.

This thesis will focus on the third problem area - the development of pricing and hedging procedures for derivative securities with particular emphasis on the use of stochastic analytic, numerical and variance reduction techniques. The main aim is to show that these advanced mathematical tools can now be applied to solve a range of difficult and challenging valuation problems. For all of the applications described in this thesis, corresponding software systems have been built which deliver fast, reliable and accurate pricing and hedging of the corresponding security. The technology represents a significant improvement over existing methods and approaches.

A number of new results are included in this thesis. Some of these refer to new pricing methods, formulas and perspectives, and others refer to extension of existing methods but applied to new classes of problems. Also, each of the applications covered includes new development of the underlying theory. The overall theoretical framework proposed, which has been driven by real world applications, should be both of independent mathematical interest and of practical value as it can be successfully applied to a wide class of valuation and hedging problems.

This thesis is divided into two parts and contains five chapters. Part I, which concentrates more on theoretical issues, presents a number of general mathematical tools which can be used for the pricing and hedging of derivative securities. Part II deals with applications demonstrating how these tools can be applied to some key valuation and hedging problems. The contents of the chapters are as follows: Chapter 1 describes a general valuation methodology for and a new approach to finding explicit
expressions for the integrands in Ito integral representations of contingent claim payoff structures. This result is established firstly for one-dimensional diffusion processes and is then extended for multidimensional diffusion processes with stopping boundaries. Using general results from the theory of measure and integration, conditions are found under which these results can be strengthened to include a wide class of non-smooth payoff functionals. These methods are applied by finding corresponding representations for one-dimensional absolutely continuous functionals, the maximum of several assets and lookback options.

A number of variance reduction techniques based on stochastic analytic techniques are outlined in Chapter 2. These methods are mainly used to improve the performance of the raw Monte Carlo estimator by finding new ones having the same expectation but smaller variance. Some new variance reduction methods will be described as well as extensions to, and new perspectives on, some existing or classical ones. These include the use of general measure transformation procedures, discrete versions of importance sampling estimators, control variates based on Ito integral representations and new approaches to stratified sampling and quasi Monte Carlo. For a number of these methods, the variance of the resulting estimators is computed explicitly. This is of considerable practical value as it provides the basis for precise controls of the factors which contribute to the variance of an estimator.

Pricing and hedging procedures for a class of foreign exchange barrier options under stochastic volatility are considered in Chapter 3. A general valuation methodology is developed using mean self-financing arguments and the minimal equivalent martingale measure. This methodology is then applied by computing the prices and hedge ratios of down-and-out calls for the Heston (1993) model. Fast and accurate valuations are obtained by using a combination of control and antithetic variates and stratified sampling techniques, together with a derivative free weak approximation. It is shown that these methods can be adapted to suit the observation frequency or fixings of the barrier option.

The pricing of discount bonds and European style contingent claims for multifactor term structure models is dealt with in Chapter 4. The approach is demonstrated by efficiently computing the prices of discount bonds and European call options on bonds for a version of the Fong & Vasicek (1991a,b) model. This version is extended to include time dependent parameters in the drift term of the short rate for the model. It is shown that option prices and corresponding hedge ratios can be computed using representations under the so-called forward measure together with appropriate combinations of stochastic and deterministic numerical methods.

In Chapter 5 the analysis provided in the previous chapter is extended to include the valuation of American options. A methodology for pricing American puts for a class of two-factor term structure models using an integral representation of the early exercise premium is described. An exact form for this representation in the case of an extended version of the Vasicek (1977) model is derived. These results are then applied,
together with appropriate stochastic and deterministic numerical methods, to compute the prices of American puts for the two-factor extended Fong & Vasicek (1991a,b) model considered in Chapter 4. A number of simulation experiments are described which show that both American option prices and the corresponding two-dimensional critical exercise boundary can be efficiently estimated.

Some of the main characteristics which distinguish the use of numerical methods in this thesis compared to some previous treatments of the subject include the following: Firstly, and most importantly, these techniques are based mainly on the application of stochastic analytic principles and the semimartingale calculus. This approach has enabled new theoretical insights to be gained and provides support and a rigorous mathematical framework for a wide class of valuation and hedging problems to be handled. In fact the successful application of these methods has only been achieved by using some of the most powerful and deepest results from stochastic analysis, supported by a range of other techniques from numerical analysis and general simulation. It has also meant that the detailed structure of specific models can be more easily exploited.

Secondly, the systematic application of higher order numerical approximations is emphasized. For example, excellent results have been achieved with derivative free, predictor-corrector and extrapolated schemes, described in Kloeden & Platen (1992) and Hofmann, Platen & Schweizer (1992).

Thirdly, variance reduction, based on the combined use of several procedures has been used. Previous researchers have applied usually one or two separate techniques, typically antithetic and control variates from general simulation. However huge overall gains can be achieved by building systems which combine a number of complementary methods. For example, as will be explained in Chapter 2, finite-difference approximations can form the basis of control variate estimates in stochastic simulations and these can be combined with quasi Monte Carlo techniques.

Finally the theory developed in this thesis has been specifically designed to cater for multidimensional pricing problems and models. This type of modelling is increasingly required for many types of path-dependent and global securities, groups of assets, and even single instruments, where either the drift or diffusion components are themselves stochastic. All of the applications covered in Chapter 3 to 5 are based on multidimensional models.

The numbering system used in this thesis is as follows: Equations are numbered by their section and number in the section in parentheses where the reference occurs within the same section or chapter. The chapter number appears as a prefix, with the full reference (again in parentheses) when the equation is referred to in other chapters. All figures, lemmas and theorems, are numbered by their chapter, section and order of appearance within a section and do not appear in parentheses. Except for Section 1.4 the use of lemmas, propositions and theorems is avoided. This is to provide a more descriptive and expository account and a style of presentation which encourages a more appropriate discussion of the practical features of the methods used and their results.
A Brief Survey of Numerical Methods in Finance

Background

There is now considerable interest both from academics and practitioners in the application of stochastic modelling and other advanced mathematical methods to support the pricing and management of derivative securities in the finance area. A derivative security or contingent claim is one whose value is dependent on, or is derived from, some other underlying asset or security such as a bond, stock or currency contract. Over the last decade the growth in the use of derivative securities has been enormous. Financial institutions are seeking new products, new ways of handling existing ones and better methods for managing the risks associated with trading in these instruments.

In the past, because of the intractability of the underlying stochastic models, the development of accurate pricing methods has been extremely difficult. Analytic solutions to valuation problems are possible only in a few specialized cases, for example the classical Black and Scholes model based on a single asset and geometric Brownian motion. Consequently, numerical techniques are required for many types of valuation problems.

The use of discrete time methods including the application of stochastic numerical procedures, is in some sense fundamental and natural to an understanding and treatment of financial markets because individual financial securities are in fact observed and traded at discrete points in time. The continuous time theory is however extremely useful in providing a more concise formalism, clarifying insights and asymptotic limits to valuation problems.

The rapid development of new derivative securities and corresponding methods for pricing and managing them can be expected to continue in the future with particular emphasis on multidimensional modelling, complex payoff structures and the integration of risk management procedures over many instruments and even across divisions within financial institutions. For these type of challenges it is likely that both stochastic and deterministic numerical methods will play an increasingly important and crucial role.

Numerical approximations are now widely used in the finance industry. Even in cases where so-called exact valuations exist, computations based on these valuations are often supported by an array of deterministic numerical procedures such as interpolation, equation solving, search techniques, differentiation and integration routines. These numerical methods have become more effective in recent years due to increases in the power of desktop workstations and computers, at reduced costs, and the widespread availability of comprehensive numerical software packages, both in the commercial and
public domains.

Broadly speaking three main categories of numerical approximations have been used for the valuation of derivative securities. These are finite-difference methods, multinomial lattices and Monte Carlo simulation. In addition, a number of analytic approximations have been proposed and applied to valuation problems.

**Finite-Difference Approximations**

Finite difference approximations are used to solve numerically the Kolmogorov backward or Feynman-Kac equation with associated boundary conditions. Subject to certain integrability and smoothness conditions, this partial differential equation must be satisfied by the valuation process viewed as a function of time and the state variables in the underlying model. In the case of American options, a free boundary problem arises which can also be expressed as a variational inequality.

The method in its implicit form was first applied to option valuation problems by Schwartz (1977) and Brennan & Schwartz (1977). The valuation of derivative securities using explicit finite-difference methods has been analysed by Hull & White (1990) and described in the monograph by Hull (1993). The relationship between option pricing and the Feynman-Kac formula has been explored by Duffie (1988) and in a wider context by Karatzas & Shreve (1988). The monograph by Wilmott, Dewynne & Howison (1993) provides a comprehensive and accessible introduction to the use of partial differentiation equation solution techniques in finance. Dempster (1994) and Dempster & Hutton (1994) present some recent and encouraging results on the use of finite-difference and related approximations for the valuation of a range of option securities. These results focus on the use of linear programming techniques combined with matrix factorization.

Also of interest is the method of lines applied to general multidimensional free boundary problems by Meyer (1977) and used for American pricing by Carr & Faguett (1994) and Meyer & van der Hoek (1994). This technique, related to finite-difference methods, usually involves discrete time approximations of the underlying partial differential equations by ordinary differential equations. In the case of multidimensional pricing, some discretization of one of the state space variables is needed and this results in a set of equations which must be solved or numerically evaluated at each time step. For the Black and Scholes model these resulting ordinary differential equations can be solved explicitly and the approach is then referred to as the analytic method of lines.

Finite-element methods can also be used to approximate the solutions of a wide class of partial differential equations. Because of the relative ease with which boundary conditions can be handled, they are better suited to free boundary problems and therefore to the pricing of certain types of American options. A description of these methods can be found in most books on numerical analysis, for example Burden & Faires (1993) and Hoffmann (1992). In addition Wilmott, Dewynne & Howison (1993) provide in-
formation on the formulation of the finite-element method for American pricing.

**Lattice Methods**

Binomial and multinomial lattice techniques constitute a popular and widely used numerical procedure for pricing a range of derivative securities. The basic binomial method values a derivative security by approximating the underlying diffusion with binomial trees for each process component where at each time step model components can move to at most two new values. The method was introduced for option pricing by Cox, Ross & Rubinstein (1979). Usually the technique is applied backwards in time and when used to price American options represents an application of the Bellman principle of dynamic programming. A faster version of the method has been developed by Breen (1991).

Approximation of the underlying diffusion by multinomial trees, where model components can move to a finite number of new states or values, leads to corresponding multinomial lattices. Trinomial and multinomial lattices have been analysed by Boyle (1988), Omberg (1988) and Boyle, Evnine & Gibbs (1989).


**Monte Carlo Simulations**

The application of Monte Carlo methods to option pricing was first described by Boyle (1977). According to the arbitrage-free pricing theory the price of many types of contingent claims can be expressed as the discounted expected value of its terminal future payoff under an appropriately changed probability measure. A standard Monte Carlo procedure would estimate this price by simulating many trajectories for the underlying stochastic model using this measure. The Monte Carlo estimate is then the discounted sample mean of the terminal payoffs for these simulated trajectories.

Monte Carlo simulations are used by both academics and practitioners to provide comparative results for other methods, and ‘rough’ estimates when no other valuation procedure is available. Since the pioneering work of Boyle, several authors have applied the technique to a variety of valuation problems. These include Hull & White (1987, 1988), Johnson & Shanno (1987) and Scott (1987) on stochastic volatility, Schwartz & Torous (1989) on mortgage-backed securities and Kenma & Vorst (1990) on the valuation of Asian options. Hofmann, Platen & Schweizer (1992), Duffie (1992) and
Duffie & Glynn (1992) have also applied Monte Carlo methods for derivative security valuation problems.

For a number of valuation problems the pricing functional can be expressed as an integral involving the densities of the underlying model components. If these densities have an explicit form, the general methods of quasi Monte Carlo can be applied. These methods are described in the monographs by Ripley (1983) and more recently Niederreiter (1992). Their applicability to financial modelling problems have been examined by Joy, Boyle & Tan (1995). Quasi Monte Carlo in its standard form involves replacing pseudo-random numbers with so-called low discrepancy point sets such as Sobol (1967) or Halton (1960) sequences. These low discrepancy sequences exhibit less deviations from uniformity compared to pseudo random numbers. Quasi Monte Carlo, in cases where densities of the model components can be determined, seem to be well-suited to higher dimensional problems. A version of quasi Monte Carlo involving the use of multipoint random variables is considered in Chapter 2 of this thesis.

Recently Hofmann, Platen & Schweizer (1992), using measure transformation techniques, and Clewlow & Carverhill (1992, 1994), using discrete versions of a martingale control variate, have introduced powerful variance reduction methods which significantly improve the performance of the raw Monte Carlo estimator. The paper by Hofmann, Platen and Schweizer is particularly significant as it represents the first successful attempt to formulate and use variance reduction techniques for financial modelling problems based on stochastic analytic methods.

The application of Monte Carlo methods in general simulation has been described by several authors. Some excellent references include Hammersley & Handscomb (1964), Ripley (1983), Bratley, Fox & Schrage (1987), Ross (1991) and Law & Kelton (1991). All of these sources provide information on variance reduction techniques. It is somewhat surprising that, since the work of Boyle (1977), these classical methods have only recently been used systematically in the finance area.

**Analytic Approximations and Related Methods**

Analytic approximations have been used for both European and American option valuations. Some examples of these in the case of American options include the quadratic approximation method developed by MacMillan (1986) and Barone-Adesi & Whaley (1987), the compound option approach used by Geske & Johnson (1984) and refined by Bunch & Johnson (1992). McKean (1965), Kim (1990), Jacka (1991), Carr, Jarrow & Mynteni (1992) and Chesney, Elliott & Gibson (1991) price American options as the sum of the corresponding European price together with an integral representation of the early exercise premium. This representation is exact but requires a backward numerical technique to determine the optimal exercise boundary and from this the price of the option. Examples of analytic approximations using Taylor series expansions and applied to European style contingent claims are given by Dothan (1987) and Hull &
Laplace transform methods have been used to value Asian options by Geman & Yor (1992) and Eydeland & Geman (1995) and Parisian options by Chesney, Jeanblanc-Picque & Yor (1995). Inversion of the Laplace transform for Parisian options requires some delicate numerical problems to be solved. These problems together with methods for handling them are presented by Cornwall et al. (1995). The fast Fourier transform is used by Carverhill & Clewlow (1990) to evaluate Asian options.

**Performance Criteria**

The performance of a numerical method is measured in terms of its speed, accuracy, flexibility, robustness and ease of implementation. The attribute of speed is related to the rate of convergence of the numerical method to a continuous time limit and is clearly of considerable importance to financial institutions. This is because these institutions often have large books of securities which must be frequently revalued and hedged. Also, sensitivity analysis studies, model calibration procedures and calculation of implied parameters often require numerically intensive computations.

The requirements for fast valuations must be balanced with the competing demands of accuracy which, for financial modelling problems, usually varies considerably, depending on the application. For example, if hedge ratios are being computed using finite difference approximations, the underlying valuation procedures may need to be very accurate. On the other hand, sensitivity analysis work may require less accuracy with more emphasis being placed on providing rapid feedback to risk managers. Also the requirements for accuracy need to be worked out for the entire valuation problem not just for a particular component. Clearly if the parameter estimation and modelling errors associated with a product are say, of the order of 5 per cent, which may easily arise for new products and certain types of exotics, it is of limited value having valuation software that is accurate to say 0.1 per cent.

The flexibility of a numerical method refers to how easily the method can be adapted to new problems and situations. This attribute is particularly important to institutions at the cutting-edge of research and development.

Robustness is measured in terms of the stability of the corresponding method. A numerical method is stable if small changes in input parameters lead to small changes in output results. Some algorithms are stable in certain regions and not in others. The differences in the modes of convergence and dynamics of stochastic processes lead to different stability criteria for stochastic systems, see for example Kloeden & Platen (1992). An important part of the work required to build valuation procedures based on stochastic numerical methods is therefore a detailed and systematic study of stability issues relating to the dynamics of the underlying stochastic system.

Finally, we can measure the ease with which a numerical method can be formulated and implemented within a subroutine library or separate application. This is related to
the complexity of the method and its capacity to be modularized and broken down into components. This is also an important attribute as software development expenditure can be considerable with time and cost overruns common.

If the various classes of numerical methods are compared and evaluated in terms of their performance, it is found that in general the fully numerical techniques are more flexible but slower and less accurate compared to analytic approximations. Stability problems depend on the specific application being considered but are more likely to be an issue for fully numerical techniques compared to analytic approximations. There are considerable differences between the methods in terms of their ease of implementation. For example quadratic approximations for American options can usually be implemented in a straightforward manner, whereas the integral representation of the early exercise premium seems to require more delicate numerical problems to be solved.

As observed by Brennan & Schwartz (1978), multinomial lattice approaches can be regarded as versions of explicit finite-difference methods. The methods are therefore roughly similar in terms of their performance capabilities. However Geske & Shastri (1985) compare the two approaches and note some differences in efficiencies for different types of applications. The main disadvantage of finite-difference and lattice approaches is that memory requirements explode exponentially with increases in the dimensionality of the model and the number of branch points used. The speed of convergence to a continuous time limit can also be extremely slow, particularly for certain classes of path dependent options. Cheuk & Vorst (1994a,b) report very slow convergence of binomial pricing methods for barrier and lookback options. Their research in the case of lookback options also indicates the sensitivity of valuations to the observation frequency of the option.

Monte Carlo techniques are generally considered to be flexible and in their basic form simple to implement although somewhat inefficient. Compared to lattice and finite-difference methods they seem to be better suited to higher dimensional problems. A common view, see for example Hull (1993) and Reider (1994) is that they are not as effective for American pricing. However Clewlow & Carverhill (1992, 1994) and Grant, Vora & Weeks (1993) have applied Monte Carlo techniques to American valuations. As will be demonstrated in this thesis, a two-factor American problem is handled using techniques related to Monte Carlo simulation. An excellent comparison and evaluation of various approximation methods for American options together with a description of a new approximation technique based on capped options is given by Broadie & Detemple (1994).
Chapter 1

Calculation of Hedge Ratios for Non-Smooth Payoff Functionals

The calculation of hedge ratios is fundamental to both the arbitrage free valuation of derivative securities and also the risk management procedures needed to replicate these instruments. In this chapter we consider the problem of finding explicit Ito integral representations of the payoff structure of a derivative security. If such a representation can be found in an explicit form, the corresponding hedge ratio can usually be easily identified and calculated.

In Section 1.1, following from the work of Bensoussan (1984), Karatzas & Shreve (1988), and Hofmann, Platen & Schweizer (1992), we propose a general framework which expresses the price dynamics of a derivative security as the conditional expectation, under a suitable measure, of the payoff structure of the security. This framework will be used throughout this thesis.

In Section 1.2 and 1.3 we apply the Markov property and the Kolmogorov backward equation to obtain explicit Ito integral representations for a class of smooth payoff functionals, firstly for one-dimensional diffusion processes and, secondly, for multidimensional diffusion processes with stopping boundaries. The extension to include stopping boundaries is needed to support the pricing and hedging of American options. These representations are related to the formula of Clark (1970) and results obtained by Haussmann (1979), Ocone (1984), Elliott & Kohlmann (1988) and Colwell, Elliott & Kopp (1991).

Applying general arguments from the theory of measure and integration, we then extend these results in Section 1.4 to include a wide class of non-smooth payoff functionals which can be expressed as the pointwise limit of smooth functionals that satisfy a uniform linear growth bound. In Section 1.5 we use these results together with certain smoothing operators to obtain explicit Ito integral representations for one-dimensional absolutely continuous functionals whose derivative is continuous except at a finite number of points. In the last two sections of this chapter we extend these results to include representations of the maximum of several assets and lookback options. The representation of the maximum of several assets provides an important tool for the computation of hedge ratios in the funds management area. In the case of lookback options we provide an example of how these methods can be adapted to the case of path-dependent options.
1.1 Contingent Claim Pricing Fundamentals

Let $W = (W^1, \ldots, W^m)$ be an $m$-dimensional Brownian motion defined on the probability space $(\Omega, \mathcal{F}, P)$. We assume that the filtration $\mathcal{F} = (\mathcal{F}_t)_{t \geq t_0}$ is the $P$-augmentation of the natural filtration of $W$. These conditions ensure, see Karatzas & Shreve (1988), Proposition 2.7.7, that $\mathcal{F}$ satisfies the usual conditions.

Let $X^{t_0, z} = \{X^{t_0, z}_t = (X^{t_0, z}_{t_1}, \ldots, X^{t_0, z}_{t_m}), t_0 \leq t \leq T\}$ be a $d$-dimensional diffusion process whose components satisfy the stochastic differential equation

$$dX^{t_0, z}_t = a^i(t, X^{t_0, z}_t) \, dt + \sum_{j=1}^m b^{i,j}(t, X^{t_0, z}_t) \, dW^j_t$$

for $t_0 \leq t \leq T$, $i \in \{1, \ldots, d\}$, where $X^{t_0, z}$ starts at time $t_0$ with initial value $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$, and $\mathbb{R}^d$ is the set of $d$-dimensional reals. We assume that appropriate growth and Lipschitz conditions apply for the drift $a^i$ and diffusion $b^{i,j}$ coefficients so that (1.1) admits a unique solution and is Markov, see for example Kloeden & Platen (1992).

In order to model the time value of money and stochastic interest rates we take the first component $X^{1, t_0, z} = r = \{r_t, t_0 \leq t \leq T\}$ to represent some instantaneous interest rate process and the second component $X^{2, t_0, z} = \beta = \{\beta_t, t_0 \leq t \leq T\}$ to model the price movements of a riskless savings account. We assume this savings account $\beta$ evolves according to the stochastic differential equation

$$d\beta_t = r_t \beta_t \, dt$$

for $t_0 \leq t \leq T$, where $\beta$ starts at time $t_0$ with initial value $\beta_{t_0} = 1$. Note that (1.2) can be solved explicitly yielding

$$\beta_t = \exp \left( \int_{t_0}^t r_s \, ds \right),$$

for $t_0 \leq t \leq T$. The vector process $X^{t_0, z}$ could include several risky assets and other securities as well as components to provide for additional specifications or features of the model such as stochastic volatility or averages of risky assets for Asian options.

In order to build a framework that will support in particular American, and certain classes of exotic option valuations and hedging, we consider a stopping time formulation as follows:

Let $\Gamma_0 \subset [t_0, T] \times \mathbb{R}^d$ be some region with $\Gamma_0 \cap (t_0, T] \times \mathbb{R}^d$ an open set and define a stopping time $\tau : \Omega \to \mathbb{R}$ by

$$\tau = \inf\{t > t_0 : (t, X^{t_0, z}_t) \notin \Gamma_0\}.$$  

(1.4)

Using the stopping time $\tau$ we define the region

$$\Gamma_1 = \{ (\tau(\omega), X^{t_0, z}_{\tau(\omega)}(\omega)) \in [t_0, T] \times \mathbb{R}^d : \omega \in \Omega \}.$$
1.1. CONTINGENT CLAIM PRICING FUNDAMENTALS

\( \Gamma_1 \) contains all points of the boundary of \( \Gamma_0 \) which can be reached by the diffusion \( X^{t_0,x} \). We now consider contingent claims with payoff structures of the form

\[ h(\tau, X^{t_0,x}_\tau), \]

where \( h : \Gamma_1 \to \mathbb{R} \) is some payoff function.

Using terminology that is applied mainly for American option pricing, we call the set \( \Gamma_0 \) the continuation region and \( \Gamma_1 \) the exercise boundary, which forms part of the stopping region. For a diffusion process \( X^{t_0,x} \) with continuous sample paths, an option is considered ‘alive’ at time \( s, t_0 \leq s \leq T, \) if \( (s, X^{t_0,x}_s) \in \Gamma_0 \). On the other hand it is ‘exercised’ or ‘stopped’ at the first time \( s, t_0 \leq s \leq T, \) that \( (s, X^{t_0,x}_s) \) touches \( \Gamma_1 \). It is assumed that \( (t_0, x) \in \Gamma_0 \) since otherwise the derivative security would be immediately ‘exercised’.

For example if we take \( \Gamma_0 = [t_0, T) \times \mathbb{R}^d \) which implies \( \tau = T \) and payoff structures of the form \( h(T, X^{t_0,x}_T) \) this formulation reduces to the case of a multidimensional European style contingent claim.

Applying the Markov property, the option pricing or valuation function \( u : \Gamma_0 \cup \Gamma_1 \to \mathbb{R} \) corresponding to the payoff structure \( h(\tau, X^{t_0,x}_\tau) \) is given by

\[ u(t, x) = \bar{E} \left( \exp \left( - \int_t^T r_s \, ds \right) h \left( \tau, X^{t,x}_\tau \right) \right), \tag{1.5} \]

for \( (t, x) \in \Gamma_0 \cup \Gamma_1 \), where \( \bar{E} \) denotes expectation under an appropriately defined probability measure \( \bar{P} \). We will not discuss here how this measure \( \bar{P} \), which is usually the risk neutral measure, should be determined. A good choice, the so called minimal equivalent martingale measure, which can be used both for complete and incomplete markets is described in Hofmann, Platen & Schweizer (1992); see also Foellmer & Schweizer (1991) and Schweizer (1991).

Define the discounted functions \( \bar{h} : \Gamma_1 \to \mathbb{R} \) and \( \bar{u} : \Gamma_0 \cup \Gamma_1 \to \mathbb{R} \) by

\[ \bar{h}(s, y) = \frac{1}{y_2} h(s, y) \]
\[ \bar{u}(t, x) = \bar{E} \left( h \left( \tau, X^{t,x}_\tau \right) \right) \tag{1.6} \]

for \( (s, y) \in \Gamma_1 \), \( (t, x) \in \Gamma_0 \cup \Gamma_1 \) with \( y = (y_1, \ldots, y_d) \in \mathbb{R}^d \), where we recall that \( X^{2,t_0,x} \) represents the price movements of the riskless savings account \( \beta \).

Let \( Z^{t_0,t_0} = \{ Z^{t_0,t}_t, \ t_0 \leq t \leq T \} \) be the solution of the stochastic differential equation

\[ dZ^{t_0,t_0}_t = 1_{\{t<\tau\}} \, dt \tag{1.7} \]

for \( t_0 \leq t \leq T \), starting at time \( t_0 \) with initial value \( t_0 \). We can write the solution to (1.7) in the form

\[ Z^{t_0,t}_t = t \wedge \tau, \tag{1.8} \]
for \( t_0 \leq t \leq T \). This expression together with the uniqueness of the solutions of (1.1) and (1.7) shows that
\[
X^{t_0,\xi}_T = X^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{\tau}
\]
and
\[
\tau = Z^{t_0,\xi}_T = Z^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{\tau}
\]
for \( t_0 \leq t \leq T \). Using these equalities, (1.5), the Markov property, equation (1.3) and the assignment \( \beta_\tau = X^{2,\xi}_{2T} = 1 \), we have
\[
u_t = u \left( t \wedge \tau, X^{t_0,\xi}_{t_\wedge \tau} \right)
\]
\[
\begin{align*}
u_t &= E \left( \exp \left( - \int_{t \wedge \tau}^\tau r_s \, ds \right) h \left( Z^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{t_\wedge \tau}, X^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{t_\wedge \tau} \right) \right) \\
&= E \left( \exp \left( - \int_{t \wedge \tau}^\tau r_s \, ds \right) h \left( Z^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{t_\wedge \tau}, X^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{t_\wedge \tau} \right) \big| \mathcal{F}_t \right) \\
&= E \left( \frac{\beta_{t_\wedge \tau}}{\beta_\tau} h \left( Z^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{t_\wedge \tau}, X^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{t_\wedge \tau} \right) \big| \mathcal{F}_t \right) \\
&= \beta_{t_\wedge \tau} E \left( \frac{1}{\beta_\tau} h \left( Z^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{t_\wedge \tau}, X^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{t_\wedge \tau} \right) \big| \mathcal{F}_t \right) \\
&= \beta_{t_\wedge \tau} E \left( \frac{1}{\beta_\tau} h \left( t \wedge \tau, X^{t_0,\xi}_{t_\wedge \tau} \right) \right) \\
&= \beta_{t_\wedge \tau} E \left( \frac{1}{\beta_\tau} h \left( \tau, X^{t_0,\xi}_{t_\wedge \tau} \right) \right) \\
&= \beta_{t_\wedge \tau} E \left( \frac{1}{\beta_\tau} h \left( t \wedge \tau, X^{t_0,\xi}_{t_\wedge \tau} \right) \right) \\
&= \beta_{t_\wedge \tau} \tilde{u} \left( t \wedge \tau, X^{t_0,\xi}_{t_\wedge \tau} \right)
\end{align*}
\]
for \( (t \wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}) \in \Gamma_0 \).

Define the martingale \( M = \{M_t : t_0 \leq t \leq T\} \) by
\[
M_t = E \left( h(\tau, X^{t_0,\xi}_{t}) \big| \mathcal{F}_t \right),
\]
for \( t_0 \leq t \leq T \). We assume that an appropriate growth condition applies for \( \tilde{h} \) so that the conditional expectation in (1.12) is well-defined. Applying once again equation (1.9), the Markov property and the definitions of \( \tilde{h} \) and \( \tilde{u} \) we have
\[
M_t = \tilde{E} \left( \tilde{h} \left( \tau, X^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{t_\wedge \tau} \right) \big| \mathcal{F}_t \right)
\]
\[
= \tilde{E} \left( \tilde{h} \left( \tau, X^{t_\wedge \tau, X^{t_0,\xi}_{t_\wedge \tau}}_{t_\wedge \tau} \right) \right) 
\]
1.2. HEDGE RATIOS FOR ONE-DIMENSIONAL DIFFUSIONS

\[ \mathbb{E} \left( \tau, X_{t \wedge \tau}^0 \right) \]

for \( t_0 \leq t \leq T \).

Consequently the \( \beta \)-discounted valuation process

\[ \bar{u} = \left\{ \bar{u}_t = \bar{u} \left( t \wedge \tau, X_{t \wedge \tau}^0 \right), t_0 \leq t \leq T \right\} \]

is a martingale. Also, from (1.11) we can determine values for the random variable \( u_t \) if corresponding values for \( \beta_t, M \) and \( \bar{u}_t \) are known. Usually it is much more convenient to compute prices via the function \( \bar{u} \) rather than \( u \). This is mainly because the martingale property associated with \( \bar{u} \) enables us to apply a number of powerful results from stochastic analysis. In particular, subject to certain integrability conditions, the price process corresponding to \( \bar{u} \) will admit an Ito integral representation, and from this hedging parameters can be determined either in an implicit or explicit form.

We will not discuss here how these hedging strategies for general valuations can be formulated. Instead we refer the reader to the papers by Hofmann, Platen & Schweizer (1992) and Karatzas (1989) for a more complete coverage; see also Bensoussan (1984) and Karatzas & Shreve (1988). Some specific examples of hedging strategies are however considered in Part II of this thesis dealing with applications.

The above analysis leading in particular to equations (1.11) and (1.13) demonstrates that the valuation of contingent claims as given by (1.5) can be reduced to the case of valuations of the form

\[ \bar{u}_t = \bar{u} \left( t \wedge \tau, X_{t \wedge \tau}^0 \right) = \mathbb{E} \left( \tau, X_{\tau \wedge \tau}^0 \right) \mid \mathcal{F}_t \]  

(1.14)

for some payoff function \( \bar{h} : \Gamma_1 \rightarrow \mathbb{R} \). Consequently in the remaining part of this chapter and the next we will assume this type of structure for our pricing and hedging problems.

In the special case where \( \tau = T \) and the payoff structure takes the form \( h(X_{T \wedge \tau}^0) \) we will refer to the corresponding equations for (1.5) and (1.11)-(1.14) as the time-independent formulations.

### 1.2 Explicit Hedge Ratios for One-Dimensional Diffusions

In this section, to allow for an easier and simpler exposition of the underlying ideas, we suppose \( W = \{ W_t, t \geq t_0 \} \) is a one-dimensional Brownian motion defined on the probability space \( (\Omega, \mathcal{F}, P) \). Consider the one-dimensional stochastic differential equation

\[ dX_t = a(t, X_t) \, dt + b(t, X_t) \, dW_t \]  

(2.1)

for \( t_0 \leq t \leq T \). Here \( a, b : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R} \) are measurable functions which have linear growth, are Lipschitz continuous in \( x \) and whose partial derivatives have linear growth.
and are Lipschitz continuous in $x$. We denote by $X_{t_0}^{t,-x} = \{X_s^{t_0} : t_0 \leq s \leq T\}$ the solution of (2.1) starting at $x \in \mathbb{R}$ at time $t_0$, $t_0 \leq t \leq T$.

In this section we consider a European style contingent claim with a payoff structure of the form

$$h \left( X_T^{t_0,-x} \right) = h \left( T, X_T^{t_0,-x} \right)$$

for $x \in \mathbb{R}$ and terminal time $T$.

Subject to certain growth bounds applying for the function $h$ these conditions for the coefficients $a$ and $b$ ensure, see Kloeden & Platen (1992), that $E((h(X_T^{t_0,-x}))^2) < \infty$ and the process $M = \{M_t, t_0 \leq t \leq T\}$ defined by

$$M_t = E \left( h \left( X_T^{t_0,-x} \right) \mid \mathcal{F}_t \right) \quad (2.2)$$

for $t_0 \leq t \leq T$, is a square integrable martingale and therefore admits a (Kunita & Watanabe (1967)) representation of the form

$$M_t = M_0 + \int_{t_0}^t \xi_s \, dW_s, \quad (2.3)$$

where $\xi$ is an $\mathcal{F}$-predictable process with

$$E \left( \int_{t_0}^T \xi_s^2 \, ds \right) < \infty.$$

The process $\xi$ is unique in the sense that if $M_t = M_0 + \int_{t_0}^t \bar{\xi}_s \, dW_s$ for some other $\mathcal{F}$-predictable process $\bar{\xi}$, then

$$E((\int_{t_0}^t (\xi_s - \bar{\xi}_s) \, dW_s)^2) = \int_{t_0}^t E((\xi_s - \bar{\xi}_s)^2) \, ds = 0.$$

A more general statement of this uniqueness property can be found in Karatzas & Shreve (1988), Exercise 3.4.22.

Finding explicit expressions for the integrand $\xi$ is of considerable practical value as it is closely related to the computation of hedge ratios in the theory of option pricing. Here we seek explicit characterizations based on an application of the Kolmogorov backward equation.

Define the scalar function $u : [t_0, T] \times \mathbb{R} \to \mathbb{R}$ by

$$u(t, x) = E(h(X_T^{t_0,-x})), \quad (2.4)$$

for $(t, x) \in [t_0, T] \times \mathbb{R}$. We assume that the function $u$ is of class $C^{1,3}$, that is, continuously differentiable with respect to $t$ and three times continuously differentiable in $x$. Expanding $u(T, X_T^{t_0,-x}) = h(X_T^{t_0,-x})$ by the Ito rule and applying the Kolmogorov backward equation yields

$$u \left( t, X_t^{t_0,-x} \right) = u(t_0, x) + \int_{t_0}^t \frac{\partial u}{\partial x} \left( s, X_s^{t_0,-x} \right) b \left( s, X_s^{t_0,-x} \right) \, dW_s, \quad (2.5)$$
for $t_0 \leq t \leq T$. Using the time-independent formulations of (1.12), (1.13) with $\tau = T$, $h = \bar{h}$ and $u = \bar{u}$, and (2.5) we have

$$M_t = E \left( h \left( X_T^{t_0, \bar{z}} \right) \mid \mathcal{F}_t \right) = u(t, X_t^{t_0, \bar{z}}) = u(t_0, \bar{z}) + \int_{t_0}^t \frac{\partial u}{\partial x} (s, X_s^{t_0, \bar{z}}) b(s, X_s^{t_0, \bar{z}}) dW_s$$

(2.6)

for $t_0 \leq t \leq T$. This result can also be obtained by applying Ito’s formula to $u(T, X_T^{t_0, \bar{z}}) = h(X_T^{t_0, \bar{z}})$, taking the conditional expectation of both sides of the resulting equation, and using the relations (2.2) and (2.5). Consequently $M_{t_0} = u(t_0, \bar{z})$ and $\xi_s = \frac{\partial u}{\partial z} (s, X_s^{t_0, \bar{z}}) b(s, X_s^{t_0, \bar{z}})$. We now have a representation of the form (2.3). However this expression for the integrand $\xi$ requires the solution of the valuation equation (2.4) to be known. Typically, in practical applications, one uses finite differences to approximate the partial derivative $\frac{\partial u}{\partial z} (s, X_s^{t_0, \bar{z}})$ and from these $\xi_s$.

We will now find an alternate characterization of the integrand $\xi$ which does not depend directly on the solution to (2.4). Let $L^0$ and $\frac{\partial}{\partial x} L^0$ be the operators

$$L^0 = \frac{\partial}{\partial s} + a \frac{\partial}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2}{\partial x^2}$$

$$\frac{\partial}{\partial x} L^0 = \frac{\partial^2}{\partial s \partial x} + \frac{\partial a}{\partial x} \frac{\partial}{\partial x} + \left(a + b \frac{\partial b}{\partial x}\right) \frac{\partial^2}{\partial x^2} + \frac{1}{2} b^2 \frac{\partial^3}{\partial x^3}$$

(2.7)

where the operator $\frac{\partial}{\partial x} L^0$ is obtained by computing the partial derivative of $L^0$ with respect to $x$. The Kolmogorov backward equation can now be written in the form

$$L^0 u(s, x) = 0,$$

for $(s, x) \in (t_0, T) \times \mathbb{R}$ with terminal condition

$$u(T, x) = h(x)$$

(2.8)

for $x \in \mathbb{R}$, so that

$$\frac{\partial}{\partial x} L^0 (u(s, x)) = 0$$

(2.9)

for $(s, x) \in (t_0, T) \times \mathbb{R}$.

Consider the linearized stochastic differential equation

$$dZ_t = \frac{\partial a}{\partial x} (t, X_t) Z_t dt + \frac{\partial b}{\partial x} (t, X_t) Z_t dW_t$$

(2.10)

for $t_0 \leq t \leq T$. Let $Z^{s, \bar{z}} = \{Z^s_t, \ s \leq t \leq T\}$ be the solution of (2.10) starting at $\bar{z} \in \mathbb{R}$ at time $s$, $t_0 \leq s \leq T$. 
Introducing the scalar function $v : [t_0, T] \times \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by

$$v(t, x, z) = \frac{\partial u}{\partial x}(t, x)z$$

(2.11)

for $(t, x, z) \in [t_0, T] \times \mathbb{R}^2$, and applying the multidimensional version of Ito's formula to $v(T, X_T^{t,x}, Z_T^{t,z})$ we have

$$v(T, X_T^{t,x}, Z_T^{t,z}) = v(t, x, z) + \int_t^T L^0 v(s, X_s^{t,x}, Z_s^{t,z}) \, ds$$

$$+ \int_t^T L^1 v(s, X_s^{t,x}, Z_s^{t,z}) \, dW_s$$

(2.12)

for $t_0 \leq t \leq T$, where

$$L^0 = \frac{\partial}{\partial s} + a \frac{\partial}{\partial x} + b \frac{\partial}{\partial x} + \frac{\partial}{\partial z}$$

$$+ \frac{1}{2} \left( b^2 \frac{\partial^2}{\partial x^2} + \left( \frac{\partial b}{\partial x} \right)^2 z^2 \frac{\partial^2}{\partial z^2} \right) + b \frac{\partial b}{\partial x} \frac{\partial^2}{\partial x \partial z}$$

and

$$L^1 = b \frac{\partial}{\partial x} + z \frac{\partial b}{\partial x \partial z}.$$ 

Calculating the partial derivatives of the function $v$ using (2.11), and applying (2.9) yields

$$L^0 v(s, x, z) = \left( \frac{\partial}{\partial x} L^0 u(s, x) \right) z$$

(2.13)

$$= 0,$$

for $(s, x, z) \in (t_0, T) \times \mathbb{R}^2$.

We now assume that $E(|v(T, X_T^{t,x}, Z_T^{t,z})|) < \infty$ for all $(t, x) \in [t_0, T] \times \mathbb{R}$. Subject to certain growth bounds applying for the derivative $\frac{\partial h}{\partial z}$ this condition will be verified in Section 1.4 of this chapter.

Consequently, taking expectations of both sides of (2.12) and using (2.8) we have

$$\frac{\partial u}{\partial x}(t, x) = v(t, x, 1)$$

$$= E \left( v \left( T, X_T^{t,x}, Z_T^{t,1} \right) \right)$$

$$= E \left( \frac{\partial u}{\partial x} \left( T, X_T^{t,x} \right) Z_T^{t,1} \right)$$

$$= E \left( \frac{\partial h}{\partial x} \left( X_T^{t,x} \right) Z_T^{t,1} \right)$$

(2.14)

for $(t, x) \in (t_0, T) \times \mathbb{R}.$
Substituting this result into (2.5) we obtain

\[ u(T, X_{T, z}) = u(t_0, z) + \int_{t_0}^{T} E \left( \frac{\partial h}{\partial x} \left( X_{t, z} \right) Z_{T, 1} \mid \mathcal{F}_s \right) b \left( s, X_{s, z} \right) dW_s. \quad (2.15) \]

Applying the Markov property of \( X \) (see the remarks following (1.1)), and (1.9) with \( \tau = T \), this representation becomes

\[ u(T, X_{T, z}) = u(t_0, z) + \int_{t_0}^{T} E \left( \frac{\partial h}{\partial x} \left( X_{T, z} \right) Z_{T, 1} \mid \mathcal{F}_s \right) b \left( s, X_{s, z} \right) dW_s. \quad (2.16) \]

Using the relations \( u(T, X_{T, z}) = h(X_{T, z}) \) and \( u(t_0, z) = E(h(X_{T, z})) \), the latter following by taking expectations of both sides of (2.15), we can express (2.16) equivalently in the form

\[ h \left( X_{T, z} \right) = E \left( h \left( X_{T, z} \right) \right) + \int_{t_0}^{T} E \left( \frac{\partial h}{\partial x} \left( X_{T, z} \right) Z_{T, 1} \mid \mathcal{F}_s \right) b \left( s, X_{s, z} \right) dW_s. \quad (2.17) \]

Thus we have obtained an explicit characterization of the integrand \( \xi \) appearing in (2.3). We see from (2.14) that the variate \( \frac{\partial h}{\partial x} \left( X_{T, z} \right) Z_{T, 1} \) is an unbiased estimator of \( \frac{\partial u}{\partial x} (t, X) \), unlike finite difference approximations. Note also that if Monte Carlo or related sampling methods are used to estimate the price functional \( u \) at the point \( (s, X_{s, z}) \), \( t_0 \leq t \leq T \), the same simulation trajectories for \( X_{T, z} \) can be used, together with new ones for \( Z_{T, 1} \), to approximate \( \frac{\partial u}{\partial x} \) at \( (s, X_{s, z}) \). This procedure is usually more accurate since it is unbiased, and more efficient, since only one simulation run is required, compared to at least two separate simulation runs which are needed for finite difference estimates.

Representations of the type (2.15) - (2.17), under different conditions, have been obtained by Elliott & Kohlmann (1988) and Colwell, Elliott & Kopp (1991) who use the Markov property and the differentiability of solutions of Itô stochastic differential equations with respect to the initial conditions. Broadie & Glasserman (1993) and Carr (1993) also derive explicit representations of hedge ratios in the case where the payoff structure is a standard European call and the diffusion process \( X_{t, z} \) follows a one-dimensional geometric Brownian motion. Our result relies on certain smoothness conditions and an application of the Kolmogorov backward equation. It has the advantage of being very simple and straightforward, and can also be extended to include stopping time boundaries and multidimensional diffusions as will be seen in the next section.

As noted by Colwell, Elliott & Kopp (1991) similar results can be obtained as an application of the Haussmann (1978) integral representation theorem. In fact, the Haussmann's integral representation theorem can be used for certain classes of path dependent securities, where the payoff function \( h \) depends on whole trajectories of the diffusion process \( X_{t, z} \). A proof of Haussmann's theorem which is related to the formula of Clark (1970) using Malliavin calculus techniques is given by Ocone (1984), see also Davis (1980) and Haussmann (1979).
CHAPTER 1. CALCULATION OF HEDGE RATIOS

However use of the Haussmann's integral representation theorem has the disadvantage of being less direct and the conditions of the theorem more difficult to establish. In fact the results presented in this section are sufficient for many types of practical valuation and hedging problems that arise in financial mathematics and can also be extended to a wide class of path dependent options. These include many which are not Frechet differentiable as is required for an application of the Haussmann’s theorem.

1.3 Explicit Hedge Ratios for Multidimensional Diffusions

In practice one is often confronted with the problem of computing hedge ratios for derivative securities associated with multicomponent portfolios. To derive explicit expressions for the hedge ratios in these cases will force us, in this section, to use more complex notations and formulations. However we can still successfully apply the basic ideas of the previous section.

Let \( W = (W^1, \ldots, W^m) \) be an \( m \)-dimensional Brownian motion and \( X^{t_0, \omega} = \{X^{t_0, \omega}_t = (X^{1, t_0, \omega}_t, \ldots, X^{d, t_0, \omega}_t), \ t_0 \leq t \leq T \} \) a \( d \)-dimensional diffusion process which satisfies equation (1.1).

Using the results obtained in Section 1.1 we let \( \tau \) be a stopping time given by (1.4), and corresponding to the continuation region \( \Gamma_0 \) and exercise boundary \( \Gamma_1 \), with \( (t_0, \omega) \in \Gamma_0 \). By taking \( h = \bar{h}, \ u = \bar{u} \) and \( P = \bar{P} \) in (1.6) we assume there is a payoff function \( h : \Gamma_1 \to \mathbb{R} \) and valuation function \( u : \Gamma_0 \cup \Gamma_1 \to \mathbb{R} \) with

\[
\tag{3.1}
 u(t, \omega) = E \left( h \left( \tau, X^{t, \omega}_t \right) \right).
\]

for \( (t, \omega) \in \Gamma_0 \cup \Gamma_1 \). We will say that the function \( f : \Gamma_0 \to \mathbb{R} \) is of class \( C^{1, \ell} \), for integers \( \ell \geq 1 \), if \( f \) is continuously differentiable with respect to \( t \) and \( \ell \)-times continuously differentiable with respect to the spatial variables \( x_1, \ldots, x_d \) on the domain \( \Gamma_0 \). In this section we also assume that the valuation function \( u \) given by (3.1) is of class \( C^{1,3} \) for the domain \( \Gamma_0 \).

As in the case for one-dimensional diffusion processes, see equation (2.5), we can apply multidimensional versions of the Ito formula for semimartingales and the Kolmogorov backward equation to obtain

\[
\tag{3.2}
 u \left( t \wedge \tau, X^{t_0, \omega}_{t \wedge \tau} \right) = u(t_0, \omega) + \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{t \wedge \tau} \frac{\partial u}{\partial x_i} \left( s, X^{t_0, \omega}_s \right) b^{i,j} \left( s, X^{t_0, \omega}_s \right) dW^j_s
\]

for \( t_0 \leq t \leq T \).

Applying (1.12) and (1.13), with \( h = \bar{h} \) and \( u = \bar{u} \), we also have

\[
M_t = E \left( h \left( \tau, X^{t_0, \omega}_t \right) \mid \mathcal{F}_t \right)
\]
Define the operators $L^0$ and $\frac{\partial}{\partial x_p} L^0$, $p \in \{1, \ldots, d\}$, by

$$
L^0 = \frac{\partial}{\partial s} + \sum_{i=1}^{d} a^i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,k=1}^{d} \sum_{j=1}^{m} b^i_j b^k_j \frac{\partial^2}{\partial x_i \partial x_k},
$$

$$
\frac{\partial}{\partial x_p} L^0 = \frac{\partial^2}{\partial x_p \partial s} + \sum_{i=1}^{d} \left( \frac{\partial a^i}{\partial x_p} \frac{\partial}{\partial x_i} + a^i \frac{\partial^2}{\partial x_p \partial x_i} \right) + \frac{1}{2} \sum_{i,k=1}^{d} \sum_{j=1}^{m} \left( \frac{\partial b^i_j}{\partial x_p} b^k_j \frac{\partial^2}{\partial x_i \partial x_k} + b^i_j \frac{\partial b^k_j}{\partial x_p} \frac{\partial^2}{\partial x_i \partial x_k} + b^i_j b^k_j \frac{\partial^3}{\partial x_p \partial x_i \partial x_k} \right),
$$

where the operators $\frac{\partial}{\partial x_p} L^0$, $p \in \{1, \ldots, d\}$ are obtained by calculating the partial derivatives of the operator $L^0$ with respect to $x_p$.

The Kolmogorov backward equation now takes the form

$$
L^0 u(t,x) = 0
$$

for $(t,x) \in \Gamma_0$ with boundary condition $u(\tau,x) = h(\tau,x)$ for $(t,x) \in \Gamma_1$. From this equation we also have

$$
\frac{\partial}{\partial x_p} L^0 u(t,x) = 0
$$

for $(t,x) \in \Gamma_0$, $p \in \{1, \ldots, d\}$.

Let $Z^{s,z} = \{Z^{t,s,z}_i = (Z^{t,1,s,z}_i, \ldots, Z^{t,d,s,z}_i), t_0 \leq s \leq t \leq T\}$ be the solution of the $d^2$-dimensional linearized stochastic differential equation

$$
dZ^{t,i} = \sum_{p=1}^{d} \frac{\partial a^i_k}{\partial x_p} Z^{t,p,i,s,z}_k dt + \sum_{p=1}^{d} \sum_{j=1}^{m} \frac{\partial b^i_j}{\partial x_p} Z^{t,p,i,s,z}_j dW^j_t
$$

for $t_0 \leq s \leq t \leq T$, $k,i \in \{1, \ldots, d\}$. We assume $Z^{s,z}$ starts at time $s$, $t_0 \leq s \leq T$ with initial value $z = (z_{1,1}, \ldots, z_{d,d}) \in \mathbb{R}^{d^2}$.

For $i \in \{1, \ldots, d\}$ we introduce the scalar functions $v_i : \Gamma_0 \cup \Gamma_1 \times \mathbb{R}^{d^2} \rightarrow \mathbb{R}$ defined by

$$
v_i(t,x,z) = \sum_{p=1}^{d} \frac{\partial u}{\partial x_p}(t,x) x_p,i
$$

for $(t,x,z) \in \Gamma_0 \times \mathbb{R}^{d^2}$ with $x = (x_1, \ldots, x_d) \in \mathbb{R}^{d}$, $z = (z_{1,1}, \ldots, z_{d,d}) \in \mathbb{R}^{d^2}$.

Let $i$, with $i \in \{1, \ldots, d\}$ be fixed. Applying the Ito formula for semimartingales to $v_i$ at time $\tau$ and $t \land \tau$, $t_0 \leq t \leq T$, using the system of processes $X = (X^1, \ldots, X^d)$ and $Z^i = (Z^{1,i}, \ldots, Z^{d,i})$ and noting that $\frac{\partial}{\partial x_p,i} v_i(t,x,z) = 0$ for $p,q \in \{1, \ldots, d\}$ and $(t,z) \in \Gamma_0 \times \mathbb{R}^{d^2}$ yields

$$
v_i(\tau,X_{\tau}^{t,z},Z_{\tau}^{t,z}) = v_i(t,z) + \int_{t \land \tau}^{T} \bar{L} \bar{u}_i v_i(s,X_s^{t,z},Z_s^{t,z}) ds
$$
CHAPTER 1. CALCULATION OF HEDGE RATIOS

\[
+ \sum_{j=1}^{m} \int_{t \in \mathcal{A}} L_j^j v_i(s, X^{t,x}_s, Z^{t,z}_s) \, dW^j_s,
\]

where

\[
\begin{align*}
L^0_i &= \frac{\partial}{\partial s} + \sum_{\ell=1}^d a^\ell \frac{\partial}{\partial x_\ell} + \sum_{k=1}^d \left( \sum_{p=1}^d \frac{\partial a^k_s}{\partial x_p} z_{p,i} \right) \frac{\partial}{\partial z_{k,i}} \\
&\quad + \frac{1}{2} \sum_{\ell,k=1}^d b^{\ell,j} b^{k,j} \frac{\partial^2}{\partial x_\ell \partial x_k} + \sum_{\ell,k=1}^d \sum_{j=1}^d b^{\ell,j} \left( \sum_{p=1}^d \frac{\partial b^{k,j}_s}{\partial x_p} z_{p,i} \right) \frac{\partial^2}{\partial x_\ell \partial z_{k,i}},
\end{align*}
\]

for \((t, x, z) \in \Gamma_0 \times \mathbb{R}^d\).

Computing the partial derivatives of \(v_i\) from (3.6), and using (3.4) we have

\[
L^0_i v_i(s, x, z) = \sum_{p=1}^d \frac{\partial^2}{\partial x_p \partial s} u(s, x) z_{p,i} + \sum_{\ell=1}^d a^\ell \sum_{p=1}^d \frac{\partial^2}{\partial x_\ell \partial x_p} u(s, x) z_{p,i} \\
&\quad + \sum_{p=1}^d \sum_{k=1}^d \frac{\partial a^k_s}{\partial x_p} z_{p,i} \frac{\partial}{\partial z_{k,i}} u(s, x) \\
&\quad + \frac{1}{2} \sum_{\ell,k=1}^d b^{\ell,j} b^{k,j} \sum_{p=1}^d \frac{\partial^3}{\partial x_\ell \partial x_p \partial x_k} u(s, x) z_{p,i} \\
&\quad + \sum_{\ell,k=1}^d \sum_{j=1}^d b^{\ell,j} \left( \sum_{p=1}^d \frac{\partial b^{k,j}_s}{\partial x_p} z_{p,i} \right) \frac{\partial^2}{\partial x_\ell \partial x_k} u(s, x) \\
&\quad = \sum_{p=1}^d \left( \frac{\partial}{\partial x_p} L^0_i u(s, x) \right) z_{p,i} \\
&\quad = 0
\]

for \((s, x, z) \in \Gamma_0 \times \mathbb{R}^d\).

Consequently if we take the initial value \(\delta = (\delta_{1,1}, \ldots, \delta_{d,d}) \in \mathbb{R}^d\), where \(\delta_{p,i}\) is the Kronecker delta given by

\[
\delta_{p,i} = \begin{cases} 
1 & : p = i \\
0 & : p \neq i 
\end{cases},
\]

for \(p, i \in \{1, \ldots, d\}\), then taking expectation of both sides of (3.7) yields

\[
\frac{\partial}{\partial x_i} u(t, x) = v_i(t, x, \xi) = E \left( v_i \left( \tau, X^{t,x}_\tau, Z^{t,z}_\tau \right) \right)
\]
1.3. HEDGE RATIOS FOR MULTIDIMENSIONAL DIFFUSIONS

\[ E \left( \sum_{p=1}^{d} \frac{\partial h}{\partial x_p} \left( \tau, X^{t_0, \bar{Z}} \right) Z_{p,i,t,\delta} \right) \]

\[ = E \left( \sum_{p=1}^{d} \frac{\partial h}{\partial x_p} \left( \tau, X^{t_0, \bar{Z}} \right) Z_{p,i,t,\delta} \right) \quad (3.9) \]

for \((t, \bar{Z}) \in \Gamma_0\) and \(i \in \{1, \ldots, d\} \).

We assume \(E(\nu_i(\tau, X^{t_0, \bar{Z}}, Z^{t_0, \delta})) < \infty\) for \(i \in \{1, \ldots, d\}\) so that the expectation in (3.9) is well-defined. As in the one-dimensional case and subject to certain growth bounds applying for the partial derivatives \(\frac{\partial h}{\partial x_1}, \ldots, \frac{\partial h}{\partial x_d}\), this condition will be verified in Section 1.4 of this chapter.

This expression for \(\frac{\partial}{\partial x_i} u(t, \bar{Z})\) can be substituted into (3.2) yielding

\[ u \left( \tau, X^{t_0, \bar{Z}} \right) = u(t_0, \bar{Z}) + \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{\tau} \gamma^{i,j} (s, X^{t_0, \bar{Z}}) dW_s, \quad (3.10) \]

where

\[ \gamma^{i,j} = E \left( \sum_{p=1}^{d} \frac{\partial h}{\partial x_p} \left( \tau, X^{t_0, \bar{Z}} \right) Z_{p,i,s,\delta} \right). \]

Using equation (1.9) and the Markov property we can write

\[ E \left( \sum_{p=1}^{d} \frac{\partial h}{\partial x_p} \left( \tau, X^{t_0, \bar{Z}} \right) Z_{p,i,s,\delta} \right) = E \left( \sum_{p=1}^{d} \frac{\partial h}{\partial x_p} \left( \tau, X^{t_0, \bar{Z}} \right) Z_{p,i,s,\delta} \mid \mathcal{F}_s \right). \]

Combining this with the representation (3.10) we obtain

\[ u \left( \tau, X^{t_0, \bar{Z}} \right) = u(t_0, \bar{Z}) + \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{\tau} \gamma^{i,j} (s, X^{t_0, \bar{Z}}) dW_s, \quad (3.11) \]

where

\[ \gamma^{i,j} = E \left( \sum_{p=1}^{d} \frac{\partial h}{\partial x_p} \left( \tau, X^{t_0, \bar{Z}} \right) Z_{p,i,s,\delta} \mid \mathcal{F}_s \right). \]

Taking expectations of both sides of (3.11) and using the boundary condition \(u(\tau, X^{t_0, \bar{Z}}) = h(\tau, X^{t_0, \bar{Z}})\) we can infer that \(u(t_0, \bar{Z}) = E(h(\tau, X^{t_0, \bar{Z}}))\). Consequently this representation becomes

\[ h \left( \tau, X^{t_0, \bar{Z}} \right) = E \left( h \left( \tau, X^{t_0, \bar{Z}} \right) \right) + \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{\tau} \gamma^{i,j} (s, X^{t_0, \bar{Z}}) dW_s, \quad (3.12) \]

where \(\gamma^{i,j}\) is as given in (3.10) or (3.11). The equations (3.11) and (3.12) should be compared to (2.16) and (2.17) for one-dimensional diffusion processes.

We remark that the formulas (3.11) and (3.12) can be expressed using matrix notation and appear in this form under stronger assumption in Colwell, Elliott & Kopp (1991) and Ocone (1984). We have used the component form because these components are involved explicitly in our proof of the result and because, for practical applications, all of the components need to be computed separately.
1.4 Hedge Ratios for Non-Smooth Payoffs

In this section we consider the important problem, both theoretically and practically, of extending the representation results obtained in the previous section to non-smooth payoff functions. This extension, for example, is required even for standard derivative securities such as the well-known European call option. In most of the known or available literature this problem is overlooked or just neglected.

As in the previous section we let \( W = (W^1, \ldots, W^m) \) be an \( m \)-dimensional Brownian motion and \( X^0 = (X^0_1, \ldots, X^0_d, \ell \), \( t_0 \leq t \leq T \) \) a \( d \)-dimensional diffusion process which satisfies (1.1).

We assume that the drift and diffusion coefficients of (1.1) have linear growth and are Lipschitz continuous so that in particular, see for example Kloeden & Platen (1992), Exercise 4.5.5 and Section 4.8,

\[
E \left( \sup_{t_0 \leq s \leq T} \| X^0_s \| \right) < K_1 < \infty
\]

for some constant \( K_1 \in \mathbb{R}^+ \), where \( \mathbb{R}^+ = \{ r \in \mathbb{R} : r > 0 \} \).

We also assume that the drift and diffusion coefficients of (3.5) have linear growth and are Lipschitz continuous so that using the same result in Kloeden & Platen (1992) there is a constant \( K_4(s) \in \mathbb{R}^+ \) which may depend on \( s \) with

\[
E \left( \sup_{s \leq T} |Z^p, x, s, \delta| \right) < K_2(s) < \infty
\]

for \( 1 \leq p, i \leq d \).

Let \( \tau \) be a stopping time given by (1.4) with continuation region \( \Gamma_0 \) and exercise boundary \( \Gamma_1 \). Consider a valuation function \( u : \Gamma_0 \cup \Gamma_1 \to \mathbb{R} \) of the form (3.1) with payoff function \( h : \Gamma_1 \to \mathbb{R} \). The following conditions will be used in the statement of the main theorem appearing in this section. Here \( \mathbb{N} \) denotes the set \( \{1, 2, \ldots \} \) of natural numbers.

**A1** There exists a sequence of functions \( h_n : \Gamma_1 \to \mathbb{R}, n \in \mathbb{N} \) of class \( C^{1, \ell} \), \( \ell \geq 3 \) such that

(a) for each \( (t, x) \in \Gamma_1 \)

\[
\lim_{n \to \infty} h_n(t, x) = h(t, x),
\]

(b) and for each \( (t, x) \in \Gamma_1, i \in \{1, \ldots, d\} \)

\[
\lim_{n \to \infty} \frac{\partial h_n}{\partial x_i}(t, x) = g_i(t, x)
\]

for some set of functions \( g_i : \Gamma_1 \to \mathbb{R} \).

**A2** (a) The functions \( h_n \) satisfy a uniform linear growth bound of the form

\[
|h_n(t, x)|^2 \leq K_3^2 \left( 1 + \|x\|^2 \right)
\]

for all \( x \in \mathbb{R}^d \) and \( n \in \mathbb{N} \), where \( \|x\|^2 = \sum_{i=1}^{d} x_i^2 \) and \( K_3 < \infty \),
1.4. HEDGE RATIOS FOR NON-SMOOTH PAYOFFS

(b) and the functions \( \frac{\partial h_n}{\partial x_i}, \ i \in \{1, \ldots, d\} \) satisfy a uniform linear growth bound of the form

\[
\sum_{i=1}^{d} \left| \frac{\partial h_n}{\partial x_i}(t, x) \right|^2 \leq K_4^2 \left( 1 + \|x\|^2 \right)
\]

for all \( x \in \mathbb{R}^d, \ n \in \mathbb{N} \), where \( \|x\| \) is as given in A2(a) above and \( K_4 < \infty \).

**Theorem 1.4.1** Suppose the valuation function \( u : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R} \) is defined according to (3.1), conditions A1 and A2 hold for the payoff function \( h \), and the random variables \( h_n(X_{T_s}^{t_0, x}), \ n \in \mathbb{N} \), can be represented in the form

\[
h_n(\tau, X_{\tau_s}^{t_0, x}) = E\left( h_n(\tau, X_{T_s}^{t_0, x}) \right) + \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{\tau} \gamma_{n,s}^{i,j} b^{i,j}(s, X_{s}^{t_0, x}) \, dw_s^{j}, \quad (4.3)
\]

where

\[
\gamma_{n,s}^{i,j} = E\left( \sum_{p=1}^{d} \frac{\partial h_n}{\partial x_p}(X_{T_s}^{t_0, x}) \, Z_{p}^{i,s, \tilde{\xi}} \bigg| \mathcal{F}_s \right).
\]

Then for \( (t_0, x) \in \Gamma_0 \), \( u \) admits the Ito integral representation

\[
u(\tau, X_{\tau_s}^{t_0, x}) = u(t_0, x) + \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{\tau} \gamma_{s}^{i,j} b^{i,j}(s, X_{s}^{t_0, x}) \, dw_s^{j}, \quad (4.4)
\]

or equivalently using A1(b),

\[
h(\tau, X_{\tau_s}^{t_0, x}) = E\left( h(\tau, X_{T_s}^{t_0, x}) \right) + \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{\tau} \gamma_{s}^{i,j} b^{i,j}(s, X_{s}^{t_0, x}) \, dw_s^{j}, \quad (4.5)
\]

where

\[
\gamma_{s}^{i,j} = E\left( \sum_{p=1}^{d} g_{p}(\tau, X_{\tau_s}^{t_0, x}) \, Z_{p}^{i,s, \tilde{\xi}} \bigg| \mathcal{F}_s \right)
\]

and \( Z_{T_s}^{i,s, \tilde{\xi}} \) is the unique solution of the stochastic differential equation (3.5) with initial value \( \tilde{\xi} \) at time \( s, \ t_0 \leq s \leq T \), as given by (3.8).

The above theorem has considerable practical and theoretical value as it allows, under general conditions, for the payoff structure of a contingent claim to be expressed as a stochastic integral. Furthermore, it provides explicit functionals for the corresponding hedge ratios which is extremely valuable because it enables these hedge ratios as well as prices to be accurately computed.

We will establish this result using two lemmas and some general results from the theory of measure and integration including use of the Dominated Convergence Theorem.
CHAPTER 1. CALCULATION OF HEDGE RATIOS

Lemma 1.4.2 Suppose the payoff function $h$ satisfies conditions A1 and A2 and the random variables $h_n(T, X_{\tau T})$, $n \in \mathbb{N}$, can be represented in the form

$$h_n(T, X_{\tau T}) = E(h_n(T, X_{\tau T})) + \sum_{j=1}^{m} \int_{t_0}^{T} \xi_j^n dW_j,$$

where $\xi_n = (\xi_1^n, \ldots, \xi_m^n)$ is a vector of $\mathcal{F}$-predictable processes for each $n \in \mathbb{N}$. Then

$$\lim_{n \to \infty} \left\| h(T, X_{\tau T}) - E(h(T, X_{\tau T})) - \sum_{j=1}^{m} \int_{t_0}^{T} \xi_j^n dW_j \right\|_2 = 0$$

where $\| \cdot \|_2 = \sqrt{E(|\cdot|^2)}$ denotes the norm in the Banach space $L^2(\Omega, \mathcal{F}_T, P)$.

Proof Applying the uniform linear growth bound $A2(a)$ we see that

$$\left| h_n(T, X_{\tau T}) \right|^2 \leq K_3^2 \left( 1 + \| X_{\tau T} \|^2 \right)$$

for all $n \in \mathbb{N}$. This shows that the random variable $|h_n(T, X_{\tau T})|^2$ is dominated by the variate $K_3^2(1 + \| X_{\tau T} \|^2)$ for all $n \in \mathbb{N}$. Now from the growth bound (4.1) we have

$$E(K_3^2 \left( 1 + \| X_{\tau T} \|^2 \right)) < K_3^2(1 + K_1) < \infty.$$

The pointwise convergence of $h_n$ given by condition A1(a) means that

$$\lim_{n \to \infty} |h_n(T, X_{\tau T}) - h(T, X_{\tau T})|^2 = 0 \quad P\text{-a.s.}$$

In fact the convergence here holds for all $\omega \in \Omega$ although we do not require this stronger result. Combining (4.7), (4.8) and (4.9) we can apply a version of the Dominated Convergence Theorem applicable to $L^p$ spaces, $p > 0$, see for example the Corollary to Theorem 2.6.3 in Shiryaev (1984), to obtain

$$\lim_{n \to \infty} E \left( |h_n(T, X_{\tau T}) - h(T, X_{\tau T})|^2 \right) = 0$$

which can be expressed using the $\| \cdot \|_2$ norm of $L^2(\Omega, \mathcal{F}_T, P)$ as

$$\lim_{n \to \infty} \left\| h_n(T, X_{\tau T}) - h(T, X_{\tau T}) \right\|_2 = 0.$$

Furthermore,

$$\left| E(h_n(T, X_{\tau T})) - E(h(T, X_{\tau T})) \right| \leq E\left( \left| h_n(T, X_{\tau T}) - h(T, X_{\tau T}) \right| \right)$$

$$= \left\| h_n(T, X_{\tau T}) - h(T, X_{\tau T}) \right\|_1$$

for all $n \in \mathbb{N}$, where $\| \cdot \|_1 = E(|\cdot|)$ denotes the norm in the Banach space $L^1(\Omega, \mathcal{F}_T, P)$. Since by Hölder’s inequality $\|f\|_1 \leq \|f\|_2$ for any $f \in L^2(\Omega, \mathcal{F}_T, P)$ then from (4.10) and (4.11) we can infer that

$$\lim_{n \to \infty} \left| E(h_n(T, X_{\tau T})) - E(h(T, X_{\tau T})) \right| = 0.$$
1.4. HEDGE RATIOS FOR NON-SMOOTH PAYOFFS

Using (4.10) and (4.12) together with the representation of \( h_n(X_{\tau_0}^{t_0, x}) \) stated in the hypothesis of the lemma gives the required result. □

Lemma 1.4.3 Let \( \{\xi^n_j = \{\xi^n_{j,s}, t_0 \leq s \leq T\}\}_{n \in \mathbb{N}} \) for each \( j \in \{1, \ldots, m\} \) be a sequence of stochastic processes, adapted to the filtration \( \mathcal{F} \) and satisfying the conditions:

(a) The sequence of random variables \( (\sum_{j=1}^{m} \int_{t_0}^{T} \xi^n_{j,s} \, dW_s^j)_{n \in \mathbb{N}} \), is well-defined and forms a Cauchy sequence in \( L^2(\Omega, \mathcal{F}_T, P) \).

(b) For each \( j \in \{1, \ldots, m\} \), there is a stochastic process \( \xi^j = \{\xi^j_s, t_0 \leq s \leq T\} \) with
\[
\lim_{n \to \infty} |\xi^n_{j,s} - \xi^j_s| = 0 \quad \text{P-a.s.} \quad \text{for all } s, t_0 \leq s \leq T.
\]

Then the random variable \( \sum_{j=1}^{m} \int_{t_0}^{T} \xi^j_s \, dW_s^j \) is well-defined and
\[
\lim_{n \to \infty} \left\| \sum_{j=1}^{m} \int_{t_0}^{T} \xi^n_{j,s} \, dW_s^j - \sum_{j=1}^{m} \int_{t_0}^{T} \xi^j_s \, dW_s^j \right\|_2 = 0.
\]

Proof In this proof we will use some general arguments from the theory of measure and integration using the Banach space \( L^{2,*} = L^2([t_0, T] \times \Omega, \mathcal{L} \otimes \mathcal{F}_T, u_L \times P) \) where \( \mathcal{L} \) is the \( \sigma \)-algebra of Lebesgue subsets of \( \mathbb{R} \) and \( u_L \) is the Lebesgue measure. We assume this Banach space is equipped with the norm
\[
\|f\|_2^* = \sqrt{\int_{[t_0,T] \times \Omega} |f|^2 \, du_L \times P}
\]
for any \( f \in L^{2,*} \).

Let us introduce the process \( 1^\tau = \{1^\tau_s = 1_{\{s \leq \tau\}}, t_0 \leq s \leq T\} \). Note that \( 1^\tau \) is \( \mathcal{F} \)-adapted since \( \tau \) is a stopping time. Also \( 1^\tau \) is right continuous and hence applying for example Proposition 1.13 in Karatzas & Shreve (1988) is measurable with respect to the product \( \sigma \)-algebra \( \mathcal{L} \otimes \mathcal{F}_T \).

In what follows we will consider the processes \( \xi^n_j, j \in \{1, \ldots, m\}, n \in \mathbb{N} \) as functions defined on \( [t_0, T] \times \Omega \). By hypothesis the Ito integrals \( \int_{t_0}^{T} \xi^n_{j,s} \, dW^j_s = \int_{t_0}^{T} 1^\tau_s \xi^n_{j,s} \, dW^j_s \), for \( j \in \{1, \ldots, m\}, n \in \mathbb{N} \) are well-defined. Consequently by definition of the Ito integral as the \( L^{2,*} \) limit of appropriately defined step functions the integrands \( 1^\tau \xi^n_j, j \in \{1, \ldots, m\}, n \in \mathbb{N} \), are also \( \mathcal{L} \otimes \mathcal{F}_T \)-measurable. Using this result and Fubini’s Theorem we have
\[
\|1^\tau(\xi^{n_1}_j - \xi^{n_2}_j)\|_2^* = \sqrt{E \left( \int_{t_0}^{T} 1^\tau_s (\xi^{n_1}_j - \xi^{n_2}_j)^2 \, ds \right)}
\]
\[
= \sqrt{E \left( \int_{t_0}^{T} (\xi^{n_1}_j - \xi^{n_2}_j)^2 \, ds \right)}
\]
\[
\leq \sqrt{E \left( \sum_{j=1}^{m} \int_{t_0}^{T} (\xi^{n_1}_j - \xi^{n_2}_j)^2 \, ds \right)}
\]
18  

CHAPTER 1. CALCULATION OF HEDGE RATIOS

\[
= \sqrt{E\left(\left(\sum_{j=1}^{m} \int_{t_0}^{T} (\xi_{n_1,s}^{j} - \xi_{n_2,s}^{j}) dW_s^{j}\right)^2\right)}
\]

\[
= \left|\sum_{j=1}^{m} \int_{t_0}^{T} \xi_{n_1,s}^{j} dW_s^{j} - \sum_{j=1}^{m} \int_{t_0}^{T} \xi_{n_2,s}^{j} dW_s^{j}\right|_2
\]  

(4.13)

for any integers \(n_1, n_2 \in \mathbb{N}\) and \(j \in \{1, \ldots, m\}\), where \(\| \cdot \|_2\) is the \(L^2(\Omega, \mathcal{F}_T, P)\) norm previously defined in the statement of Lemma 1.4.2. By hypothesis the random variables \((\sum_{j=1}^{m} \int_{t_0}^{T} \xi_{n,j}^{j} dW_s^{j})_{n \in \mathbb{N}}\) form a Cauchy sequence in \(L^2(\Omega, \mathcal{F}_T, P)\) and therefore from (4.13) the functions \((1^T \xi_{n}^{j})_{n \in \mathbb{N}}\), for fixed \(j, j \in \{1, \ldots, m\}\) form a Cauchy sequence in \(L^2,*\).

This means that there is a function \(\tilde{\xi}^j : [t_0, T] \times \Omega \rightarrow \mathbb{R}\), \(\tilde{\xi}^j \in L^2,*\), for each \(j \in \{1, \ldots, m\}\), with

\[
\lim_{n \to \infty} \left\|1^T \xi_{n}^{j} - \tilde{\xi}^j\right\|_2 = 0.
\]  

(4.14)

Let \(D = \{(s, \omega) \in [t_0, T] \times \Omega : 1^T_s (\omega) = 1\}\) so that \(D \in \mathcal{L} \otimes \mathcal{F}_T\), since \(1^T\) is measurable with respect to the product \(\sigma\)-algebra \(\mathcal{L} \otimes \mathcal{F}_T\). Using the definition of the \(\| \cdot \|_*^2\) norm we have for \(j \in \{1, \ldots, m\}\).

\[
\left\|1^T \xi_{n}^{j} - 1^T \tilde{\xi}^j\right\|_*^2 = \sqrt{\int_D |\xi_{n}^{j} - \tilde{\xi}^j|^2 \, du_L \times P}
\]

\[
= \sqrt{\int_D |1^T \xi_{n}^{j} - \tilde{\xi}^j|^2 \, du_L \times P}
\]

\[
\leq \sqrt{\int_{[t_0, T] \times \Omega} |1^T \xi_{n}^{j} - \tilde{\xi}^j|^2 \, du_L \times P}
\]

\[
= \left\|1^T \xi_{n}^{j} - \tilde{\xi}^j\right\|_2^2
\]

so that from (4.14)

\[
\lim_{n \to \infty} \left\|1^T \xi_{n}^{j} - 1^T \tilde{\xi}^j\right\|_2^2 = 0.
\]  

(4.15)

Consequently the \(L^2,*\) limit of \(1^T \xi_{n}^{j}, n \in \mathbb{N}, j \in \{1, \ldots, m\}\) as \(n \to \infty\), can be written in the form \(1^T \tilde{\xi}^j\).

We know from the theory of measure and integration, see for example Widom (1969), that any Cauchy sequence in an \(L^2\) Banach space will have a subsequence converging almost everywhere and this limit is the \(L^2\)-limit. For a proof of this result in a probabilistic setting see Shiryaev (1984), Theorems 2.10.2 and 2.10.5. Applying this result and Fubini’s Theorem we see that for each \(j \in \{1, \ldots, m\}\)

\[
E\left(\left|\sum_{s \in \mathbb{N}^+} (\xi_{s,n_1,s}^{j} - \xi_{s,n_2,s}^{j})^2\right|\right) = 0
\]  

(4.16)

for some subsequence \(n_i, i \in \mathbb{N}\) of positive integers, for all \(s \in A\), where \(A \subseteq [t_0, T]\) is some Lebesgue measurable set with \(u_L([t_0, T]\setminus A) = 0\), and \([t_0, T]\setminus A = \{s \in [t_0, T] : s \not\in A\}\).
HEDGE RATIOS FOR NON-SMOOTH PAYOFFS

Equation (4.16) shows that for fixed integer \( j \in \{1, \ldots, m\}\), and \( s \in A\), \((\xi_{n,s}^j, s)_{i \in \mathbb{N}}\) is a Cauchy sequence in \(L^2(\Omega, \mathcal{F}_T, P)\). This means there is a subsequence of \((n_i)_{i \in \mathbb{N}}\), say \((n_k(s))_{k \in \mathbb{N}}\), which may depend on \( s \), with

\[
\lim_{k \to \infty} \left| \int_s^{T_k} \xi_{n_k(s), s}^j \, d\xi_s^j - \int_s^{T_k} \xi_s^j \, d\xi_s^j \right|^2 = 0 \quad P\text{-a.s.} \tag{4.17}
\]

for all \( s \in A \). Therefore from condition (b) in the statement of Lemma 1.4.3 we see that for each \( j \in \{1, \ldots, m\} \)

\[
\int_s^{T_k} \xi_{n_k(s), s}^j \, d\xi_s^j = \int_s^{T_k} \xi_s^j \, d\xi_s^j \quad P\text{-a.s.}
\]

for all \( s \in A \).

Applying this result, Fubini’s Theorem and recalling that \( \| \cdot \|_2 \) denotes the norm of \( L^2(\Omega, \mathcal{F}_T, P) \) we have for any integers \( j \in \{1, \ldots, m\} \) and \( n \in \mathbb{N} \) the relation

\[
\| \int_s^{T_k} \xi_{n,s}^j - \int_s^{T_k} \xi_s^j \|_2^2 = \sqrt{E \left( \int_{t_0}^{T_k} \xi_{n,s}^j - \xi_s^j \right)^2 \, ds}
\]

\[
= \sqrt{E \left( \int_{t_0}^{T_k} \xi_{n,s}^j - \xi_s^j \right)^2 \, ds}
\]

\[
= \sqrt{E \left( \int_{t_0}^{T_k} \xi_{n,s}^j - \xi_s^j \right)^2 \, ds}
\]

\[
= \sqrt{E \left( \int_{t_0}^{T_k} \xi_{n,s}^j - \xi_s^j \right)^2 \, ds}
\]

\[
= \sqrt{E \left( \int_{t_0}^{T_k} \xi_{n,s}^j - \xi_s^j \right)^2 \, ds}
\]

\[
= \| \int_{t_0}^{T_k} \xi_{n,s}^j \, dW_s^j - \int_{t_0}^{T_k} \xi_s^j \, dW_s^j \|_2 . \tag{4.18}
\]

Combining (4.15) and (4.18) which hold for each \( j \in \{1, \ldots, m\} \) we can infer that

\[
\lim_{n \to \infty} \left\| \sum_{j=1}^m \int_{t_0}^T \xi_{n,s}^j \, dW_s^j - \sum_{j=1}^m \int_{t_0}^T \xi_s^j \, dW_s^j \right\|_2 = 0
\]
Proof of Theorem 1.4.1  We will now show that the conditions required for an application of Lemmas 1.4.2 and 1.4.3 can be satisfied for suitable choices of processes \( \xi_n, n \in \mathbb{N} \), and \( \xi \) under the assumptions A1 and A2.

For integers \( j \in \{1, \ldots, m\} \), \( n \in \mathbb{N} \) and \( s \in \mathbb{R} \), \( t_0 \leq s \leq T \) define

\[
\xi_{n,s}^j = \sum_{i=1}^{d} \gamma_i \beta_i^j \left( s, X_t^{t_0, \xi} \right),
\]

(4.19)

\[
\zeta_s^j = \sum_{i=1}^{d} \gamma_i \beta_i^j \left( s, X_t^{t_0, \xi} \right),
\]

(4.20)

where \( \gamma_i \) and \( \gamma_i' \), \( i \in \{1, \ldots, d\} \), \( n \in \mathbb{N} \) are as given in the representations (4.3) and (4.5), respectively.

Substituting (4.19) into (4.6) we obtain the representation (4.3) which is assumed to be true by the hypothesis of Theorem 1.4.1. Also, as given in the statement of Theorem 1.4.1 we assume that the payoff function \( h \) satisfies conditions A1 and A2. Consequently applying the results of Lemma 1.4.2 we see that the random variables \( \sum_{j=1}^{m} \int_{t_0}^{T} \zeta_{n,s}^j \, dW_t^j, n \in \mathbb{N} \), are well-defined and form a Cauchy sequence in \( L^2(\Omega, \mathcal{F}_T, \mathbb{P}) \). This verifies condition (a) in the statement of Lemma 1.4.3.

To establish condition (b) of the lemma we will use the uniform linear growth bound A2(b) and the pointwise limit A1(b). Let \( p, i \in \{1, \ldots, d\} \) and \( s \in \mathbb{R} \), \( t_0 \leq s \leq T \) be fixed. For \( n \in \mathbb{N} \) define

\[
\psi_{n,p,i,s} = \frac{\partial h_n}{\partial x_p} \left( \tau, X_t^{t_0, \xi} \right) Z_t^{p,i,s}.
\]

(4.21)

so that from condition A2(b) we have the inequality

\[
|\psi_{n,p,i,s}| \leq \sqrt{K_4^2 \left( 1 + \|X_t^{t_0, \xi}\|^2 \right) \|Z_t^{p,i,s}\|}
\]

\[
\leq K_4 \left( 1 + \|X_t^{t_0, \xi}\| \right) \|Z_t^{p,i,s}\|
\]

for \( t_0 \leq s \leq T \). If we let \( \Psi_{p,i,s} = K_4 \left( 1 + \|X_t^{t_0, \xi}\| \right) \|Z_t^{p,i,s}\| \), then using Hölder’s and Minkowski’s inequalities, and the bounds (4.1) and (4.2) we have

\[
E(\Psi_{p,i,s}) \leq \left( K_4 \left( 1 + \|X_t^{t_0, \xi}\| \right) \right) \|Z_t^{p,i,s}\|_2
\]

\[
\leq K_4 \left( 1 + \|X_t^{t_0, \xi}\|_2 \right) \|Z_t^{p,i,s}\|_2
\]

\[
< \infty
\]

(4.22)

for \( t_0 \leq s \leq T \). Also the pointwise limit condition A1(b) shows that

\[
\lim_{n \to \infty} \left| \frac{\partial h_n}{\partial x_p} \left( X_t^{t_0, \xi} \right) - g_p \left( X_t^{t_0, \xi} \right) \right| \|Z_t^{p,i,s}\| = 0 \quad P\text{-a.s.}
\]

(4.23)
1.4. HEDGE RATIOS FOR NON-SMOOTH PAYOFFS

for \( t_0 \leq s \leq T \). These results show that the random variable \( \psi_{n,p,i,s} \) is dominated by \( \Psi_{p,i,s} \) which has finite expectation by (4.22) and has an almost sure limit as given by (4.23).

We can now apply the Dominated Convergence Theorem for conditional expectations, see for example Shiryayev (1984) Theorem 2.7.2, yielding

\[
\lim_{n \to \infty} E \left( \left( \frac{\partial h_n}{\partial x_p} \left( X_{t_0}^{s} \right) - g_p \left( X_{t_0}^{s} \right) \right) Z_{t_0}^{p,i,s} \mid \mathcal{F}_s \right) = 0 \quad \text{P-a.s.} \tag{4.24}
\]

for \( t_0 \leq s \leq T \) and \( p, i \in \{1, \ldots, d\} \). From the definition of \( \xi_{n,s}^i \) and \( \xi_s^i \) given by (4.19) and (4.20), respectively, and applying (4.24) we can infer that

\[
\lim_{n \to \infty} |\xi_{n,s}^i - \xi_s^i| = 0 \quad \text{P-a.s.} \tag{4.25}
\]

for each \( j \in \{1, \ldots, m\} \) and \( t_0 \leq s \leq T \). This establishes condition (b) in the statement of Lemma 1.4.3.

Finally we note that since the random variables \( \gamma_{n,s}^i, \ i \in \{1, \ldots, d\}, \ n \in \mathbb{N} \) are expressed as conditional expectations with respect to \( \mathcal{F}_s \), as can be seen from (4.3), they are \( \mathcal{F}_s \)-measurable. Also \( b^{i,j}(s, X_s^{t_0, x}) \), \( i \in \{1, \ldots, d\}, \ j \in \{1, \ldots, m\} \) is \( \mathcal{F}_s \)-measurable since \( X_s^{t_0, x} \) is \( \mathcal{F}_s \)-measurable, and therefore \( \xi_s^i, \ j \in \{1, \ldots, m\}, \ n \in \mathbb{N} \), as given by (4.19) will be adapted to the filtration \( \mathcal{F} \).

Consequently the conditions required for an application of Lemma 1.4.3 are satisfied with \( \xi_s^i \) and \( \xi_s^j \) given by (4.19), (4.20), respectively. Combining the results of Lemmas 1.4.2 and 1.4.3 yields

\[
\left\| u \left( \tau, X_{\tau}^{t_0, x} \right) - u(t_0, x) - \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{\tau} \gamma_{s}^{i} b^{i,j} \left( s, X_{s}^{t_0, x} \right) dW_{s}^{j} \right\|_2 = 0
\]

or equivalently

\[
u \left( \tau, X_{\tau}^{t_0, x} \right) = u(t_0, x) + \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{\tau} \gamma_{s}^{i} b^{i,j} \left( s, X_{s}^{t_0, x} \right) dW_{s}^{j} \quad \text{P-a.s.}
\]

which completes the proof of Theorem 1.4.1.

\[\square\]

We remark that the smoothness condition, where we assume that \( h_n \) is of class \( C^{1,\ell}, \ \ell \geq 3 \) as stated in A1 has not been explicitly used in the proof of Theorem 1.4.1. In practice such a smoothness condition is required for the results of Section 2 and 3 to be applied so that the representation (4.3), used in the statement of Theorem 1.4.1, can be obtained. A method for constructing these approximating functions for one-dimensional diffusion processes and a class of absolutely continuous payoff functions will be considered in the next section. Finally we note that the condition \( E(\sum_{i=1}^{d} \frac{\partial h}{\partial x_p} (\tau, X_{\tau}^{t_0, x}) Z_{\tau}^{p,i,s,d}) < \infty \) for \( t_0 \leq t \leq T \) required in the proof of (2.14) and (3.9) can also be obtained from the inequality (4.22) which depends on the conditions (4.1), (4.2) and A2(b).
1.5 Absolutely Continuous Payoff Functions

In this and the two following sections we will provide examples of applications of Theorem 1.4.1. These examples demonstrate the wide applicability of the theorem but are not needed in the remaining parts of the thesis. The reader who is more interested in practically oriented results could therefore omit these sections and proceed directly to Chapter 2.

Here we will show that conditions A1 and A2 required for an application of Theorem 1.4.1 are satisfied for a class of one-dimensional absolutely continuous functions. This result, together with Theorem 1.4.1, will then be applied to show that these functions admit an Ito integral representation of the form (2.16) or (2.17) for a wide class of one-dimensional diffusion processes $X^{t_0, x}$.

For a one-dimensional payoff function $h$ with $\tau = T$ and $h$ of the form $h(x) = h(T, x)$, for $x \in \mathbb{R}$ these conditions can be simplified as follows:

**A1** There exists a sequence of functions $h_n : \mathbb{R} \to \mathbb{R}$, $n \in \mathbb{N}$, of class $C^\ell$ for some integer $\ell \geq 3$ such that

(a) for each $x \in \mathbb{R}$

\[
\lim_{n \to \infty} h_n(x) = h(x),
\]

(b) and there is a function $g : \mathbb{R} \to \mathbb{R}$ such that for each $x \in \mathbb{R}$

\[
\lim_{n \to \infty} h'_n(x) = g(x).
\]

**A2** The functions $h_n$ and $h'_n$ satisfy uniform linear growth bounds of the form

(a) $|h_n(x)|^2 \leq K_2^2 (1 + |x|^2)$

(b) $|h'_n(x)|^2 \leq K_2^2 (1 + |x|^2)$

for all $x \in \mathbb{R}$ and $n \in \mathbb{N}$ with $K_3, K_4 < \infty$.

Let $h$ be an absolutely continuous function of the form

\[
h(x) = h(0) + \int_0^x g(s) \, ds \tag{5.1}
\]

for $x \in \mathbb{R}$, where both $h$ and $g$ satisfy linear growth bounds of the form

(a) $|h(x)|^2 \leq A_0 (1 + |x|^2)$

(b) $|g(x)|^2 \leq B_0 (1 + |x|^2)$

for all $x \in \mathbb{R}$ with $A_0, B_0 < \infty$.

We assume that $g$ is continuous except possibly at a finite number of points $x_1, \ldots, x_N$ with $x_1 < x_2 < \ldots < x_N$ and that the right hand limits are well-defined and satisfy

\[
g^+(x_i) = \lim_{\delta \to 0^+} g(x_i + \delta) = g(x_i)
\]
1.5. ABSOLUTELY CONTINUOUS PAYOFF FUNCTIONS

for each \( i \in \{1, \ldots, N\} \). These conditions and (5.1) show that \( h'(x) = g(x) \) for all \( x \in \mathbb{R}\backslash\{x_1, \ldots, x_N\} \). Note that standard payoff functions such as European calls or puts, where \( h(x) = (x-K)^+ \) or \( h(x) = (K-x)^+ \), respectively, are absolutely continuous functions of the form (5.1).

For a continuous function \( f : \mathbb{R} \to \mathbb{R} \) define the function \( I_{n,m}(f) : \mathbb{R} \to \mathbb{R} \), for integers \( n \geq 1, m \geq 0 \) iteratively as follows

\[
I_{n,0}(f)(x) = f(x)
\]

\[
I_{n,m+1}(f)(x) = n \int_{x}^{x + \frac{1}{n}} I_{n,m}(f)(s) \, ds
\]

for \( x \in \mathbb{R} \), where \( I_{n,m} \) can be interpreted as a smoothing operator defined on the set of real-valued continuous functions with domain \( \mathbb{R} \).

We will show that \( h \) satisfies conditions \( A1^* \) and \( A2^* \) using the approximating function \( I_{n,m}(h), n \in \mathbb{N} \) for any integer \( m \geq 3 \).

The Lebesgue integral in (5.4) is used to ensure that the functions \( I_{n,m+1}(f), n \in \mathbb{N} \) are of class \( C^{m+1} \) if \( I_{n,m}(f) \) is of class \( C^{m} \) for integers \( m \geq 0 \). In fact computing the derivative of \( I_{n,m}(f) \) using (5.4) we see that

\[
(I_{n,m+1}(f))'(x) = n \left( I_{n,m}(f) \left( x + \frac{1}{n} \right) - I_{n,m}(f)(x) \right)
\]

for all \( x \in \mathbb{R} \). In particular, since the function \( h \) is absolutely continuous, then \( I_{n,m}(h) \) will be of class \( C^m \) for all integers \( n \geq 1, m \geq 1 \).

We will now show that

\[
|I_{n,m}(h)(x)| \leq A_m \sqrt{1 + |x|^2}
\]

for all \( x \in \mathbb{R} \), and integers \( n \geq 1, m \geq 0 \) where \( A_m < \infty \) is some constant which depends on \( m \).

Applying the inequality \((a + \frac{1}{n})^2 \leq 2(a^2 + \frac{1}{n}) \leq 2(a^2 + 1)\), for \( a \in \mathbb{R} \), \( n \in \mathbb{N} \) and assuming (5.6) holds for some fixed integer \( m \geq 0 \), for all integers \( n \geq 1 \), we have

\[
|I_{n,m+1}(h)(x)| \leq n \int_{x}^{x + \frac{1}{n}} |I_{n,m}(h)(s)| \, ds
\]

\[
\leq n \int_{x}^{x + \frac{1}{n}} A_m \sqrt{1 + |s|^2} \, ds
\]

\[
\leq A_m \sqrt{1 + \left( |x| + \frac{1}{n} \right)^2}
\]

\[
\leq A_m \sqrt{3 + 2|x|^2}
\]

\[
\leq A_{m+1} \sqrt{1 + |x|^2},
\]

(5.7)
where $A_{m+1} = \sqrt{3} A_m$. Since $I_{n,0}(h) = h$, condition (5.2) shows that (5.6) holds for $m = 0$. Consequently, by an induction argument using (5.7), we see that (5.6) is valid for all integers $n \geq 1$, $m \geq 0$. This shows that for any integer $m \geq 0$, the functions $I_{n,m}(h)$, $n \in \mathbb{N}$ satisfy $A2^*(a)$.

Since $I_{n,m}(h)$, for integers $n \geq 1$, $m \geq 1$ is of class $C^m$, as previously noted, and $I_{n,0}(h) = h$ is continuous, the function $I_{n,m}(h)$ is continuous for all integers $n \geq 1$, $m \geq 0$. From definition (5.4) and the Mean Value Theorem we can show that for any $x \in \mathbb{R}$ and integers $n \geq 1$, $m \geq 0$

$$I_{n,m+1}(h)(x) = I_{n,m}(h)(\eta),$$

for some $\eta \in [x, x + \frac{1}{n}]$. Applying this result to the functions $I_{n,m}(h), \ldots, I_{n,1}(h)$ we can infer that

$$I_{n,m}(h)(x) = I_{n,0}(h)(\eta_1) = h(\eta_1)$$

for some $\eta_1 \in [x, x + \frac{m}{n}]$. Taking the limit as $n \to \infty$ yields

$$\lim_{n \to \infty} I_{n,m}(h)(x) = h(x)$$

(5.8)

for any integer $m \geq 1$ and $x \in \mathbb{R}$. Consequently for any integer $m \geq 1$ condition $A1^*(a)$ is satisfied with the approximating functions $I_{n,m}(h)$, $n \in \mathbb{N}$.

To verify condition $A2^*(b)$ we will show that

$$|(I_{n,m}(h))'(x)| \leq B_m \sqrt{1 + |x|^2}$$

(5.9)

for all $x \in \mathbb{R}$, and integers $n \geq 1$, $m \geq 0$, where $B_m < \infty$ is a constant which depends on $m$.

By (5.5) we can write

$$(I_{n,m+1}(h))'(x) = n \left(I_{n,m}(h) \left(x + \frac{1}{n}\right) - I_{n,m}(h)(x)\right)$$

$$= n \int_{x}^{x + \frac{1}{n}} (I_{n,m}(h))'(s) \, ds,$$  

(5.10)

for all $x \in \mathbb{R}$ and integers $n \geq 1$, $m \geq 1$. Since $I_{n,0}(h) = h$ we also have from (5.5) and (5.1) the relation

$$(I_{n,1}(h))'(x) = n \left(h \left(x + \frac{1}{n}\right) - h(x)\right)$$

$$= n \int_{x}^{x + \frac{1}{n}} g(s) \, ds.$$  

(5.11)

If we assume that (5.9) holds for some integer $m \geq 1$ and all $n \in \mathbb{N}$ then using the inequality $(a + \frac{1}{n})^2 \leq 2 \left(a^2 + 1\right)$, we have from (5.10) and (5.11) the inequalities

$$|(I_{n,m+1}(h))'(x)| \leq n \int_{x}^{x + \frac{1}{n}} B_m \sqrt{1 + |s|^2} \, ds$$
for all \( x \in \mathbb{R} \) and integers \( n \geq 1 \), where \( B_{m+1} = \sqrt{3} B_m \). If we use equation (5.11) and apply the inequality (5.3) a similar proof shows that

\[
(I_{n,1}(h))'(x) \leq B_1 \sqrt{1 + |x|^2},
\]

where \( B_1 = \sqrt{3} B_0 \). This means that (5.9) is valid for \( m = 1 \) and hence by an induction argument using (5.12) we see that (5.9) holds for all integers \( m \geq 0 \). This means that the sequence of functions \( I_{n,m}(h) \), \( n \in \mathbb{N} \), satisfy conditions A2* (b) for any integer \( m \geq 0 \).

To find the pointwise limits of \( (I_{n,m}(h))' \), as \( n \to \infty \) for a fixed integer \( m \geq 0 \) suppose \( x \not\in \{x_1, \ldots, x_N\} \). Let \( \alpha = \min_{1 \leq i \leq N} \{|x - x_i|\} \). Then for all integers \( n > \frac{m}{\alpha} \), \( g \) is continuous on the interval \([x, x + \frac{m}{n}]\) and \( (I_{n,m}(h))' \), for integers \( m \geq 1 \), is of class \( C^{m-1} \), as previously noted and hence is continuous on \( \mathbb{R} \). From this fact, (5.10) and using the Mean Value Theorem we have for any integers \( n \geq 1 \), \( m \geq 1 \) the relation

\[
(I_{n,m+1}(h))'(x) = (I_{n,m}(h))'(\eta)
\]

for some \( \eta \in [x, x + \frac{1}{n}] \). A similar argument and (5.11) shows that

\[
I_{n,1}(h)'(x) = g(\eta_1)
\]

for some \( \eta_1 \in [x, x + \frac{1}{n}] \). Applying these results to the functions \( (I_{n,m}(h))', \ldots, (I_{n,1}(h))' \) we can infer that for any integer \( n \geq 1 \), \( m \geq 1 \),

\[
(I_{n,m}(h))'(x) = g(\eta_2)
\]

for some \( \eta_2 \in [x, x + \frac{m}{n}] \). This shows that

\[
\lim_{n \to \infty} (I_{n,m}(h))'(x) = g(x)
\]

(5.13)

for all integers \( m \geq 1 \). If \( x \in \{x_1, \ldots, x_N\} \), then for all integers \( n > \frac{m}{\beta} \), where \( \beta = \min_{1 \leq i \leq N-1} \{|x_{i+1} - x_i|\} \), \( g \) is continuous on the interval \([x, x + \frac{m}{n}]\) and as noted above \( (I_{n,m}(h))' \), \( m \geq 1 \) is continuous on \( \mathbb{R} \) and therefore the limit (5.13) also applies. It follows that the approximating functions \( I_{n,m}(h) \), \( n \in \mathbb{N} \), for any integer \( m \geq 0 \) satisfies condition A1* (b).

Summarizing these results, we have shown that the conditions A1* and A2* hold for \( h \) using the approximating functions \( I_{n,m}(h) \), \( n \in \mathbb{N} \), for any integer \( m \geq 3 \).

Let \( X^{t_0,x} \) be a one-dimensional diffusion process satisfying (2.1) with drift \( a \) and diffusion \( b \) coefficients of class \( C^{m,m} \), \( m \geq 1 \), with uniformly bounded derivatives. From
CHAPTER 1. CALCULATION OF HEDGE RATIOS

Theorem 4.8.6 in Kloeden & Platen (1992), due to Mikulevicius (1983), we know that the valuation function \( u_{n,m} : [t_0, T] \times \mathbb{R} \to \mathbb{R} \) given by

\[
 u_{n,m}(t, x) = E(I_{n,m}(h)(X_t^{t_0, x})) \tag{5.14}
\]

for \((t, x) \in [t_0, T] \times \mathbb{R} \) and \( n \in \mathbb{N} \) is of class \( C^{1,m} \) for all even integers \( m \geq 2 \). In particular for \( m \geq 4 \), \( u_{n,m} \) will be of class \( C^{1,4} \) and hence, using the results in Section 1.2 will admit an Ito integral representation of the form (2.16) or (2.17). If (4.1) holds for the diffusion process \( X_t^{t_0, x} \), and the linearized process \( Z^{s,1} \), \( t_0 \leq s \leq T \), given by (2.10) satisfies (4.2) then applying Theorem 1.4.1, we see that the valuation function \( u : [t_0, T] \times \mathbb{R} \to \mathbb{R} \) given by (2.4) will also admit a representation of the form (2.16) or (2.17). Thus we have found an Ito integral representation of the random variable \( u(T, x) = h(X_T) \) with explicit expressions for the integrand even in the case where the derivative of \( h \) is discontinuous at a finite number of points.

This result, which includes a wide class of non-smooth payoff functions, justifies the effort that was needed to deal with the technical difficulties encountered in the previous section.

1.6 Maximum of Several Assets

In this section we will apply Theorem 1.4.1 again to obtain explicit representations for non-smooth functionals of the maximum of several assets as for example are used in basket options. This is a challenging problem and will lead to some complex notations and formulations. However the representation that will be derived is of considerable practical value and again illustrates the power and scope of Theorem 1.4.1.

Let \( h : \mathbb{R} \to \mathbb{R} \) be a payoff function which satisfies conditions A1* and A2* given in Section 1.5. We assume \( I_{n,4}(h) \), for \( n \in \mathbb{N} \) are the approximating functions of class \( C^{1,4} \) as given by (5.4). We now consider payoff structures of the form

\[
 h \left( \max \left( X_T^{t_0, x} \right) \right) = h \left( \max \left( X_T^{t_0, x, 1}, \ldots, X_T^{t_0, x, d} \right) \right) \tag{6.1}
\]

where \( X_T^{t_0, x} = \{X_t^{t_0, x}, t_0 \leq t \leq T\} \) is the solution of the \( d \)-dimensional stochastic differential equation (1.1) starting at time \( t_0 \) with initial value \( x = (x_1, \ldots, x_d) \in \mathbb{R}^d \) and \( X_T^{t_0, x, 1}, \ldots, X_T^{t_0, x, d} \) are the components of \( X_T^{t_0, x} \).

In this section we assume the corresponding valuation function \( u : [t_0, T] \times \mathbb{R}^d \to \mathbb{R} \) is given by the time-independent version of (1.6) with \( \tau = T \), \( u = \tilde{u}, h = \tilde{h} \) and \( \tilde{P} = P \) so that

\[
 u(t, x_1, \ldots, x_d) = E \left( h \left( \max \left( X_T^{t, x_1}, \ldots, X_T^{t, x_d} \right) \right) \right)
\]

for \( x = (x_1, \ldots, x_d) \in \mathbb{R}^d \) and \( t_0 \leq t \leq T \).

Our aim will be to find an explicit representation of \( u \) of the form (4.4) or (4.5) using Theorem 1.4.1. We do this by constructing an appropriate sequence of approximating functions for the payoff functional \( h \circ \max : \mathbb{R}^d \to \mathbb{R} \) corresponding to \( u \) given by \( h \circ \max(x) = h(\max(x)) \), for \( x \in \mathbb{R}^d \).
1.6. MAXIMUM OF SEVERAL ASSETS

Let \( n \in \mathbb{N} \) and define \( f_n : \mathbb{R}^d \to \mathbb{R} \) by

\[
f_n(x) = (5n)^4 \int_{C_n} \int_{C_n} \int_{C_n} \int_{C_n} \max \left( x - \alpha^{(n)} - \sum_{k=1}^{4} y^{(k)} \right) dy^{(1)} \cdots dy^{(4)}
\]

for \( x \in \mathbb{R}^d \) where \( \alpha^{(n)} = (\frac{1}{n}, \frac{2}{n}, \ldots, \frac{d}{n}) \in \mathbb{R}^d \) and \( C_n \) is the \( d \)-dimensional cube of length \( \frac{1}{5n} \) given by

\[
C_n = \left\{ a \in \mathbb{R}^d : a = (a_1, \ldots, a_d), 0 \leq a_i \leq \frac{1}{5n} \text{ for } i \in \{1, \ldots, d\} \right\}.
\]

If we take \( \tau = T \) we can remove the time parameter \( t \) from the formulation of conditions A1 and A2 given in Section 1.4. For example condition A1(a) becomes

A1\(^*(a)\) For each \( x \in \mathbb{R}^d \)

\[
\lim_{n \to \infty} h_n(x) = h(x).
\]

We refer to this representation of conditions A1 and A2 as the time-independent formulation with \( \tau = T \). Note that \( \alpha^{(n)} \in \mathbb{R}^d \) has components \( \alpha_i^{(n)} = \frac{i}{n} \) for \( i \in \{1, \ldots, d\} \). We will show that the time-independent formulation of conditions A1 and A2 with \( \tau = T \) hold for the payoff function \( h \circ \max \) using the approximating functions \( h_n = I_{n,A}(h) \circ f_n \), \( n \in \mathbb{N} \) given by \( I_{n,A}(h) \circ f_n(x) = I_{n,A}(h)(f_n(x)) \) for \( x \in \mathbb{R}^d \).

From the definition of the functions \( f_n \), given by (6.2) we see that

\[
|f_n(x)| \leq (5n)^4 \int_{C_n} \int_{C_n} \int_{C_n} \int_{C_n} \max \left( x - \alpha^{(n)} - \sum_{k=1}^{4} y^{(k)} \right) dy^{(1)} \cdots dy^{(4)}
\]

\[
\leq (5n)^4 \int_{C_n} \int_{C_n} \int_{C_n} \int_{C_n} \left( |\max(x)| + \frac{d+1}{n} \right) dy^{(1)} \cdots dy^{(4)}
\]

\[
\leq \|x\| + \frac{d+1}{n}
\]

(6.3)

for all \( x \in \mathbb{R}^d \) and \( n \in \mathbb{N} \). Consequently from the linear growth bound that applies for \( I_{n,A}(h) \) given by condition A2\(^*(a)\) and the inequality \((a+b)^2 \leq 2(a^2 + b^2)\) we have

\[
|I_{n,A}(h) \circ f_n(x)|^2 = |I_{n,A}(h)(f_n(x))|^2
\]

\[
\leq K_3^2 (1 + |f_n(x)|^2)
\]

\[
\leq K_3^2 \left( 1 + \left( \|x\| + \frac{d+1}{n} \right)^2 \right)
\]

\[
\leq K_3^2 \left( 1 + 2 \left( \frac{d+1}{n} \right)^2 + 2 \|x\|^2 \right)
\]

\[
\leq 2K_3^2 \left( 1 + (d+1)^2 \right) (1 + \|x\|^2)
\]

\[
\leq K_3^2 (1 + \|x\|^2)
\]

(6.4)
for all $x \in \mathbb{R}^d$, and $n \in \mathbb{N}$, where $K_2^2 = 2 K_2^3 (1 + (d + 1)^2)$ and $K_3 = A_4$ as given by (5.6). This result shows that condition A2(a) holds in the time-independent case with $\tau = T$ for the approximating functions $h_n = I_{n,\delta}(h) \circ f_n, n \in \mathbb{N}$.

From the Mean Value Theorem and the definition of $f_n$ given by (6.2) we know that for any $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$

$$f_n(x) = \max \left( x - \alpha^{(n)} - 4 \beta^{(n)} \right)$$

for some $\beta^{(n)} \in C_n$. This shows that

$$\lim_{n \to \infty} f_n(x) = \max(x)$$

for all $x \in \mathbb{R}^d$. By condition A1*(a) (see (5.8)) this means that

$$\lim_{n \to \infty} I_{n,\delta}(h) \circ f_n(x) = \lim_{n \to \infty} I_{n,\delta}(h)(f_n(x)) = h(\max(x)) = h \circ \max(x)$$

for all $x \in \mathbb{R}^d$. This validates condition A1(a) in the time-independent case with $\tau = T$ for the payoff function $h \circ \max$ using the approximating functions $h_n = I_{n,\delta}(h) \circ f_n$.

From definition (6.2) it is a straightforward but tedious calculation to show that the functions $f_n, n \in \mathbb{N}$ are of class $C^4$. To calculate pointwise limits of the partial derivatives of $f_n$ as $n \to \infty$ we define $\pi : \mathbb{R}^d \to \{1, \ldots, d\}$ by the equation

$$\pi(x) = \min_{1 \leq i \leq d} \{ i : x_i = \max(x) \}.$$  

In addition, for $\zeta > 0$ and $i \in \{1, \ldots, d\}$ define $\zeta^{(i)} = (\zeta^{(i)}_1, \ldots, \zeta^{(i)}_d) \in \mathbb{R}^d$ by the rule $\zeta^{(i)}_k = 0$ for $i \neq k$ and $\zeta^{(i)}_k = \zeta$ for $i = k$.

Let $i \in \{1, \ldots, d\}, n \in \mathbb{N}$ and $x \in \mathbb{R}^d$ be fixed with $y^{(1)}, \ldots, y^{(4)} \in C_n$. If $i \neq \pi(x)$ then either $x_i < \max(x)$ or there is an integer $j, j < i$ with $x_j = x_i = \max(x)$.

If $x_j < \max(x)$ we choose some integer $j$ with $x_j = \max(x)$ and $x_i < x_j$. Then for $0 \leq \zeta \leq \frac{x_j - x_i}{2}$ and $n > 2(d+1) \frac{x_j - x_i}{x_j - x_i}$, so that $\frac{d+1}{n} < \frac{x_j - x_i}{2}$, we can show

$$x_i + \zeta + \left( \alpha^{(n)}_j - \alpha^{(n)}_i \right) + \frac{4}{k=1} \left( y^{(k)}_j - y^{(k)}_i \right) \leq x_j$$

or

$$x_i + \zeta - \alpha^{(n)}_i - \frac{4}{k=1} y^{(k)}_i \leq x_j - \alpha^{(n)}_j - \frac{4}{k=1} y^{(k)}_j$$

for all $y^{(1)}, \ldots, y^{(4)} \in C_n$. This means that for $0 \leq \zeta \leq \frac{x_j - x_i}{2}$ and $n > 2(d+1) \frac{x_j - x_i}{x_j - x_i}$ we have the relation

$$\max \left( x + \zeta^{(i)} - \alpha^{(n)} - \frac{4}{k=1} y^{(k)} \right) = \max \left( x - \alpha^{(n)} - \frac{4}{k=1} y^{(k)} \right)$$

(6.8)
1.6. MAXIMUM OF SEVERAL ASSETS

for all \(y(1), \ldots, y(4) \in C_n\).

If \(x_j = x_i = \max(x), j < i\), then for \(0 < \zeta \leq \frac{1}{5n}\) and using the identity \(\alpha_i^{(n)} - \alpha_j^{(n)} = \frac{i-j}{n}\) we can show

\[
x_i + \zeta + \sum_{k=1}^{4} (y_j^{(k)} - y_i^{(k)}) \leq x_j + (\alpha_i^{(n)} - \alpha_j^{(n)})
\]

or

\[
x_i + \zeta - \alpha_i^{(n)} - \sum_{k=1}^{4} y_i^{(k)} \leq x_j - \alpha_j^{(n)} - \sum_{k=1}^{4} y_j^{(k)}
\]

for all \(y(1), \ldots, y(4) \in C_n\) and \(n \in \mathbb{N}\). Again this means that (6.8) holds for all \(0 < \zeta \leq \frac{1}{5n}\) and \(y(1), \ldots, y(4) \in C_n\).

Combining these two results and the definition of \(f_n\) given by (6.2) we have, in the case where \(i \neq \pi(x)\), and for sufficiently large \(n\) and sufficiently small \(\zeta\), the identity \(f_n(x + \zeta^{(i)}) - f_n(x) = 0\). Letting \(\zeta \to 0\) we therefore obtain

\[
\lim_{n \to \infty} \frac{\partial f_n}{\partial x_i}(x) = 0. \tag{6.9}
\]

If \(i = \pi(x)\) then \(x_i = \max(x)\) and \(x_j < x_i\) for \(j < i, j \in \{1, \ldots, d\}\). Thus for \(j < i\) and \(n > \frac{d+1}{\min_{j<i} |x_i - x_j|}\), so that \(\frac{d+1}{n} < \min_{j<i} |x_i - x_j|\) we have the inequality

\[
x_j + (\alpha_i^{(n)} - \alpha_j^{(n)}) + \sum_{k=1}^{4} (y_i^{(k)} - y_j^{(k)}) \leq x_j + \frac{d+1}{n} \leq x_i,
\]

for all \(y(1), \ldots, y(4) \in C_n\).

If \(j > i, j \in \{1, \ldots, d\}\) and \(n > \frac{d+1}{\min_{j<i} |x_i - x_j|}\), then \(\alpha_i^{(n)} - \alpha_j^{(n)} = \frac{i-j}{n} \leq -\frac{1}{n}\) and therefore we can infer that

\[
x_j + (\alpha_i^{(n)} - \alpha_j^{(n)}) + \sum_{k=1}^{4} (y_j^{(k)} - y_i^{(k)}) \leq x_j \leq x_i,
\]

for all \(y(1), \ldots, y(4) \in C_n\).

These inequalities show that if \(i = \pi(x)\) and \(j \neq i, j \in \{1, \ldots, d\}\), then for sufficiently large \(n\)

\[
x_j - \alpha_j^{(n)} - \sum_{k=1}^{4} y_j^{(k)} \leq x_i - \alpha_i^{(n)} - \sum_{k=1}^{4} y_i^{(k)}
\]

or

\[
\max \left( x - \alpha^{(n)} - \sum_{k=1}^{4} y^{(k)} \right) = x_i - \alpha_i^{(n)} - \sum_{k=1}^{4} y_i^{(k)},
\]

for all \(y(1), \ldots, y(4) \in C_n\) and hence for \(\zeta > 0\) and sufficiently large \(n\)

\[
\max \left( x + \zeta^{(i)} - \alpha^{(n)} - \sum_{k=1}^{4} y^{(k)} \right) = x_i + \zeta - \alpha_i^{(n)} - \sum_{k=1}^{4} y_i^{(k)}.
\]
Consequently if \( i = \pi(x) \), then from the definition of \( f_n \) given by (6.2) we see that for sufficiently large \( n \), \( f_n(x + \zeta(i)) - f_n(x) = \zeta \). Taking the limit as \( n \to \infty \) we therefore obtain

\[
\lim_{n \to \infty} \frac{\partial f_n}{\partial x_i}(x) = 1. 
\] (6.10)

For \( i \in \{1, \ldots, d\} \) define \( q_i, Q_i : \mathbb{R}^d \to \mathbb{R} \) by

\[
q_i(x) = \begin{cases} 
1 & : i = \pi(x) \\
0 & : i \neq \pi(x) 
\end{cases} 
\] (6.11)

and

\[
Q_i(x) = g(\max(x))q_i(x)
\]

for \( x \in \mathbb{R}^d \), where \( \lim_{n \to \infty} (I_{n,4}(h))'(y) = g(y), y \in \mathbb{R} \) as given by condition A1*(b) (see (5.13)).

Using this definition, (6.9) and (6.10) we have

\[
\lim_{n \to \infty} \frac{\partial f_n}{\partial x_i}(x) = q_i(x)
\]

for all \( x \in \mathbb{R}^d \). Thus by the chain rule, condition A1*(b) and (6.5) we see that

\[
\lim_{n \to \infty} \frac{\partial}{\partial x_i} I_{n,4}(h)(f_n(x)) = \lim_{n \to \infty} (I_{n,4}(h))(f_n(x)) \frac{\partial f_n}{\partial x_i}(x)
\]

\[
= g(\max(x))q_i(x)
\]

\[
= Q_i(x)
\] (6.12)

for all \( x \in \mathbb{R}^d \) and \( i \in \{1, \ldots, d\} \). This proves that condition A1(b) holds in the time-dependent case with \( \tau = T \) for the approximating functions \( \frac{\partial}{\partial x_i} (I_{n,4}(h) \circ f_n), n \in \mathbb{N}, \ i \in \{1, \ldots, d\} \) with pointwise limit functions \( Q_i \).

A straightforward calculation using the definition of \( f_n \), shows that for any integers \( n \in \mathbb{N}, i \in \{1, \ldots, d\}, x \in \mathbb{R}^d \) and \( \zeta > 0 \) we have the inequality

\[
f_n(x + \zeta^i) - f_n(x) \leq \zeta
\]

and hence

\[
\left| \frac{\partial f_n}{\partial x_i}(x) \right| \leq 1.
\]

Consequently applying condition A2*(b) (see (5.9)) that holds for \( (I_{n,4}(h))' \), (6.3) and similar arguments used in the derivation of (6.4) we can infer that

\[
\left| \frac{\partial}{\partial x_i} I_{n,4}(h)(f_n(x)) \right|^2 \leq \left| (I_{n,4}(h))'(f_n(x)) \right|^2
\]

\[
\leq K_4^2 \left( 1 + |f_n(x)|^2 \right)
\]

\[
\leq K_4^2 \left( 1 + (\|x\| + \frac{d+1}{n})^2 \right)
\]

\[
\leq K_8^2 \left( 1 + \|x\|^2 \right)
\]
1.7. EXPLICIT HEDGE RATIOS FOR LOOKBACK OPTIONS

for all \( x \in \mathbb{R} \) and \( i \in \{1, \ldots, d\} \), where \( K^2_\theta = 2K^2_\theta (1 + (d + 1)^2) \). This verifies that condition A2(b) holds in the time-independent case with \( \tau = T \) for the approximating functions \( \frac{\partial}{\partial x_i} (I_{n,4}(h) \circ f_n), n \in \mathbb{N} \).

We will now assume that the valuation function \( u_n : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R} \) given by

\[
u_n(t, x) = E \left( I_{n,4}(h) \left( f_n \left(x^{t_0, x}_{T}, \ldots, x^{d, t_0, x}_{T}\right)\right) \right) \tag{6.13}
\]

for \( x = (x_1, \ldots, x_d) \in \mathbb{R}^d \), \( t_0 \leq t \leq T \) is of class \( C^{1,3} \). For example, using Theorem 4.8.6 in Kloeden & Platen (1992) we see that if the drift \( a^i \) and diffusion \( b^{i,j} \) coefficients have uniformly bounded derivatives and are of class \( C^{4,4} \) then \( u_n \) will be of class \( C^{1,4} \). Thus applying the results of Section 3, the payoff function \( I_{n,4}(h) \circ f_n(x^{1, t_0, x}_{T}, \ldots, x^{d, t_0, x}_{T}) \) will admit an Ito integral representation of the form (4.3). Since the payoff function \( h \circ \max : \mathbb{R}^d \rightarrow \mathbb{R} \) with approximating functions \( h_n = I_{n,4}(h) \circ f_n, n \in \mathbb{N} \), satisfies conditions A1 and A2 in Section 1.4 in the time-independent case with \( \tau = T \) we can apply Theorem 1.4.1 with \( \tau = T \) to the functional \( h \circ \max \) and the diffusion process \( X^{t_0, x}_{T} \) yielding

\[
\begin{align*}
\left( \max \left(X^{t_0, x}_{T}\right) \right) & = E \left( h \left( \max \left(X^{t_0, x}_{T}\right) \right) \right) + \sum_{j=1}^{d} \sum_{i=1}^{m} \int_{t_0}^{T} \gamma^{i,j}_s \left( s, X^{t_0, x}_{s} \right) dW^j_s \tag{6.14}
\end{align*}
\]

where

\[
\gamma^{i}_s = E \left( \sum_{p=1}^{d} Q_p \left(X^{t_0, x}_{s}\right) \mathbb{Z}^{p,i,s,\delta}_T \left| F_s \right. \right)
\]

for \( i \in \{1, \ldots, d\} \) and \( \mathbb{Z}^{p,i,s,\delta}_T \) is the unique solution of the stochastic differential equation (3.5) with initial value at time \( s, s \in [t_0, T] \) of \( \delta \) as given by (3.8).

Thus we have found an explicit representation of the payoff structure (6.1). This is the case when \( h \) is an absolutely continuous function of the form (5.1) and the approximating valuation functions \( u_n, n \in \mathbb{N} \), are sufficiently smooth.

1.7 Explicit Hedge Ratios for Lookback Options

In this section another important application of Theorem 1.4.1 related to lookback options will be discussed.

Let \( X^{t_0, x}_{T} = \{X^{t_0, x}_{t}, t_0 \leq t \leq T\} \) be the solution of the one-dimensional stochastic differential equation (2.1) starting at time \( t_0 \) with initial value \( x \in \mathbb{R} \). We assume \( h : \mathbb{R} \rightarrow \mathbb{R} \) is a payoff function for which conditions A1* and A2* given in Section 1.5 hold. We now consider so called lookback options with payoff structures of the form

\[
h \left( \sup_{t_0 \leq t \leq T} X^{t_0, x}_{t} \right). \tag{7.1}
\]

To ensure that option prices for this payoff structure are well-defined we also assume that the mean square bound (4.1) holds. Our aim will be to find an explicit representation for the payoff structure (7.1) of the form (4.5).
For integers $m \in \mathbb{N}$ and $i \in \{1, \ldots, d_m\}$, where $d_m = 2^m$, let $(t)_{\Delta_m}$ be an equispaced time discretization of the interval $[t_0, T]$ of the form $t_0 < t_1 < \ldots < t_{d_m} = T$ with step size $\Delta_m = (T - t_0)/d_m$. We also denote by $\bar{X}^{t_0, \bar{x}} = \{\bar{X}^{t_0, \bar{x}} = (X_1^{t_0, \bar{x}}, \ldots, X_{d_m}^{t_0, \bar{x}}), t_0 \leq t \leq T\}$ the unique solution of the $d_m$-dimensional stochastic differential equation with components

$$dX_t^{i, t_0, \bar{x}} = a\left(t, X_t^{i, t_0, \bar{x}}\right) dt + b\left(t, X_t^{i, t_0, \bar{x}}\right) dW_t \quad (7.2)$$

for $t_0 \leq t \leq T$, $i \in \{1, \ldots, d_m\}$, starting at time $t_0$ with initial value $\bar{x} = (\bar{x}, \ldots, \bar{x}) \in \mathbb{R}^{d_m}$. For simplicity we will use the symbol $\bar{x}$ for both, the initial value of the vector diffusion $\bar{X}^{t_0, \bar{x}}$ at time $t_0$, and the initial value of the 1-dimensional diffusion $X_1^{t_0, \bar{x}}$ at time $t_0$.

For each integer $i \in \{1, \ldots, d_m\}$, the evolution of the component process $X_t^{i, t_0, \bar{x}}$ is stopped at time $t_i$ and we can write

$$X_t^{i, t_0, \bar{x}} = X_{t \wedge t_i}^{i, t_0, \bar{x}} \quad (7.3)$$

for $t_0 \leq t \leq T$. For the discretization grid with step size $\Delta_m$ define

$$M_T^{m, t_0, \bar{x}} = \max_{i \in \{1, \ldots, d_m\}} X_T^{i, t_0, \bar{x}} \quad (7.4)$$

and

$$M_T^{\infty, t_0, \bar{x}} = \sup_{t_0 \leq t \leq T} X_t^{0, \bar{x}}.$$  

Let $h_n = I_{n,4}(h)$, $n \in \mathbb{N}$ be the sequence of approximating functions for $h$ as given by (5.4) and consider the payoff structure

$$h_n \left(\max_{i \in \{1, \ldots, d_m\}} X_T^{i, t_0, \bar{x}}\right) = h_n \left( M_T^{m, t_0, \bar{x}}\right) \quad (7.5)$$

for $n \in \mathbb{N}$. As in Section 1.6 we assume that the valuation function $u_{n,m}$ given by

$$u_{n,m}(t, x_1, \ldots, x_{d_m}) = E\left(h_n \left( M_T^{m, t_0, \bar{x}}\right)\right)$$

for $x = (x_1, \ldots, x_{d_m}) \in \mathbb{R}^{d_m}$ and $t_0 \leq t \leq T$ is of class $C^{1,3}$ and hence equation (6.14), will admit a representation of the form

$$h_n \left( M_T^{m, t_0, \bar{x}}\right) = E\left(h_n \left( M_T^{m, t_0, \bar{x}}\right)\right) + \sum_{i=1}^{d_m} \int_{t_0}^{T} \gamma_s^i 1_{\{s < t_i\}} b\left(t, X_t^{i, t_0, \bar{x}}\right) dW_s \quad (7.6)$$

where

$$\gamma_s^i = E \left( \sum_{p=1}^{d_m} Q_p \left( \bar{X}_T^{t_0, \bar{x}} \right) Z_T^{p, i, \bar{s}, \bar{x}} \big| \mathcal{F}_s \right)$$

$$= E \left( \sum_{p=1}^{d_m} h_{n_p} \left( M_T^{m, t_0, \bar{x}}\right) q_p \left( \bar{X}_T^{t_0, \bar{x}} \right) Z_T^{p, i, \bar{s}, \bar{x}} \big| \mathcal{F}_s \right) \quad (7.6)$$
for \( n, m \in \mathbb{N} \), with the functions \( Q_p \) and \( q_p \), \( p \in \{1, \ldots, d_m\} \), defined by (6.12) and (6.11), respectively, and \( Z^{s,\delta} \) is the unique solution of the \( d_m \)-dimensional stochastic differential equation (3.5) with initial value \( \delta \) given by (3.8) at time \( s, t_0 \leq s \leq T \).

Note that an indicator function is included as part of the integrand in (7.6) because it forms part of the diffusion coefficient appearing in (7.2). Also, as this representation depends on an application of Theorem 1.4.1 we assume the growth bound (4.2) holds for the process \( Z^{s,\delta} \).

Using the specific structure that applies for the underlying vector diffusion \( \bar{X}^{t_0,\xi} \) given by (7.2) the stochastic differential equation for the component \( Z^{p,i,s,\delta} \), \( p, i \in \{1, \ldots, d_m\} \), can be simplified and expressed in the form

\[
dZ^{p,i,s,\delta}_t = 1_{\{t < t_p\}} Z^{p,i,s,\delta}_t \left( \frac{\partial a}{\partial x} \left( t, X^{p,t_0,\xi}_t \right) dt + \frac{\partial b}{\partial x} \left( t, X^{p,t_0,\xi}_t \right) dW_t \right)
\]

for \( t_0 \leq s \leq t \leq T \). The solution to (7.7) for each \( p, i \in \{1, \ldots, d_m\} \), has an exponential form and therefore applying the initial condition \( \delta_{p,i} = 0 \), for \( p \neq i \) given by (3.8), we obtain

\[
Z^{p,i,s,\delta}_t = 0 \quad \text{P.-a.s.}
\]

for \( t_0 < s \leq t \leq T \). If we let \( Z^{s,1} = \{Z_s^{1,1}; s \leq t \leq T\} \) be the solution of the one-dimensional linearized stochastic differential equation (2.10) then from (7.7) and the initial condition \( \delta_{p,i} = 1 \) for \( p = i \) we can write

\[
Z^{i,i,s,\delta}_t = Z^{i,1}_t
\]

for \( t_0 \leq t \leq T \), \( i \in \{1, \ldots, d_m\} \). By (7.8) the representation given by (7.6) can also be simplified and expressed as

\[
h_n \left( M_{T}^{m,t_0,\xi} \right) = E \left( h_n \left( M_{T}^{m,t_0,\xi} \right) \right) + \sum_{i=0}^{d_m} \int_{t_0}^{T} E \left( h'_n \left( M_{T}^{m,t_0,\xi} \right) q_i \left( \bar{X}_{T}^{t_0,\xi} \right) Z^{i,i,s,\delta}_T 1_{\{s < t\}} | \mathcal{F}_s \right) b(s, \bar{X}_s^{t_0,\xi}) dW_s
\]

for \( n, m \in \mathbb{N} \). For each \( m \in \mathbb{N} \) and discretization grid \( (t)_{\Delta m} \) let \( \tau_m = \pi(\bar{X}_T^{t_0,\xi}) \), where \( \pi : \mathbb{R}^{d_m} \to \{1, \ldots, d_m\} \) is given by (6.7). Using this function we define \( \tau_m : \Omega \to \mathbb{R} \) by

\[
\tau_m = t_{\pi_m}.
\]
The variable $\tau_m$ is a random time which is $\mathcal{F}_T$-measurable but $t \wedge \tau_m$ may not be $\mathcal{F}_T$-measurable and consequently $\tau_m$ will not in general be a stopping time.

From the definition of the random time $\tau_m$ and $q_i$, $i \in \{1, \ldots, d_m\}$ together with (7.9) and (7.10) we see that

$$h_n \left( M^{m,t_0}_T \right) = E \left( h_n \left( M^{m_0}_T \right) \right)$$

$$+ \int_0^T E \left( h' \left( M^{m,t_0}_T \right) Z_{\tau_m} \mathbb{1}_{\{s < \tau_m\}} \right| \mathcal{F}_s \right) b \left( s, X^{t_0}_s \right) dW_s.$$

(7.12)

By continuity of the sample paths of $X^{t_0}_t$ we know that for a fixed integer $n \in \mathbb{N}$

$$\lim_{m \to \infty} h_n \left( M^{m_0,t_0}_T \right) = h_n \left( M^{\infty,t_0}_T \right) \quad \text{P-a.s.,}$$

(7.13)

where $M^{\infty,t_0}_T$ is given in (7.4). Also, the linear growth of $h_n$, given by condition A2*(a), means that

$$\left| h_n \left( M^{m_0,t_0}_T \right) \right|^2 \leq K_3^2 \left( 1 + \left| M^{m_0,t_0}_T \right|^2 \right)$$

$$\leq K_3^2 \left( 1 + \sup_{t_0 \leq t \leq T} \left| X^{t_0}_t \right|^2 \right).$$

(7.14)

Note that this bound applies uniformly for both integer variables $n$ and $m$. In addition from the mean square bound (4.1) we see that

$$E \left( K_3^2 \left( 1 + \sup_{t_0 \leq t \leq T} \left| X^{t_0}_t \right|^2 \right) \right) < \infty.$$

This result together with (7.13) and (7.14) means that we can apply a version of the Dominated Convergence Theorem applicable for $L^p$ spaces, $p > 0$, see for example the Corollary to Theorem 2.6.3 in Shiryayev (1984), to obtain

$$\lim_{m \to \infty} \left\| h_n \left( M^{m_0,t_0}_T \right) - h_n \left( M^{\infty,t_0}_T \right) \right\|_2 = 0$$

(7.15)

for all integers $n \in \mathbb{N}$. Furthermore by Hölder’s inequality

$$\left| E \left( h_n \left( M^{m_0,t_0}_T \right) \right) - E \left( h_n \left( M^{\infty,t_0}_T \right) \right) \right|$$

$$\leq E \left( \left| h_n \left( M^{m_0,t_0}_T \right) - h_n \left( M^{\infty,t_0}_T \right) \right| \right)$$

$$= \left\| h_n \left( M^{m_0,t_0}_T \right) - h_n \left( M^{\infty,t_0}_T \right) \right\|_1$$

$$\leq \left\| h_n \left( M^{m_0,t_0}_T \right) - h_n \left( M^{\infty,t_0}_T \right) \right\|_2$$

so that from (7.15)

$$\lim_{m \to \infty} \left| E \left( h_n \left( M^{m_0,t_0}_T \right) \right) - E \left( h_n \left( M^{\infty,t_0}_T \right) \right) \right| = 0$$

(7.16)
for all $n \in \mathbb{N}$.

Since $d_m = 2^m$ for $m \in \mathbb{N}$ any discretization point $t_i$, $i \in \{1, \ldots, d_m\}$, belonging to the grid $(t)_{\Delta_m}$ will be an element of the set of discretization points for any grid $(t)_{\Delta_m'}$ with $m' \in \mathbb{N}$ and $m' \geq m$. This means that $(\tau_m)_{m \in \mathbb{N}}$ is a non-decreasing sequence of random times and thus $\lim_{m \to \infty} \tau_m(\omega)$ exists for each $\omega \in \Omega$. Let

$$\lim_{m \to \infty} \tau_m(\omega) = \tau_\infty(\omega) \quad (7.17)$$

for some random time $\tau_\infty : \Omega \to \mathbb{R}$ and let $s \in [t_0, T]$ be fixed. If $\omega \in \Omega$ and $s < \tau_\infty(\omega)$, then by (7.17) there is an $M > 0$ such that for $m > M$ we have $s < \tau_m(\omega)$. If $s \geq \tau_\infty(\omega)$ then clearly $s \geq \tau_m(\omega)$ for all $m \in \mathbb{N}$. Thus

$$\lim_{m \to \infty} 1_{\{s < \tau_m\}} = 1_{\{s < \tau_\infty\}} \quad P\text{-a.s.} \quad (7.18)$$

Combining this result, (7.13) and by continuity of the function $h'_n$ and the sample paths of $Z^{s,1}$ we can infer that

$$\lim_{m \to \infty} h'_n \left( M^{m,t_0,x}_T \right) Z^{s,1}_{\tau_m} 1_{\{s < \tau_m\}} = h'_n \left( M^{\infty,t_0,x}_T \right) Z^{s,1}_{\tau_\infty} 1_{\{s < \tau_\infty\}} \quad P\text{-a.s.} \quad (7.19)$$

for any $n \in \mathbb{N}$.

Applying the linear growth bound $A2^*(b)$, we know that for integers $n, m \in \mathbb{N}$ and $t_0 \leq s \leq T$

$$\left| h'_n \left( M^{m,t_0,x}_T \right) Z^{s,1}_{\tau_m} 1_{\{s < \tau_m\}} \right| \leq K_4 \sqrt{\left( 1 + \left| M^{m,t_0,x}_T \right|^2 \right) \left| Z^{s,1}_{\tau_m} \right|} \leq K_4 \sqrt{\left( 1 + \sup_{t_0 \leq t \leq T} \left| X^{t_0,x}_t \right|^2 \right) \left| Z^{s,1}_{\tau_m} \right|}. \quad (7.20)$$

Also, applying the mean square bounds (4.1) and (4.2) together with Hölder’s inequality we have

$$E \left( K_4 \sqrt{\left( 1 + \sup_{t_0 \leq t \leq T} \left| X^{t_0,x}_t \right|^2 \right) \left| Z^{s,1}_{\tau_m} \right|} \right) \leq K_4 \left\| \sqrt{\left( 1 + \sup_{t_0 \leq t \leq T} \left| X^{t_0,x}_t \right|^2 \right)} \right\|_2 \left\| Z^{s,1}_{\tau_m} \right\|_2 \quad < \infty \quad (7.21)$$

for $t_0 \leq t \leq T$. This inequality combined with (7.19) and (7.20) means that we can apply the Dominated Convergence Theorem for conditional expectations, see for example Shiryaev (1984), Theorem 2.7.2 to obtain

$$\lim_{m \to \infty} E \left( h'_n \left( M^{m,t_0,x}_T \right) Z^{s,1}_{\tau_m} 1_{\{s < \tau_m\}} \big| \mathcal{F}_s \right) = E \left( h'_n \left( M^{\infty,t_0,x}_T \right) Z^{s,1}_{\tau_\infty} 1_{\{s < \tau_\infty\}} \big| \mathcal{F}_s \right) \quad P\text{-a.s.} \quad (7.22)$$

for all $n \in \mathbb{N}$ and $t_0 \leq s \leq T$. 

Now define the process $\xi_m$, $m \in \mathbb{N}$, and $\xi_\infty$ by
\[
\xi_m = E \left( h_n^m \left( M_T^{m,t_0,\xi} \right) Z_{\tau_m}^{s,1} 1_{\{s < \tau_m\}} | \mathcal{F}_s \right) b \left( s, X_s^{t_0,\xi} \right), \\
\xi_\infty = E \left( h_n^m \left( M_T^{\infty,t_0,\xi} \right) Z_{\tau_\infty}^{s,1} 1_{\{s < \tau_\infty\}} | \mathcal{F}_s \right) b \left( s, X_s^{t_0,\xi} \right).
\]
Equation (7.22) shows that
\[
\lim_{m \to \infty} \xi_m = \xi_\infty \text{ P-a.s.}
\]
and therefore with these choices for $\xi_m$, $m \in \mathbb{N}$ and $\xi_\infty$, condition (b) is satisfied in the statement of Lemma 1.4.3.

In addition, the limits (7.15) and (7.16) combined with the representation (7.12) show that the Ito integral in (7.12) form a Cauchy sequence in $L^2(\Omega, \mathcal{F}_T, P)$ for $m = 1, 2, \ldots$ and fixed $n \in \mathbb{N}$. Moreover using similar arguments to those given in Section 1.4, see the commentary following equation (4.24), it can be shown that the integrand in the Ito integral of (7.12) is $\mathcal{F}$-adapted with continuous sample paths.

Consequently condition (a) is satisfied in the statement of Lemma 1.4.3. These results mean that Lemma 1.4.3 can be applied which together with (7.15) and (7.16) shows that
\[
\begin{align*}
\text{for } n \in \mathbb{N}, \\
\text{we have an explicit representation for a lookback option for the smooth payoff function } h_n, \\
\text{and these methods also extend to the non-smooth payoff function } h.
\end{align*}
\]
In fact using conditions A1*, A2* (a) and similar arguments to those given for the proof of (7.15), (7.16), (7.19) and (7.20) we can show
\[
\begin{align*}
\lim_{n \to \infty} \left\| h_n \left( M_T^{\infty,t_0,\xi} \right) - h \left( M_T^{\infty,t_0,\xi} \right) \right\|_2 = 0, \\
\lim_{n \to \infty} \left| E \left( h_n \left( M_T^{\infty,t_0,\xi} \right) \right) - E \left( h \left( M_T^{\infty,t_0,\xi} \right) \right) \right| = 0, \\
\lim_{n \to \infty} h_n' \left( M_T^{\infty,t_0,\xi} \right) Z_{\tau_m}^{s,1} 1_{\{s < \tau_m\}} = g \left( M_T^{\infty,t_0,\xi} \right) Z_{\tau_m}^{s,1} 1_{\{s < \tau_m\}}, \text{ P-a.s.} \\
|h_n' \left( M_T^{\infty,t_0,\xi} \right) Z_{\tau_m}^{s,1} 1_{\{s < \tau_m\}}| \leq K_4 \sqrt{\left( 1 + \sup_{t_0 \leq t \leq T} |X_t^{t_0,\xi}|^2 \right) \left| Z_{\tau_m}^{s,1} \right|},
\end{align*}
\]
with the last two relations and (7.21) showing that
\[
\begin{align*}
\lim_{n \to \infty} E \left( h_n' \left( M_T^{\infty,t_0,\xi} \right) Z_{\tau_m}^{s,1} 1_{\{s < \tau_m\}} | \mathcal{F}_s \right) \\
= E \left( g \left( M_T^{\infty,t_0,\xi} \right) Z_{\tau_m}^{s,1} 1_{\{s < \tau_m\}} | \mathcal{F}_s \right) \text{ P-a.s.} \quad (7.25)
\end{align*}
\]
1.7. EXPLICIT HEDGE RATIOS FOR LOOKBACK OPTIONS

The first two equations in (7.24) and relation (7.23) show that the Ito integral in (7.23) is a Cauchy sequence in \( L^2(\Omega, \mathcal{F}_T, P) \) for \( n = 1, 2, \ldots \). Also from (7.25) we see that

\[
\lim_{n \to \infty} E \left( h_n \left( M_T^{\infty, t_0, \bar{x}} \right) Z_{\tau_{\infty}} 1_{\{s < \tau_{\infty}\}} \mid \mathcal{F}_s \right) b \left( s, X_{s}^{t_0, \bar{x}} \right)
= E \left( g \left( M_T^{\infty, t_0, \bar{x}} \right) Z_{\tau_{\infty}} 1_{\{s < \tau_{\infty}\}} \mid \mathcal{F}_s \right) b \left( s, X_{s}^{t_0, \bar{x}} \right) \quad P\text{-a.s.}
\]  

(7.26)

Furthermore, the random variables on both sides of the above equation are clearly \( \mathcal{F}_s \)-measurable and consequently by an application of Lemma 1.4.3 and using again the first two equations in (7.23) we can infer that

\[
h(M_T^{\infty, t_0, \bar{x}}) = E \left( h \left( M_T^{\infty, t_0, \bar{x}} \right) \right)
+ \int_{t_0}^{T} E \left( g \left( M_T^{\infty, t_0, \bar{x}} \right) Z_{\tau_{\infty}} 1_{\{s < \tau_{\infty}\}} \mid \mathcal{F}_s \right) b \left( s, X_{s}^{t_0, \bar{x}} \right) dW_s.
\]

(7.27)

Thus we have found a representation of the payoff structure (7.1) of the form (4.5), with the non-smooth payoff function \( h \) which satisfies conditions A1* and A2* given in Section 1.5.

We remark that finding an explicit Ito integral representation of the payoff structure for a lookback option is generally considered as one of the more difficult problems concerning the computation of hedge ratios for path-dependent options. For example Foellmer (1991) using methods employed by M. Schweizer applies Clark’s formula to obtain an explicit representation in the simple case, where \( X^{t_0, \bar{x}} \) is a geometric Brownian motion and \( h \) is the identity function.

Other path-dependent options such as Asian options can be also be handled using the methods developed in this section. In fact, the analysis is often simpler because the state space may only need to be increased by a single extra variable.

Our methods can also be adapted to the case, where the underlying model is continuous but trading activities or observations are restricted to certain times say \( t_0 < t_1 < \ldots < t_{d_m} = T \) with \( d_m = 2^m \).

The payoff structure (7.5) with \( h \) replacing \( h_n, n \in \mathbb{N} \), becomes

\[
h \left( \max_{1 \leq i \leq d_m} X_{T_i}^{t_0, \bar{x}} \right) = h \left( M_T^{m, t_0, \bar{x}} \right).
\]

Applying similar arguments to those used to obtain the representation (7.27), with \( M_T^{m, t_0, \bar{x}} \) replacing \( M_T^{\infty, t_0, \bar{x}} \), and \( \tau_m \) replacing \( \tau_{\infty} \), we can show that

\[
h \left( M_T^{m, t_0, \bar{x}} \right) = E \left( h \left( M_T^{m, t_0, \bar{x}} \right) \right)
+ \int_{t_0}^{T} E \left( g \left( M_T^{m, t_0, \bar{x}} \right) Z_{\tau_{m}} 1_{\{s < \tau_{m}\}} \mid \mathcal{F}_s \right) b \left( s, X_{s}^{t_0, \bar{x}} \right) dW_s.
\]

(7.28)
In this case we have therefore found an explicit Ito integral representation of the payoff structure of a lookback option with discrete observations or fixings. We do not require the limiting arguments needed in the derivation of (7.13), (7.15)–(7.19) and (7.22).

Although most sections of this chapter appear rather technical, they provide powerful and constructive methods which enables a wide class of payoff structures to be represented in the form of stochastic integrals with explicit expressions for the integrands. Using these stochastic integrals we can find corresponding explicit formulas to compute the hedge ratios for the underlying derivative security.
Chapter 2

Variance Reduction Techniques

Naive Monte Carlo estimates of the expectation appearing in (1.1.5), and from which derivative security prices can be computed, can be very expensive in terms of computer resource usage. In this chapter we investigate the important problem of finding efficient variance reduced estimators for the expectation of functionals of Ito diffusion processes.

A number of unbiased variance reduced estimators are considered, the first of which is based on a measure transformation procedure which is related to the classical variance reduction technique of importance sampling. The application of measure transformations to reduce the variance of functionals of Ito diffusion processes was first proposed by Milstein (1988) and has been subsequently developed and used for financial modelling problems by Hofmann, Platen & Schweizer (1992) and Fournie, Lasry & Touzi (1995). We extend this method so that it can be applied to a wide class of diffusion processes and payoff structures and find explicit expressions for the variance of the resulting estimators. These calculations are important because they provide precise information on what factors contribute to the variance and what factors need to be controlled in any practical implementation of the method.

In Section 3 we use discrete time numerical methods to approximate the continuous time estimators previously obtained and to compute the variance of the resulting discrete time estimators. We also find discrete approximations for the integrands appearing in the Ito integral representations of the discounted price processes of derivative securities. These methods are of interest theoretically because they use only simple properties of stochastic processes and are of practical value because of their close association with the actual implementation that is likely on a digital computer.

Following this, we consider the application of control variate methods and introduce a class of Ito integral control variates which extend and provide an elegant continuous time formulation of the martingale variates proposed by Clewlow & Carverhill (1992, 1994). We show how the variance of linear combinations of these control variates can be minimized using least-squares analysis and orthogonalization techniques. In the last section of this chapter we provide some new perspectives and insights on how classical variance reduction methods such as conditioning, stratified sampling, use of antithetic variates and quasi Monte Carlo can be adapted and applied for derivative security valuation problems.

We describe three alternative approaches to quasi Monte Carlo and recommend one which is based on the use of multipoint approximations of Wiener increments and which provides some useful practical extensions to the work of Barraquand (1993).
2.1 Measure Transformations

In this section we apply a measure transformation method and the Ito formula to build variance reduced estimators that can be used for a wide class of derivative security valuation problems. The construction of these estimators requires only very weak integrability conditions to hold for the underlying payoff structure and in this sense significantly extends the results of Milstein (1988), Hofmann, Platen & Schweizer (1992) and Fournie, Lasry & Touzi (1995).

As in Chapter 1 we let $W = (W^1, \ldots, W^m)$ be an $m$-dimensional Wiener process defined on the probability space $(\Omega, \mathcal{F}, P)$, where $\mathcal{F} = (\mathcal{F}_t)_{t_0 \leq t \leq T}$ denotes the $P$-augmentation of the natural filtration of $W$. Let $M = \{M_t, t_0 \leq t \leq T\}$ be a square integrable martingale with respect to the filtration $\mathcal{F}$ and measure $P$ with the Kunita-Watanabe representation

$$M_t = M_{t_0} + \sum_{j=1}^{m} \int_{t_0}^{t} \xi^j_s dW^j_s,$$

where $\xi = (\xi^1, \ldots, \xi^m)$ is a vector of $\mathcal{F}$-predictable processes with

$$E \left( \int_{t_0}^{T} ||\xi_s||^2 \, ds \right) < \infty$$

for $t_0 \leq t \leq T$.

Let $(\Omega, \mathcal{F}, \tilde{P})$ be another probability space with the same sample space $\Omega$ and filtration $\mathcal{F}$ but a different probability measure $\tilde{P}$.

We will say that an adapted process $G = \{G_t, t_0 \leq t \leq T\}$ is an unbiased estimator under $\tilde{P}$ for $E(M_T) = M_{t_0}$ at time $t$ if

$$\tilde{E}(G_t) = E(M_T),$$

where $\tilde{E}$ denotes expectation with respect to the measure $\tilde{P}$. Ideally we seek unbiased estimators of $E(M_T)$ since in these cases we only need to deal with the process $G$ in order to approximate $E(M_T)$. If in addition, the inequality

$$\tilde{\text{Var}}(G_t) < \text{Var}(M_T)$$

holds, where $\tilde{\text{Var}}(G_t)$ denotes the variance of $G_t$ under $\tilde{P}$ we will say that $G$ is a variance reduced unbiased estimator under $\tilde{P}$ for $E(M_T)$ at time $t, \ t_0 \leq t \leq T$. For many practical valuation problems we require variance reduced estimators of $E(M_T)$ because the statistical error associated with simulation estimates of $\tilde{E}(G_t), \ t_0 \leq t \leq T$, depends on the variance under $\tilde{P}$ of $G_t$. This error will be manageable if this variance is small. Clearly these definitions extend to the case where we replace $t$ with some stopping time $\tau$.

Let $d = (d^1, \ldots, d^m)$ be a vector of adapted measurable processes with

$$E \left( \exp \left( \frac{1}{2} \int_{t_0}^{T} ||d_s||^2 \, ds \right) \right) < \infty.$$
2.1. MEASURE TRANSFORMATIONS

Also, let \( P_t , t_0 \leq t \leq T \) denote the restriction of the measure \( P \) to the \( \sigma \)-algebra \( \mathcal{F}_t \). Using the Girsanov transformation we know there is a corresponding measure \( \tilde{P}_t \) such that the process \( \tilde{W} = (\tilde{W}^1, \ldots, \tilde{W}^m) \) given by

\[
\tilde{W}_t^j = W_t^j - \int_{t_0}^{t} d\tilde{W}_s^j \quad (1.4)
\]

for \( t_0 \leq t \leq T, \ j \in \{1, \ldots, m\} \) is an \( m \)-dimensional Wiener process under \( \tilde{P}_t \), where \( \tilde{P}_t \) is defined using the Radon-Nikodym derivative

\[
\frac{d\tilde{P}_t}{dP_t} = \theta_t = \exp \left\{ -\frac{1}{2} \int_{t_0}^{t} ds \frac{d}{ds} + \sum_{j=1}^{m} \int_{t_0}^{t} ds dW_s^j \right\}
\]

and \( \theta = \{ \theta_t, t_0 \leq t \leq T \} \) is a process which is the solution of the integral equation

\[
\theta_t = 1 + \sum_{j=1}^{m} \int_{t_0}^{t} ds \theta_s dW_s^j \quad (1.5)
\]

for \( t_0 \leq t \leq T, \ \text{and} \ j \in \{1, \ldots, m\} \).

From (1.3) it can be shown that \( \theta \) is a square integrable martingale under the measure \( P_t \) with \( E(\theta_T) = 1, \ t_0 \leq t \leq T \). This result was shown by Novikov (1972). Since the process \( \theta \) is a martingale, and \( P_T = P \), then for any event \( A \in \mathcal{F}_T \) we have

\[
\tilde{P}_T(A) = \int_A \theta_T = \int_A \theta_T dP_T = \int_A E(\theta_T | \mathcal{F}_T) dP_T = \int_A \theta_T dP_T = \tilde{P}_T(A).
\]

This means that \( \tilde{P}_T \) is actually the restriction of \( \tilde{P}_T \) to the \( \sigma \)-algebra \( \mathcal{F}_T \).

Let \( \tilde{M} = \{ \tilde{M}_t, 0 \leq t \leq T \} \) be a square integrable martingale with respect to the filtration \( \mathcal{F} \) and the measure \( \tilde{P} \), with representation

\[
\tilde{M}_t = \tilde{M}_{t_0} + \sum_{j=1}^{m} \int_{t_0}^{t} \tilde{x}_s^j d\tilde{W}_s^j, \quad (1.6)
\]

where \( \tilde{x} = (\tilde{x}^1, \ldots, \tilde{x}^m) \) is an \( \mathcal{F} \)-predictable process with

\[
\tilde{E} \left( \int_{t_0}^{T} \| \tilde{x}_s \|^2 ds \right) < \infty. \quad (1.7)
\]

For the moment we will regard the process \( \tilde{x} \) as being unspecified except for the requirement that \( \tilde{x} \) be \( \mathcal{F} \)-predictable and that the mean square integrability condition (1.7) holds. We assume that

\[
\tilde{E}(\tilde{M}_T) = E(M_T). \quad (1.8)
\]

This condition will be verified later on for a wide class of contingent claims.

Equation (1.8) together with (1.1) and (1.6) imply the relations

\[
E(M_t) = E(M_T) = \tilde{E}(\tilde{M}_T) = \tilde{E}(\tilde{M}_t) \quad (1.9)
\]

for \( t_0 \leq t \leq T \).
CHAPTER 2. VARIANCE REDUCTION TECHNIQUES

Since the restrictions of the measures \( P = P_T \) and \( \tilde{P} = \tilde{P}_T \) to \( \mathcal{F}_t \) are \( P_t \) and \( \tilde{P}_t \) respectively, we have, using the Radon-Nikodym derivative \( \theta_t = \frac{d\tilde{P}_t}{dP_t} \), the relations

\[
\mathbb{E}(\tilde{M}_t) = \int_\Omega \tilde{M}_t \, d\tilde{P}_T = \int_\Omega \tilde{M}_t \, d\tilde{P}_t = \int_\Omega \tilde{M}_t \, dP_t = \int_\Omega \tilde{M}_t \, dP_T = \mathbb{E}(\tilde{M}_t \theta_t) \quad \text{(1.10)}
\]

for \( t_0 \leq t \leq T \). Combining (1.9) and (1.10) we see that the process \( \tilde{M} \theta = \{ \tilde{M}_t \theta_t, 0 \leq t \leq T \} \) is an unbiased estimator under \( P \) for \( \mathbb{E}(M_T) \) at time \( t \), \( t_0 \leq t \leq T \).

We will now compute the variance of the estimator \( \tilde{M} \theta \) under the measure \( P \). To do this we first express the martingale \( \tilde{M} \) as a semimartingale using the Wiener process \( W \) rather than \( \tilde{W} \).

Applying (1.4) and (1.6) we have

\[
\tilde{M}_t = \tilde{M}_{t_0} - \sum_{j=1}^m \int_{t_0}^t d_j \tilde{\xi}_j \, ds + \sum_{j=1}^m \int_{t_0}^t \tilde{\xi}_j \, dW_j. \quad (1.11)
\]

Expanding the estimator \( \tilde{M} \theta \) by the Ito rule together with (1.5) and (1.11), the integral equation for \( \tilde{M}_t \theta_t \) becomes

\[
\tilde{M}_t \theta_t = \tilde{M}_{t_0} + \sum_{j=1}^m \int_{t_0}^t \theta_s (\tilde{\xi}_j + d_j \tilde{M}_s) \, dW_j. \quad (1.12)
\]

for \( t_0 \leq t \leq T \).

This shows that \( \tilde{M} \theta \) is an \((\mathcal{F}, P)\)-martingale. From this formula we can also verify the fact, previously noted, that the process \( \tilde{M} \theta \) is an unbiased estimator under \( P \) for \( \mathbb{E}(M_t) \) at time \( t \), \( t_0 \leq t \leq T \). The variance of the product \( \tilde{M}_t \theta_t \) under \( P \), denoted by \( \text{Var}(\tilde{M}_t \theta_t) \), can now be computed using the equations

\[
\text{Var}(\tilde{M}_t \theta_t) = \mathbb{E} \left( \left( \sum_{j=1}^m \int_{t_0}^t \theta_s (\tilde{\xi}_j + d_j \tilde{M}_s) \, dW_j \right)^2 \right) = \int_{t_0}^t \mathbb{E} \left( \sum_{j=1}^m (\tilde{\xi}_j + d_j \tilde{M}_s)^2 \right) \, ds \quad (1.13)
\]

for \( t_0 \leq t \leq T \).

Consequently if the inequality \( \tilde{M}_t > 0 \) holds \( P\text{-a.s.} \) for \( t_0 \leq t \leq T \) and we choose \( d_j = -\tilde{\xi}_j / \tilde{M}_s \) for all \( j \in \{1, \ldots, m\} \), and \( t_0 \leq s \leq T \), then the variance of the random
2.1. MEASURE TRANSFORMATIONS

variable $\tilde{M}_t \theta_t$ under $P$ is reduced to zero for any $t$, $0 \leq t \leq T$. Of course in this form we have assumed that the integrand process $\tilde{\xi}$ and the martingale $\tilde{M}$ can somehow be determined. In practice this condition is difficult to satisfy. Even in the case where an explicit form exists for the process $\tilde{\xi}$ they usually cannot be easily computed.

An alternative is to find approximations $\tilde{\xi} = (\tilde{\xi}^1, \ldots, \tilde{\xi}^m)$ and $\tilde{M}$ for the processes $\tilde{\xi}$ and $\tilde{M}$, respectively, with $\tilde{M}_t > 0$ $P$-a.s., for all $0 \leq t \leq T$ and set

$$d_t^j = -\frac{\tilde{\xi}_t^j}{\tilde{M}_t},$$

for $0 \leq t \leq T$ and $j \in \{1, \ldots, m\}$, so that from (1.13)

$$\text{Var} (\tilde{M}_t \theta_t) = E \left( \left( \sum_{j=1}^m \int_{t_0}^t \theta_s \left( \tilde{\xi}_s^j - \tilde{\xi}_s^j \frac{\tilde{M}_s}{M_s} \right) dW_s^j \right)^2 \right)$$

$$= \int_{t_0}^t E \left( \theta_s^2 \sum_{j=1}^m \left( \tilde{\xi}_s^j - \tilde{\xi}_s^j \frac{\tilde{M}_s}{M_s} \right)^2 \right) ds.$$  (1.15)

This formula for the variance of $\tilde{M}_t \theta_t$ under $P$ is an important practical result as it tells us exactly what factors need to be controlled to reduce the variance of the estimator $\tilde{M} \theta$. In particular it shows that the process $\theta^2$ also contributes to the variance of $\tilde{M}_t \theta_t$. If $\theta^2$ is allowed to explode, then Var$(\tilde{M}_t \theta_t)$ may become unacceptably large. Stability problems can also easily arise if $\tilde{M}_t \theta_t$ is being approximated using stochastic numerical methods and $\theta$ is not constrained.

To see how these problems can be controlled suppose the process $\tilde{M}$ is bounded from below by the value $\alpha > 0$, that is

$$\tilde{M}_t(\omega) \geq \alpha$$

for all $(t, \omega) \in [t_0, T] \times \Omega$ and the approximation $\tilde{\xi}$ satisfies a mean square integrability bound of the form (1.7). Then for a suitably large choice of $\alpha$, the process $d = (d^1, \ldots, d^m)$, given by (1.14) will be small, in a mean square sense, since $|d_t^j| = |\tilde{\xi}_t^j/\tilde{M}_t| \leq \frac{1}{\alpha} |\tilde{\xi}_t^j|$ for $0 \leq t \leq T$ and $j \in \{1, \ldots, m\}$. This will ensure that the exponential $\theta_t = \frac{d_t^j}{\tilde{M}_t}$, $0 \leq t \leq T$ does not become unbounded but remains close to 1.

Unfortunately for most derivative security valuation problems a lower bound condition of this type will usually not hold. However if we apply the above analysis to the processes $\tilde{M}^\alpha$, $\tilde{\xi}^\alpha$ and $\tilde{\xi}^\alpha_j$ given by

$$\tilde{M}^\alpha_t = \alpha + \tilde{M}_t, \quad \tilde{\xi}^\alpha_t = \tilde{\xi}_t,$$

for $0 \leq t \leq T$, $j \in \{1, \ldots, m\}$ then using the assignment $d_t^j = -\frac{\tilde{\xi}^\alpha_t}{\tilde{M}^\alpha_t}$ for $0 \leq t \leq T$, $j \in \{1, \ldots, m\}$ and the above arguments we can generally ensure that the
corresponding process \( \theta \) will be close to 1. Consequently \( \text{Var}(\tilde{M}_t \theta_t) \) will be small for the approximations \( \tilde{M}^\alpha \) and \( \tilde{e}^\beta \), \( j \in \{1, \ldots, m\} \) and \( \alpha \) sufficiently large.

We remark that, in the above construction, we could define \( \theta_t \) and \( \tilde{M}_t \) directly by equations (1.5) and (1.11). This would enable us to build estimators of the form (1.12) without a measure transformation. However we also need to verify condition (1.8) and have a means for building the process \( \tilde{M} \), so that the estimator \( \tilde{M}_t \theta_t \) based on (1.12) can be computed. In practice, for most types of financial valuation problems, this is achieved via a measure transformation. Also the machinery developed in this section is needed in the next section when we construct other variance reduced estimators.

Note that the above variance reduction procedure can be applied iteratively because the estimator \( \tilde{M} \theta \) is a martingale under \( P \) with an Ito integral representation of the form (1.12). We assume of course that the integrands in (1.12) satisfies a mean square integrability condition of the form (1.7).

Thus if we let \( M^{(1)} = \tilde{M} \theta \), then we can apply the above procedure to \( M^{(1)} \) rather than the original \( M \). The new estimator will be of the form \( M^{(2)} = M^{(1)} \theta^{(1)} \) for a suitably chosen process \( \theta^{(1)} \). In a similar fashion, additional estimators \( M^{(3)}, \ldots, M^{(k)} \) for some positive integer \( k \) can also be built.

For a large class of valuation problems in financial mathematics we can find good analytic approximations for the valuation process \( M \) and integrands \( \tilde{e}_j \), \( j \in \{1, \ldots, m\} \). Using these we can construct the estimator \( M^{(1)} \). To construct \( M^{(2)} \) we require approximations of the integrands \( \theta(\tilde{e}_j + d^i M^{(1)}) \), \( j \in \{1, \ldots, m\} \) appearing in (1.12). Analytic approximations are usually more difficult to find for this type of integrand. An effective alternative is to use discrete time methods as explained in Section 2.3 below.

### 2.2 An Alternative Variance Reduced Estimator

In this section we will compute the variance of the estimator \( M \theta^{-1} \) under the measure \( \tilde{P} \). From the relation \( \theta_t^{-1} = \frac{dP}{d\tilde{P}} \), we have

\[
E(M_t) = \int_{\Omega} M_t dP_t
\]

\[
= \int_{\Omega} M_t d\tilde{P}_t
\]

\[
= \int_{\Omega} M_t \theta_t^{-1} d\tilde{P}_t
\]

\[
= \int_{\Omega} M_t \theta_t^{-1} d\tilde{P}_T
\]

\[
= \tilde{E}(M_t \theta_t^{-1})
\]

for \( t_0 \leq t \leq T \). This expression can be compared to (1.10) for the estimator \( \tilde{M} \theta \). Since \( E(M_t) = E(M_T) \) the process \( M \theta^{-1} = \{M_t \theta_t^{-1}, \ t_0 \leq t \leq T\} \) is thus an unbiased
2.2. AN ALTERNATIVE VARIANCE REDUCED ESTIMATOR

estimator under the measure \( \tilde{P} = \tilde{P}_T \) for \( E(M_T) \) at time \( t \), \( t_0 \leq t \leq T \). This result does not use the \((\mathcal{F}, \tilde{P})\)-martingale \( \tilde{M} \) and in particular does not require condition (1.8) to be verified.

We now express the random variables \( M_t \) and \( \theta_t \) as semimartingales using the Wiener process \( \tilde{W} \) rather than \( W \). Using (1.1), (1.4) and (1.5) we have

\[
M_t = M_{t_0} + \sum_{j=1}^{m} \int_{t_0}^{t} \xi_s^j d\tilde{W}_s^j + \sum_{j=1}^{m} \int_{t_0}^{t} \xi_s^j d\tilde{W}_s^j, \tag{2.2}
\]

\[
\theta_t = 1 + \sum_{j=1}^{m} \int_{t_0}^{t} (d_s^j)^2 \theta_s d\tilde{W}_s^j + \sum_{j=1}^{m} \int_{t_0}^{t} d_s^j \theta_s d\tilde{W}_s^j. \tag{2.3}
\]

Applying the above expressions and Ito’s formula, the integral equation for \( M_t \theta_t^{-1} \) becomes

\[
M_t \theta_t^{-1} = M_{t_0} + \sum_{j=1}^{m} \int_{t_0}^{t} \theta_s^{-1} (\xi_s^j - d_s^j M_s) d\tilde{W}_s^j. \tag{2.4}
\]

Consequently \( M \theta^{-1} \) is an \((\mathcal{F}, \tilde{P})\)-martingale. We can also verify the result, previously noted, that the process \( M \theta^{-1} \) is an unbiased estimator under \( \tilde{P} \) for \( E(M_T) \) at time \( t \), \( t_0 \leq t \leq T \). If we denote by \( \tilde{\text{Var}}(M_t \theta_t^{-1}) \) the variance of the estimator \( M_t \theta_t^{-1} \) under \( \tilde{P} \), then

\[
\tilde{\text{Var}}(M_t \theta_t^{-1}) = \tilde{E} \left( \left( \sum_{j=1}^{m} \int_{t_0}^{t} \theta_s^{-2} (\xi_s^j - d_s^j M_s)^2 d\tilde{W}_s^j \right)^2 \right) \]

\[
= \int_{t_0}^{t} \tilde{E} \left( \theta_s^{-2} \sum_{j=1}^{m} (\xi_s^j - d_s^j M_s)^2 \right) ds. \tag{2.5}
\]

This should be compared to the variance of \( \tilde{M}_t \theta_t \) under \( P \) given by (1.13). If the inequality \( M_t > 0 \) holds \( \tilde{P} \) a.s. for all \( t_0 \leq t \leq T \) and we choose \( d_s^j = \xi_s^j / M_t \) for each \( j \in \{1, \ldots, m\} \), then the variance of \( M_t \theta_t^{-1} \) under \( \tilde{P} \) is reduced to zero. As previously noted the estimator \( M \theta^{-1} \) is a martingale under \( \tilde{P} \). Consequently, if the integrands in (2.4) satisfy a mean square integrability condition of the form (1.7), then we can repeat the above procedure to find a new measure \( \tilde{P}^{(1)} \) and a new exponential process \( \theta^{(1)} \) such that \( M \theta^{-1} (\theta^{(1)})^{-1} \) is an unbiased variance reduced estimator under \( \tilde{P}^{(1)} \) for \( \tilde{E}(M_T \theta_T^{-1}) = E(M_T) \). As in the previous section this methodology can be applied iteratively to build more refined estimators.

Note that compared to the estimators \( \tilde{M} \theta, \tilde{M}^{(1)} \theta^{(1)}, \ldots \) obtained under \( P \) we do not need to deal with successive iterates of the martingale process \( \tilde{M}, \tilde{M}^{(1)}, \ldots \), as the original martingale \( M \) remains the same. However we may need to construct a new measure at each iteration step.

For most derivative security pricing problems we are given a random variable \( H : \Omega \rightarrow \mathbb{R} \) and a valuation (martingale) process of the form

\[
M_t = E \left( H \mid \mathcal{F}_t \right).
\]
CHAPTER 2. VARIANCE REDUCTION TECHNIQUES

This provides a very general framework for modelling many types of asset dynamics, contingent claims and sources of uncertainty. Because our variance reduction procedure obtained under $\hat{P}$ requires only $M$ to be square integrable and they can therefore be applied to a wide class of valuation problems.

We will now consider a more explicit form for the martingale process $M$ given by (1.1). Our task will be to verify condition (1.8) and find another expression for $\text{Var}(M_t)$ as given by (1.13). Let $X^{t_0, x} = \{X^{t_0, x}_t, \ t_0 \leq t \leq T\}$ be a $d$-dimensional diffusion process which satisfies the system of stochastic differential equations (1.1.1) starting at time $t_0$ with initial value $x \in \mathbb{R}^d$. As in Section 1.1 we assume appropriate growth and smoothness conditions apply for the drift and diffusion coefficients so that the solution to (1.1.1) is unique and is a Markov process.

We also take the valuation function $u : \Gamma_0 \cup \Gamma_1 \rightarrow \mathbb{R}$ as given by (1.1.6) with $u = \hat{u}$ for some suitable choice of payoff functional $h : \Gamma_1 \rightarrow \mathbb{R}$ with $h = \hat{h}$. The regions $\Gamma_0$ and $\Gamma_1$ and stopping time $\tau$ are as given in Section 1.1. Our task is therefore to construct variance reduced estimators for the martingale $M$ given by

$$M_t = E\left(h\left(\tau, X^{t_0, x}_{t \wedge \tau}\right) \mid \mathcal{F}_t\right) = u\left(t \wedge \tau, X^{t_0, x}_{t \wedge \tau}\right) \quad (2.6)$$

for $t_0 \leq t \leq T$ (see (1.1.13)). We assume appropriate growth bounds apply for $h$ so that $M$ is square integrable under $P$.

Suppose that the process $d = (d^1, \ldots, d^m)$ has the form

$$d_t^j = d^j(t, X^{t_0, x}_t) \quad (2.7)$$

for some suitable choice of real-valued functions $d^j : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ for $t_0 \leq t \leq T$, $j \in \{1, \ldots, m\}$.

Following Kloeden & Platen (1992) we let $\tilde{X}^{t_0, x} = \{\tilde{X}^{t_0, x}_t, t_0 \leq t \leq T\}$ be a $d$-dimensional diffusion process starting at time $t_0$ with initial value $x \in \mathbb{R}^d$ and integral representation

$$\tilde{X}^{t_0, x}_t = x + \int_{t_0}^{t} a\left(s, \tilde{X}^{t_0, x}_s\right) \, ds + \sum_{j=1}^{m} \int_{t_0}^{t} b^j\left(s, \tilde{X}^{t_0, x}_s\right) \, dW^j_s. \quad (2.8)$$

We can now express $\tilde{X}^{t_0, x}$ as a semimartingale using $W$ rather than $\hat{W}$. Thus from (1.4) and (2.8) we have

$$\tilde{X}^{t_0, x}_t = x + \int_{t_0}^{t} \left[ a\left(s, \tilde{X}^{t_0, x}_s\right) - \sum_{j=1}^{m} b^j\left(s, \tilde{X}^{t_0, x}_s\right) d^j\left(s, \tilde{X}^{t_0, x}_s\right) \right] \, ds$$

$$+ \sum_{j=1}^{m} \int_{t_0}^{t} b^j\left(s, \tilde{X}^{t_0, x}_s\right) \, dW^j_s. \quad (2.9)$$

Define the stopping time $\tilde{\tau} : \Omega \rightarrow \mathbb{R}$ by

$$\tilde{\tau} = \inf\left\{ t > 0 : \left( t, \tilde{X}^{t_0, x}_t \right) \notin \Gamma_0 \right\},$$
2.2. AN ALTERNATIVE VARIANCE REDUCED ESTIMATOR

the martingale process $\tilde{M}$ by

$$
\tilde{M}_t = \tilde{E} \left( h \left( \tilde{\tau}, \tilde{X}_{\tilde{\tau}}^{t_0, \varepsilon} \right) \mid \mathcal{F}_t \right)
$$

(2.10)

and the region $\tilde{\Gamma}_1$ by

$$
\tilde{\Gamma}_1 = \left\{ \left( \tau(\omega), \tilde{X}_{\tilde{\tau}}^{t_0, \varepsilon}(\omega) \right) \in [t_0, T] \times \mathbb{R}^d : \omega \in \Omega \right\}
$$

for $t_0 \leq t \leq T$. As is the case for the martingale $M$ we assume that $\tilde{M}$ is square integrable with respect to $\tilde{P}$ and that $\tilde{\Gamma}_1 \subseteq \Gamma_1$ so that the random variable $h(\tilde{\tau}, \tilde{X}_{t_0, \varepsilon})$ is well defined by continuity of the sample paths of $\tilde{X}_{t_0, \varepsilon}$.

Using equations (1.1.1) and (2.8) we see that $X_{t_0, \varepsilon}$ and $\tilde{X}_{t_0, \varepsilon}$ are weak solutions of the same system of stochastic differential equations. Consequently from (2.6) and (2.10) we see that

$$
\tilde{E}(\tilde{M}_T) = \tilde{E} \left( h \left( \tilde{\tau}, \tilde{X}_{\tilde{\tau}}^{t_0, \varepsilon} \right) \right) = E \left( h \left( \tau, X_{t_0, \varepsilon}^{t_0, \varepsilon} \right) \right) = E(M_T)
$$

(2.11)

which verifies condition (1.8).

Let $\tilde{u}: \Gamma_0 \cup \Gamma_1 \rightarrow \mathbb{R}$ be some approximation to $u$ with

$$
\tilde{u} \left( \tilde{\tau}, \tilde{X}_{t_0, \varepsilon}^{t_0, \varepsilon} \right) = u \left( \tilde{\tau}, \tilde{X}_{t_0, \varepsilon}^{t_0, \varepsilon} \right).
$$

(2.12)

We assume that the functions $u$ and $\tilde{u}$ are of class $C^{1,2}$; that is these functions have continuous first order time, and second order spatial partial derivatives.

To simplify the notation in what follows we will use the operators $L^0$, $\tilde{L}^0$ and $L^j$, $j \in \{1, \ldots, m\}$ defined by

$$
L^0 = \frac{\partial}{\partial t} + \sum_{i=1}^{d} a_i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,k=1}^{d} b^{i,j} b^{k,j} \frac{\partial^2}{\partial x_i \partial x_k},
$$

$$
\tilde{L}^0 = \frac{\partial}{\partial t} + \sum_{i=1}^{d} \left( a_i - \sum_{j=1}^{m} b^{j,i} d^{j} \right) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,k=1}^{d} b^{i,j} b^{k,j} \frac{\partial^2}{\partial x_i \partial x_k},
$$

$$
L^j = \sum_{i=1}^{d} b^{i,j} \frac{\partial}{\partial x_i}.
$$

The Kolmogorov backward equation applied to $u$ in the region $\Gamma_0$ can be expressed in the form

$$
L^0 u(s, X^{t_0, \varepsilon}) = 0
$$

(2.13)

and

$$
L^0 u(s, \tilde{X}_{t_0, \varepsilon}^{t_0, \varepsilon}) = 0
$$

for $(s, X^{t_0, \varepsilon}) \in \Gamma_0$ and $(s, \tilde{X}_{t_0, \varepsilon}^{t_0, \varepsilon}) \in \Gamma_0$, respectively.

If we expand $u(t \wedge \tilde{\tau}, \tilde{X}_{t \wedge \tilde{\tau}}^{t_0, \varepsilon})$ using Ito's formula for semimartingales and (2.8), we have from (2.13) the relation

$$
u \left( t \wedge \tilde{\tau}, \tilde{X}_{t \wedge \tilde{\tau}}^{t_0, \varepsilon} \right) = u(t_0, \varepsilon) + \sum_{j=1}^{m} \int_{t_0}^{t \wedge \tilde{\tau}} L^j u \left( s, \tilde{X}_{s}^{t_0, \varepsilon} \right) d\tilde{W}_s
$$

(2.14)
CHAPTER 2. VARIANCE REDUCTION TECHNIQUES

Writing this as an integral equation using the Wiener process \( W \) rather than \( \tilde{W} \) we obtain from (1.4) the expression

\[
\begin{align*}
  u(t, \pi, \tilde{X}_{t+t}) &= u(t_0, \pi) - \sum_{j=1}^{m} \int_{t_0}^{t+t} d^j (s, \tilde{X}_{s+t}) L^j u(s, \tilde{X}_{s+t}) \, ds \\
  &\quad + \sum_{j=1}^{m} \int_{t_0}^{t+t} L^j u(s, \tilde{X}_{s+t}) \, dW^j_s.
\end{align*}
\] (2.15)

In a similar manner we can determine \( \bar{u}(t, \pi, \tilde{X}_{t+t}) \) using Ito's formula for semimartingales and (2.8) to obtain

\[
\begin{align*}
  \bar{u}(t, \pi, \tilde{X}_{t+t}) &= u(t_0, \pi) + \int_{t_0}^{t+t} L^0 \bar{u}(s, \tilde{X}_{s+t}) \, ds \\
  &\quad + \sum_{j=1}^{m} \int_{t_0}^{t+t} L^j \bar{u}(s, \tilde{X}_{s+t}) \, dW^j_s.
\end{align*}
\] (2.16)

which, from (1.4), can be written in the form

\[
\begin{align*}
  \bar{u}(t, \pi, \tilde{X}_{t+t}) &= u(t_0, \pi) + \int_{t_0}^{t+t} \left( L^0 \bar{u}(s, \tilde{X}_{s+t}) \\
  &\quad - \sum_{j=1}^{m} \int_{t_0}^{t+t} d^j (s, \tilde{X}_{s+t}) L^j \bar{u}(s, \tilde{X}_{s+t}) \right) \, ds \\
  &\quad + \sum_{j=1}^{m} \int_{t_0}^{t+t} L^j \bar{u}(s, \tilde{X}_{s+t}) \, dW^j_s.
\end{align*}
\] (2.17)

Applying, once again, Ito's formula for semimartingales using (2.15), (2.17) and (1.5) we have the relations

\[
\begin{align*}
  u(t, \pi, \tilde{X}_{t+t}) \theta_{t+t} &= u(t_0, \pi) \\
  &\quad + \sum_{j=1}^{m} \int_{t_0}^{t+t} \theta_s \left[ L^j u(s, \tilde{X}_{s+t}) + u(s, \tilde{X}_{s+t}) d^j (s, \tilde{X}_{s+t}) \right] \, dW^j_s,
\end{align*}
\] (2.18)

\[
\begin{align*}
  \bar{u}(t, \pi, \tilde{X}_{t+t}) \theta_{t+t} &= \bar{u}(t_0, \pi) + \int_{t_0}^{t+t} L^0 \bar{u}(s, \tilde{X}_{s+t}) \theta_s \, ds \\
  &\quad + \sum_{j=1}^{m} \int_{t_0}^{t+t} \theta_s \left[ L^j \bar{u}(s, \tilde{X}_{s+t}) + \bar{u}(s, \tilde{X}_{s+t}) d^j (s, \tilde{X}_{s+t}) \right] \, dW^j_s
\end{align*}
\] (2.19)

for \((t, X_t) \in \Gamma_0\) and \((t, \tilde{X}_t) \in \Gamma_0\), respectively. These equations can also be obtained from Ito expansions using (2.9) and (1.5).
2.2. AN ALTERNATIVE VARIANCE REDUCED ESTIMATOR

Combining (2.10), (2.14) and the equality 
\[ h(\tau, X_{t_0}^{\tau, \bar{x}}) = u(\tau, X_{t_0}^{\tau, \bar{x}}) \]
we see that
\[
\tilde{M}_t = \tilde{E} \left( h(\tau, X_{t_0}^{\tau, \bar{x}}) \mid \mathcal{F}_t \right) \\
= \tilde{E} \left( u(\tau, X_{t_0}^{\tau, \bar{x}}) \mid \mathcal{F}_t \right) \\
= u(t_0, \bar{x}) + \sum_{j=1}^{m} \int_{t_0}^{t} L^j u\left(s, X_s^{t_0, \bar{x}}\right) d\bar{W}_s^j \\
= u\left(t \wedge \tilde{\tau}, X_{t \wedge \tilde{\tau}}^{t_0, \bar{x}}\right).
\]
(2.20)

This relation can also be derived using more general arguments as given in Section 1.1, see in particular (1.1.13).

If we expand \( u(\tau, X_{t_0}^{\tau, \bar{x}}) \) using Ito's formula for semimartingales, (2.13) and (1.1.1), then we have
\[
u(\tau, X_{t_0}^{\tau, \bar{x}}) = u(t_0, \bar{x}) + \sum_{j=1}^{m} \int_{t_0}^{\tau} L^j u\left(s, X_s^{t_0, \bar{x}}\right) d\bar{W}_s^j.
\]

Consequently from (2.20) and taking the expectation, under \( P \), of both sides of this relation we see that
\[
E(M_T) = E \left( u\left(\tau, X_{t_0}^{\tau, \bar{x}}\right) \right) = u(t_0, \bar{x}).
\]
(2.21)

Applying this result and (2.18) we can infer that
\[
E \left( u \left(\tau \wedge \tilde{\tau}, X_{t \wedge \tilde{\tau}}^{t_0, \bar{x}}\right) \theta_{t \wedge \tilde{\tau}} \right) = u(t_0, \bar{x}) = E(M_T).
\]
(2.22)

This expression shows that the process \( \{u(\tau \wedge \tilde{\tau}, X_{t \wedge \tilde{\tau}}^{t_0, \bar{x}}) \theta_{t \wedge \tilde{\tau}}, t_0 \leq t \leq T\} \) is an unbiased estimator under \( P \) for \( E(M_T) \) at time \( t \wedge \tilde{\tau} \).

Furthermore, using (2.12) and (2.22) with \( t = T \) we have
\[
E \left( \bar{u}(\tau, X_{t_0}^{\tau, \bar{x}}) \theta_{\tilde{\tau}} \right) = E \left( u \left(\tau, X_{t_0}^{\tau, \bar{x}}\right) \theta_{\tilde{\tau}} \right) = u(t_0, \bar{x}) = E(M_T)
\]
so that the random variable \( \bar{u}(\tilde{\tau}, X_{t_0}^{\tau, \bar{x}}) \theta_{\tilde{\tau}} \) is an unbiased estimator for \( E(M_T) \) at time \( \tilde{\tau} \).

If we assume that \( \bar{u}(t, x) > 0 \) for all \( (t, x) \in \Gamma_0 \) and we choose the function \( d = (d^1, \ldots, d^m) \) by
\[
d^j(t, x) = -L^j \bar{u}(t, x)/\bar{u}(t, x)
\]
for \((t, x) \in \Gamma_0 \) and \( j \in \{1, \ldots, m\} \), then using this choice for \( d \), together with equation (2.19), we can compute the variance of \( \bar{u}(\tilde{\tau}, X_{t_0}^{\tau, \bar{x}}) \theta_{\tilde{\tau}} \) under \( P \) by the relation
\[
\text{Var} \left( \bar{u}(\tilde{\tau}, X_{t_0}^{\tau, \bar{x}}) \theta_{\tilde{\tau}} \right) = \text{Var} \left( \int_{t_0}^{\tilde{\tau}} L^0 \bar{u}(s, X_s^{t_0, \bar{x}}) \theta_s ds \right).
\]
Using (2.13), this formula can also be written in the form
\[
\text{Var} \left( \bar{u}(\tilde{\tau}, X_{t_0}^{\tau, \bar{x}}) \theta_{\tilde{\tau}} \right) = \text{Var} \left( \int_{t_0}^{\tilde{\tau}} \left( L^0 \bar{u}(s, X_s^{t_0, \bar{x}}) - L^0 u(s, X_s^{t_0, \bar{x}}) \right) \theta_s ds \right).
\]
(2.24)
which is a more convenient expression of the variance for some applications. This should be compared with the variance obtained from (2.18) which is given by the equation

\[
\text{Var}\left(\bar{u}\left(\tau, \bar{X}^{t_0 \rightarrow x}\right) \theta_\tau\right)
\]

\[
= E\left(\sum_{j=1}^{m} \int_{t_0}^{\tau} \theta_s \left[ L^j u\left(s, \bar{X}^{t_0 \rightarrow x}\right) - L^j \bar{u}\left(s, \bar{X}^{t_0 \rightarrow x}\right) \frac{u\left(s, \bar{X}^{t_0 \rightarrow x}\right)}{\bar{u}\left(s, \bar{X}^{t_0 \rightarrow x}\right)} \right] dW^j_s\right)^2
\]

\[
= E\left(\int_{t_0}^{\tau} \theta_s^2 \sum_{j=1}^{m} \left[ L^j u\left(s, \bar{X}^{t_0 \rightarrow x}\right) - L^j \bar{u}\left(s, \bar{X}^{t_0 \rightarrow x}\right) \frac{u\left(s, \bar{X}^{t_0 \rightarrow x}\right)}{\bar{u}\left(s, \bar{X}^{t_0 \rightarrow x}\right)} \right]^2 ds\right).
\]

(2.25)

This relation can also be obtained by substituting (2.23) into (1.13). Thus for a suitable choice of the function \(d\) we can make the random variable \(\bar{u}(\tau, \bar{X}^{t_0 \rightarrow x})\theta_\tau\) an unbiased variance reduced estimator for \(E(M_T)\) at time \(\tau\).

These methods provide the basis for powerful variance reduction and error minimization procedures. The explicit formulas for the variance of estimators given by (1.13), (2.5), (2.24) and (2.25) mean that exact controls of the variance can be built as well as being of theoretical interest.

We emphasize that the smoothness conditions used in the derivation of (2.24) and (2.25) are not required in the more general formulations leading to (1.13) and (2.5). Also, condition (1.8), required for the estimator \(\tilde{M}\) \(\theta\), can often be easily verified in practice for reasonable choices of \(\tilde{M}\). The approximating function \(\bar{u} : \Gamma_0 \cup \Gamma_1 \rightarrow \mathbb{R}\) can be any choice which satisfies the terminal payoff condition (2.12) and is \(C^{1,2}\) smooth. This approximating function may be constructed iteratively using some combination of analytic solutions, numerical methods and/or appropriate interpolation routines. For example, it may correspond to the pricing function for a similar model, where analytic or more accurate pricing can be made.

### 2.3 Discrete Time Variance Reduced Estimators

In this section we consider the application of discrete-time numerical methods that can be used to either approximate a given continuous time estimator or to build new (discrete) variance reduced estimators.

To illustrate the principles involved we will first consider the problem of discretizing the product estimator \(\tilde{M}\) \(\theta\) for \(E(M_T)\) considered in Section 2.1, where \(M\) is a given square integrable \((\mathcal{F}, \mathcal{P})\)-martingale with representation (1.1) and satisfying the integrability condition (1.2). We take \(\tilde{M}\) to be an \((\mathcal{F}, \tilde{\mathcal{P}})\)-martingale expressible in the form (1.6) and satisfying (1.7), where \(\theta\) is the Radon-Nikodym derivative given by (1.5). The process \(d = (d^1, \ldots, d^m)\) is assumed to be a free parameter process, satisfying (1.3)
and which is used to construct $\theta$ to reduce the variance. Note that the definitions of unbiased and variance reduced unbiased estimators can be formulated for discrete processes in a similar fashion to that described in Section 2.1.

Let $(t)_\Delta$ be an equi-spaced time discretization of the interval $[t_0, T]$ of the form

$$t_0 < t_1 < \ldots < t_N = T$$

with step size

$$\Delta = \frac{(T - t_0)}{N}, \quad N = 1, 2, \ldots$$

Let $M^\Delta = \{M^\Delta_k \mid k \in \{0, \ldots, N - 1\}\}$, $\tilde{M}^\Delta_k$ and $\theta^\Delta_k$, represent Euler (weak) approximations of the processes $M$, $\tilde{M}$ and $\theta$ given by (1.1), (1.11) and (1.5), respectively, of the form

$$M^\Delta_{k+1} = M^\Delta_k + \sum_{j=1}^{m} \xi^j_k \Delta W^j_k,$$  \hfill (3.2)

$$\tilde{M}^\Delta_{k+1} = \tilde{M}^\Delta_k - \sum_{j=1}^{m} d^j_k \xi^j_k \Delta + \sum_{j=1}^{m} \xi^j_k \Delta W^j_k,$$  \hfill (3.3)

$$\theta^\Delta_{k+1} = \theta^\Delta_k + \sum_{j=1}^{m} d^j_k \theta^\Delta_k \Delta W^j_k,$$  \hfill (3.4)

for $k \in \{0, \ldots, N - 1\}$ with initial values $M^\Delta_0 = M_0$, $\tilde{M}^\Delta_0 = M_0$ and $\theta^\Delta_0 = 1$. We will take the increments $\Delta W_k = (\Delta W^{1}_k, \ldots, \Delta W^{m}_k)$, $k \in \{0, \ldots, N - 1\}$, to be either a collection of independent Gaussian random variables with expectation 0 and variance $\Delta$ under $P$ or a collection of multipoint random variables again with expectation 0 and variance $\Delta$ under $P$. We require values for only the first and second moments of these increments in what follows. For example we could take a set of two point random variables with

$$P \left( \Delta W^j_k = \pm \sqrt{\Delta} \right) = \frac{1}{2}$$  \hfill (3.5)

for $k \in \{0, \ldots, N - 1\}$ and $j \in \{1, \ldots, m\}$, see for example Kloeden & Platen (1992) and Hofmann, Platen & Schweizer (1992).

Multiplying (3.3) by (3.4) the product $M^\Delta_{k+1} \theta^\Delta_{k+1}$ can be expressed in the form

$$M^\Delta_{k+1} \theta^\Delta_{k+1} = \tilde{M}^\Delta_0 \theta^\Delta_0 - \sum_{j=1}^{m} d^j_k \xi^j_k \Delta + \sum_{j=1}^{m} \xi^j_k \theta^\Delta_k \Delta W^j_k$$

$$+ \sum_{j=1}^{m} \tilde{M}^\Delta_k d^j_k \theta^\Delta_k \Delta W^j_k - \left( \sum_{j=1}^{m} d^j_k \xi^j_k \Delta \right) \left( \sum_{j=1}^{m} d^j_k \theta^\Delta_k \Delta W^j_k \right)$$

$$+ \left( \sum_{j=1}^{m} \xi^j_k \Delta W^j_k \right) \left( \sum_{j=1}^{m} d^j_k \theta^\Delta_k \Delta W^j_k \right).$$  \hfill (3.6)
We can write
\[
\left( \sum_{j=1}^{m} \hat{\xi}_k^j \Delta W_k^j \right) \left( \sum_{j=1}^{m} d_k^j \theta_k^j \Delta W_k^j \right) = \sum_{j=1}^{m} \hat{\xi}_k^j d_k^j \theta_k^j (\Delta W_k^j)^2 + \sum_{j_1, j_2=1 \atop j_1 \neq j_2}^{m} \hat{\xi}_k^{j_1} d_k^{j_1} \theta_k^{j_1} \Delta W_k^{j_1} \Delta W_k^{j_2}
\]
(3.7)
and consequently using the fact that \( E((\Delta W_k^j)^2) = \Delta \) and \( E(\Delta W_k^j \Delta W_k^\ell) = 0 \) for \( j, \ell \in \{1, \ldots, m\} \), \( k \in \{0, \ldots, N - 1\} \) and \( j \neq \ell \), the latter following from the independence property of the increments \( \Delta W_k^i \), we have the relation
\[
E \left( \left( \sum_{j=1}^{m} \hat{\xi}_k^j \Delta W_k^j \right) \left( \sum_{j=1}^{m} d_k^j \theta_k^j \Delta W_k^j \right) \right) = E \left( \theta_k^2 \sum_{j=1}^{m} \hat{\xi}_k^j d_k^j \right) \Delta.
\]
From this result, again the independence property of the increments \( \Delta W_k^j \), and taking expectations of both sides of (3.6) and the initial value \( \theta_0^2 = 1 \) we have
\[
E \left( \tilde{M}_{k+1}^\Delta \theta_{k+1}^2 \right) = E \left( \tilde{M}_k^\Delta \theta_k^2 \right) = E \left( \tilde{M}_0^\Delta \theta_0^2 \right) = M_0.
\]
(3.8)
for \( k \in \{0, \ldots, N-1\} \). This means that \( \tilde{M}^\Delta \theta^2 \) is an unbiased estimator for \( E(M_N) = M_0 \) at time \( t_k, \ k \in \{0, \ldots, N-1\} \).

Applying once again the independence property of the increments \( \Delta W_k^j \) we see that for all \( k, \ell \in \{0, \ldots, N-1\}, \ j_1, j_2 \in \{1, \ldots, m\} \) with \( \ell < k \) the random variables \( \Delta W_k^{j_1} \) and \( \Delta W_k^{j_2} \) will both be independent of any Borel measurable functions of the variates \( d_k^{j_1}, \hat{\xi}_k^{j_1}, \theta_k^{j_1}, M_k, d_k^{j_2}, \hat{\xi}_k^{j_2}, \theta_k^{j_2}, M_\ell, \Delta W_k^{j_1}, \Delta W_\ell^{j_2} \). In particular \( (\Delta W_k^{j_1})^2 - \Delta \) will be independent of \( d_k^{j_1} \hat{\xi}_k^{j_1} \theta_k^{j_1} d_k^{j_2} \hat{\xi}_k^{j_2} \theta_k^{j_2} [(\Delta W_k^{j_1})^2 - \Delta] \).

Therefore if we denote by \( C_{k, \ell} \) the covariance
\[
C_{k, \ell} = \text{Cov} \left( \tilde{M}_{k+1}^\Delta \theta_{k+1}^2 - \tilde{M}_k \theta_k^2, \tilde{M}_{\ell+1}^\Delta \theta_{\ell+1}^2 - \tilde{M}_\ell \theta_\ell^2 \right),
\]
then from (3.6), (3.7), (3.8) and the independence properties stated above and for integers \( \ell < k \), we have, after simplification, the equations
\[
C_{k, \ell} = E \left( \left( \sum_{j=1}^{m} d_k^{j_1} \hat{\xi}_k^{j_1} \theta_k^{j_1} \left[ (\Delta W_k^{j_1})^2 - \Delta \right] \right) \left( \sum_{j=1}^{m} d_\ell^{j_2} \hat{\xi}_\ell^{j_2} \theta_\ell^{j_2} \left[ (\Delta W_\ell^{j_2})^2 - \Delta \right] \right) \right)
= \sum_{j_1, j_2=1}^{m} E \left( d_k^{j_1} \hat{\xi}_k^{j_1} \theta_k^{j_1} d_\ell^{j_2} \hat{\xi}_\ell^{j_2} \theta_\ell^{j_2} \left[ (\Delta W_k^{j_1})^2 - \Delta \right] \left[ (\Delta W_\ell^{j_2})^2 - \Delta \right] \right)
= \sum_{j_1, j_2=1}^{m} E \left( d_k^{j_1} \hat{\xi}_k^{j_1} \theta_k^{j_1} d_\ell^{j_2} \hat{\xi}_\ell^{j_2} \theta_\ell^{j_2} \left[ (\Delta W_k^{j_1})^2 - \Delta \right] \right) E \left( \left[ (\Delta W_\ell^{j_2})^2 - \Delta \right] \right)
= 0.
\]
(3.9)
Also, for any \( k \in \{0, \ldots, N - 1\} \) and using the initial value \( \theta_0 = 1 \) we can write
\[
\bar{M}_k^\Delta \theta_k^\Delta = \sum_{\ell=0}^{k-1} \left( \bar{M}_{\ell+1}^\Delta \theta_{\ell+1}^\Delta - \bar{M}_\ell^\Delta \theta_\ell^\Delta \right) + M_0
\]
so that applying (3.9), the variance of \( \bar{M}_k^\Delta \theta_k^\Delta, \ k \in \{1, \ldots, N\} \) can be computed from the formula
\[
\text{Var} \left( \bar{M}_k^\Delta \theta_k^\Delta \right) = \sum_{\ell=0}^{k-1} E \left( \left( \bar{M}_{\ell+1}^\Delta \theta_{\ell+1}^\Delta - \bar{M}_\ell^\Delta \theta_\ell^\Delta \right)^2 \right).
\] (3.10)
For small values of \( \Delta \) we can ignore all terms in the product \( (\bar{M}_{\ell+1}^\Delta \theta_{\ell+1}^\Delta - \bar{M}_\ell^\Delta \theta_\ell^\Delta)^2 \), \( \ell \in \{1, \ldots, k - 1\} \) obtained from (3.6) whose expectation under \( P \) is 0 or of order \( \Delta^q \) for \( q \geq \frac{3}{2} \). This means that all terms on the right hand side of (3.6) except
\[
\sum_{j=1}^{m} \hat{\xi}_k^j \Delta W_k^j \quad \text{and} \quad \sum_{j=1}^{m} \bar{M}_k^\Delta d_k^j \theta_k^\Delta \Delta W_k^j
\]
can be ignored, and using the property \( E((\Delta W_k^j)^2) = \Delta \) we can approximate the variance of \( \bar{M}_k^\Delta \theta_k^\Delta, \ k \in \{1, \ldots, N\} \) by
\[
\text{Var}(\bar{M}_k^\Delta \theta_k^\Delta) \approx \sum_{\ell=0}^{k-1} E \left( \left( \theta_\ell^\Delta \right)^2 \left( \sum_{j=1}^{m} (\hat{\xi}_k^j + d_k^j \bar{M}_\ell^\Delta)^2 \right) \right) \Delta.
\] (3.11)
As in Section 2.1 we let \( \hat{\xi}_\ell = (\hat{\xi}_\ell^1, \ldots, \hat{\xi}_\ell^m) \) and \( \bar{M}_\ell^\Delta \) be approximations for \( \xi_\ell = (\hat{\xi}_\ell^1, \ldots, \hat{\xi}_\ell^m) \) and \( M_\ell^\Delta \), \( \ell \in \{0, \ldots, N - 1\} \), respectively. If the inequality \( \bar{M}_\ell^\Delta > 0 \) holds \( P\text{-a.s.} \) for all \( \ell \in \{0, \ldots, N - 1\} \) we can choose \( d_k^j = -\frac{\hat{\xi}_k^j}{\bar{M}_\ell^\Delta} \). Substituting this value into (3.11) we obtain
\[
\text{Var}(\bar{M}_k^\Delta \theta_k^\Delta) \approx \sum_{\ell=0}^{k-1} E \left( \left( \theta_\ell^\Delta \right)^2 \sum_{j=1}^{m} \left( \hat{\xi}_k^j - \hat{\xi}_\ell^j \bar{M}_\ell^\Delta \bar{M}_\ell^\Delta \right)^2 \right) \Delta
\] (3.12)
for \( k \in \{1, \ldots, N\} \).

The variance formulas (3.11) and (3.12) for the estimator \( \bar{M}^\Delta \theta^\Delta \) should be compared to the variance of the continuous time version of the estimator given by (1.13) and (1.15), respectively. Note that with the above formulation we have used only very basic properties of discrete time stochastic processes. In the case where the increments \( \Delta W_k^j \) are two-point random variables with probabilities given by (3.5) these calculations can be further simplified since in this case \( (\Delta W_k^j)^2 = \Delta \) for \( j \in \{1, \ldots, m\} \), \( k \in \{0, \ldots, N - 1\} \). Similar expressions for the variance given by (3.11) and (3.12) are obtained if we replace the Euler approximations (3.2) to (3.4) by other higher order weak approximations, although for these approximations the above computations become more involved and complex.

The variance reduction procedure discussed in Section 2.2 can also be formulated in a discrete-time framework. However in this case we need to be more careful regarding
how the problem should be formulated as some choices lead to biased estimates of $E(M_T)$.

To see how an unbiased estimator can be constructed let $M^\Delta$ and $\theta^\Delta$ be Euler (weak) approximations of the processes $M$ and $\theta$ as given by (3.2) and (3.4), respectively. Observing the form of (3.4) and the initial condition $\theta_0^\Delta = 1$ an induction argument shows that

$$\theta_k^\Delta = \prod_{\ell=0}^{k-1} \left( 1 + \sum_{j=1}^m d_{\ell}^j \Delta W_{\ell}^j \right)$$

(3.13)

for $k \in \{1, \ldots, N-1\}$.

Define the measure $\bar{P}^\Delta : \mathcal{F}_T \to [0, 1]$ by

$$\bar{P}^\Delta(A) = \int_A \theta_k^\Delta dP$$

(3.14)

for $A \in \mathcal{F}_T$.

We will now assume that the increments $\Delta W_k = (\Delta W_k^1, \ldots, \Delta W_k^m)$, $k \in \{0, \ldots, N-1\}$ form a collection of two-point random variables with mean 0 and variance $\Delta$ under $P$ with probabilities given by (3.5). In addition we assume that

$$\left| \sum_{j=1}^m d_{k}^j \Delta W_{k}^j \right| \leq K_1 \sqrt{\Delta}$$

(3.15)

for some constant $K_1 \in \mathbb{R}^+$ for all $k \in \{0, \ldots, N-1\}$ and any discretization grid $(t)$. This assumption clearly depends on the properties of the random variables $d_{k}^j$, $j \in \{1, \ldots, m\}$, $k \in \{0, \ldots, N-1\}$. For a discussion on how the growth of these variates can be constrained see the commentary following (1.16) in Section 2.1. We restrict our attention to two point random increments because it simplifies the calculations in what follows, however this analysis can, in fact, be extended to a wider class of multipoint random variables.

Applying (3.13) and the independence property of the increments $\Delta W_k^j$ we see that $E(\theta_k^\Delta) = 1$. Also for sufficiently fine discretization grids $(t)$, (3.15) shows that $\theta_k^\Delta > 0$ for all $k \in \{0, \ldots, N-1\}$. This means that both the probability measure $\bar{P}^\Delta$ and the quotient $M_k^\Delta/\theta_k^\Delta$, $k \in \{0, \ldots, N-1\}$ are well defined.

Using the definition of $\bar{P}^\Delta$ given by (3.14), again the independence property of the increments $\Delta W_k^j$, and (3.13) we can infer that

$$E^\Delta \left( \frac{M_k^\Delta}{\theta_k^\Delta} \right) = E \left( \frac{M_k^\Delta}{\theta_k^\Delta} \right)$$

$$= E(M_k^\Delta) E \left( \prod_{\ell=k}^{N-1} \left( 1 + \sum_{j=1}^m d_{\ell}^j \Delta W_{\ell}^j \right) \right)$$

$$= E(M_k^\Delta)$$

$$= M_0$$

(3.16)
for $k \in \{0, \ldots, N\}$, where $\bar{E}^\Delta$ denotes expectation with respect to the measure $\bar{P}^\Delta$. This shows that $M_k^\Delta / \theta_k^\Delta$ is an unbiased estimator under $\bar{P}^\Delta$ for $E(M_N) = M_0$ at time $t_k$, $k \in \{0, \ldots, N\}$.

It is important to know the probabilities of the outcomes of $\Delta W_k = (\Delta W^1_k, \ldots, \Delta W^m_k), k \in \{0, \ldots, N - 1\}$ if we undertake a Monte Carlo simulation to approximate the value $E(M_N) = M_0$ using the estimator $M_N^\Delta / \theta_N^\Delta$ under the measure $\bar{P}^\Delta$. To determine these probabilities we let $A^*_k = \{\omega : \Delta W^j_k = +\sqrt{\Delta}\}$ for $j \in \{1, \ldots, m\}, k \in \{0, \ldots, N - 1\}$. From the definition of $\bar{P}^\Delta$ given by (3.14), the independence property of the increments $\Delta W^p_k$, and (3.13) we have

$$
\bar{P}^\Delta (A^*_k) = \bar{E}^\Delta \left(1_{A^*_k}\right)
= E \left(1_{A^*_k} \theta_N^\Delta \right)
= E \left(1_{A^*_k} \prod_{t=1}^{N-1} \left(1 + \sum_{p=1}^m d^p_t \Delta W^p_t\right)\right)
= E \left(1_{A^*_k} \left(1 + \sum_{p=1}^m d^p_t \Delta W^p_t\right)\right)
= \frac{1 + d^j_k \sqrt{\Delta}}{2}.
$$

A similar expression holds for $\bar{P}^\Delta (\{\omega : \Delta W^j_k = -\sqrt{\Delta}\})$. These results can be summarized in the form

$$
\bar{P}^\Delta (\Delta W^j_k = \pm \sqrt{\Delta}) = \left(\frac{1 \pm \sqrt{\Delta} d^j_k}{2}\right).
$$

(3.17)

These probabilities can also be used to show that

$$
\bar{E}^\Delta (\Delta W^j_k) = d^j_k \Delta,
\bar{E}^\Delta ((\Delta W^j_k)^2) = \Delta,
$$

(3.18)

for $j \in \{1, \ldots, m\}, 0 \leq k \leq N$.

We now introduce the random variables $\eta^\Delta_k, \phi^\Delta_k, \psi^\Delta_k$ defined by

$$
\eta^\Delta_k = \sum_{j=1}^m \xi^j_k \Delta W^j_k,
\psi^\Delta_k = \sum_{j=1}^m d^j_k \Delta W^j_k,
\phi^\Delta_k = \eta^\Delta_k - M_k^\Delta \psi^\Delta_k,
$$
for \( k \in \{1, \ldots, N - 1\} \). From the definition of \( \psi_k^\Delta \), equation (3.13) can be expressed in the form

\[
\theta_k^\Delta = \prod_{\ell=0}^{k-1} (1 + \psi_\ell^\Delta)
\]  

so that \( \theta_{k+1}^\Delta = \theta_k^\Delta (1 + \psi_k^\Delta) \). Combining this result and the Euler approximation (3.2) we can write for \( k \in \{0, \ldots, N - 1\} \)

\[
\frac{M_{k+1}^\Delta}{\theta_{k+1}^\Delta} - \frac{M_k^\Delta}{\theta_k^\Delta} = \frac{M_{k+1}^\Delta}{\theta_{k+1}^\Delta} - \frac{M_k^\Delta}{\theta_k^\Delta} (1 + \psi_k^\Delta)
\]

\[
= \frac{\psi_k^\Delta}{\theta_{k+1}^\Delta}.
\]  

(3.20)

For integers \( k, \ell \in \{0, \ldots, N - 1\} \), let

\[
\tilde{\text{Cov}}^\Delta \left( \frac{\phi_k^\Delta}{\theta_{k+1}^\Delta}, \frac{\phi_\ell^\Delta}{\theta_{\ell+1}^\Delta} \right) = \tilde{E}^\Delta \left( \left[ \frac{\phi_k^\Delta}{\theta_{k+1}^\Delta} - \tilde{E} \left( \frac{\phi_k^\Delta}{\theta_{k+1}^\Delta} \right) \right] \left[ \frac{\phi_\ell^\Delta}{\theta_{\ell+1}^\Delta} - \tilde{E} \left( \frac{\phi_\ell^\Delta}{\theta_{\ell+1}^\Delta} \right) \right] \right)
\]

denote the covariance of the random variables \( \phi_k^\Delta / \theta_{k+1}^\Delta \) and \( \phi_\ell^\Delta / \theta_{\ell+1}^\Delta \) under the measure \( \tilde{P}^\Delta \). From (3.16) and (3.20) we see that

\[
\tilde{E}^\Delta \left( \frac{\phi_k^\Delta}{\theta_{k+1}^\Delta} \right) = \tilde{E}^\Delta \left( \frac{M_{k+1}^\Delta}{\theta_{k+1}^\Delta} \right) - \tilde{E}^\Delta \left( \frac{M_k^\Delta}{\theta_k^\Delta} \right) = 0.
\]  

(3.21)

Consequently applying the representation (3.19) we have for integers \( k, \ell \in \{0, \ldots, N - 1\} \) with \( \ell < k \)

\[
\tilde{\text{Cov}}^\Delta \left( \frac{\phi_k^\Delta}{\theta_{k+1}^\Delta}, \frac{\phi_\ell^\Delta}{\theta_{\ell+1}^\Delta} \right) = \tilde{E}^\Delta \left( \frac{\phi_k^\Delta}{\theta_{k+1}^\Delta} \frac{\phi_\ell^\Delta}{\theta_{\ell+1}^\Delta} \right)
\]

\[
= \tilde{E}^\Delta \left( \frac{\phi_k^\Delta}{\theta_{k+1}^\Delta} \frac{\phi_\ell^\Delta}{\theta_{\ell+1}^\Delta} \right)
\]

\[
= \tilde{E} \left( \frac{\eta_k^\Delta - M_k^\Delta \psi_k^\Delta}{\theta_{k+1}^\Delta} \frac{\phi_\ell^\Delta}{\theta_{\ell+1}^\Delta} \right) \frac{\theta_\ell^\Delta}{\theta_{\ell+1}^\Delta}.
\]  

(3.22)

Noting that the increments \( \Delta W_{k,j}^\Delta, j \in \{1, \ldots, m\} \) appear only in the terms \( \eta_k^\Delta \) and \( \psi_k^\Delta \) and not in the terms \( M_k^\Delta, \theta_{k+1}^\Delta, \phi_k^\Delta \) and \( \theta_\ell^\Delta / \theta_{\ell+1}^\Delta = \prod_{p=k+1}^{N-1} (1 + \psi_p^\Delta) \), and using the independence property of the increments \( \Delta W_{k,j}^\Delta \), we can infer from (3.22) that for \( k, \ell \in \{0, \ldots, N - 1\} \) with \( \ell < k \)

\[
\tilde{\text{Cov}}^\Delta \left( \frac{\phi_k^\Delta}{\theta_{k+1}^\Delta}, \frac{\phi_\ell^\Delta}{\theta_{\ell+1}^\Delta} \right) = E(\eta_k^\Delta) E \left( \frac{\psi_\ell^\Delta}{\theta_{\ell+1}^\Delta} \left( \frac{\theta_k^\Delta}{\theta_{k+1}^\Delta} \right) \right)
\]

\[
- E(\psi_k^\Delta) E \left( \frac{M_k^\Delta \phi_\ell^\Delta}{\theta_{\ell+1}^\Delta} \left( \frac{\theta_k^\Delta}{\theta_{k+1}^\Delta} \right) \right)
\]

\[
= 0.
\]  

(3.23)
2.3. DISCRETE TIME VARIANCE REDUCED ESTIMATORS

Also for \( k \in \{0, \ldots, N - 1\} \) we can write \( M_k^\Delta / \theta_k^\Delta = \sum_{\ell=0}^{k-1} (M_{\ell+1}^\Delta / \theta_{\ell+1}^\Delta - M_\ell^\Delta / \theta_\ell^\Delta) + M_0 \) since \( \theta_0^\Delta = 1 \). Combining (3.20), (3.21) and (3.23) we obtain

\[
\tilde{\text{Var}}^\Delta \left( \frac{M_k^\Delta}{\theta_k^\Delta} \right) = \tilde{\text{Var}}^\Delta \left( \sum_{\ell=0}^{k-1} \left( \frac{M_{\ell+1}^\Delta}{\theta_{\ell+1}^\Delta} - \frac{M_\ell^\Delta}{\theta_\ell^\Delta} \right) + M_0 \right) = \sum_{\ell=0}^{k-1} \tilde{E}^\Delta \left( \left( \frac{\phi_\ell^\Delta}{\theta_\ell^\Delta} \right)^2 \right).
\]

(3.24)

If we assume we can ignore all terms in the expansions of \( (\phi_\ell^\Delta / \theta_\ell^\Delta)^2 \) given by

\[
\left( \frac{\phi_\ell^\Delta}{\theta_\ell^\Delta} \right)^2 = \left( \frac{\phi_\ell^\Delta}{\theta_\ell^\Delta} \right)^2 \left( \frac{1}{1 + \psi_\ell} \right)^2 = \left( \frac{\phi_\ell^\Delta}{\theta_\ell^\Delta} \right)^2 \left( 1 - \psi_\ell + (\psi_\ell)^2 - \ldots \right)^2
\]

whose expectation under \( \tilde{P}^\Delta \) is 0 or of order \( \Delta^q \) for \( q \geq \frac{3}{2} \), then the variance of \( M_k^\Delta / \theta_k^\Delta \) under \( \tilde{P}^\Delta \) as given by (3.24) can be approximated by

\[
\tilde{\text{Var}}^\Delta \left( \frac{M_k^\Delta}{\theta_k^\Delta} \right) \approx \sum_{\ell=0}^{k-1} \tilde{E}^\Delta \left( \left( \frac{\phi_\ell^\Delta}{\theta_\ell^\Delta} \right)^2 \right) = \sum_{\ell=0}^{k-1} \tilde{E}^\Delta \left( \left( \frac{1}{\theta_\ell^\Delta} \right)^2 \sum_{j=1}^{m} (\xi_j^\Delta - M_\ell^\Delta d_j^\Delta) \Delta W_j^\Delta \right)^2
\]

(3.25)

We also know from the independence property of the increments \( \Delta W_j^\Delta \) and (3.19) that

\[
\tilde{E}^\Delta \left( \Delta W_{j_1}^\Delta \Delta W_{j_2}^\Delta \right) = E \left( \Delta W_{k}^{j_1} \Delta W_{k}^{j_2} \theta_N^\Delta \right)
\]

\[
= E \left( \Delta W_{k}^{j_1} \Delta W_{k}^{j_2} \prod_{\ell=0}^{N-1} (1 + \psi_\ell^\Delta) \right)
\]

\[
= E \left( \Delta W_{k}^{j_1} \Delta W_{k}^{j_2} (1 + \psi_k^\Delta) \right)
\]

\[
= E \left( \Delta W_{k}^{j_1} \Delta W_{k}^{j_2} \right) + E \left( \Delta W_{k}^{j_1} \Delta W_{k}^{j_2} \Delta W_{k}^{j_2} \right)
\]

\[
+ E \left( \Delta W_{k}^{j_1} \Delta W_{k}^{j_2} \Delta W_{k}^{j_2} \right) = 0
\]

for \( k \in \{0, \ldots, N - 1\} \), \( j_1, j_2 \in \{1, \ldots, m\} \) with \( j_1 \neq j_2 \). Consequently since for any \( \ell \in \{0, \ldots, N - 1\} \) and \( j \in \{1, \ldots, m\} \) the random variable \( \Delta W_{j}^{j} \) will be independent of the variates \( (1/\theta_\ell^\Delta)^2, \xi_j^\Delta, M_\ell^\Delta \) and \( d_j^\Delta \), as these variates include increments \( \Delta W_{j}^{j} \), \( j \in \{1, \ldots, m\} \) with index values \( p \) only upto but excluding the value \( \ell \), then from (3.25) and (3.18) we have

\[
\tilde{\text{Var}} \left( \frac{M_k^\Delta}{\theta_k^\Delta} \right) \approx \sum_{\ell=0}^{k-1} \tilde{E}^\Delta \left( \left( \frac{1}{\theta_\ell^\Delta} \right)^2 \sum_{j=1}^{m} (\xi_j^\Delta - M_\ell^\Delta d_j^\Delta)^2 \right) \Delta.
\]

(3.26)
This should be compared to the variance obtained from (2.5) for the estimator $M^\theta$.

We will now consider the problem of directly estimating the integrands in the martingale representation (1.1) using discrete time methods. Let $X_k^\Delta = (X_k^\Delta_1, \ldots, X_k^\Delta_d)$, $k \in \{0, \ldots, N-1\}$ be an Euler approximation of the process $X$ given by (1.1.1) of the form

$$X_{k+1}^\Delta = X_k^\Delta + a^i(t_k, X_k^\Delta) \Delta + \sum_{j=1}^{m} b^{ij}(t_k, X_k^\Delta) \Delta W_k^j$$

for $i \in \{1, \ldots, d\}$, $k \in \{0, \ldots, N-1\}$ with initial value $X_0^\Delta = X_0 = \bar{x} \in \mathbb{R}^m$, where $\Delta W_k^j$ are two point random variables with probabilities given by (3.5) and $(t)^\Delta$ is the equi-spaced time discretization of the interval $[t_0, T]$ given by (3.1). For simplicity we consider a valuation functional of the form $h(X_N^\Delta)$.

That is, we consider only functionals of the terminal value of the approximation $X^\Delta$ corresponding to European style derivative securities. We assume that $X^\Delta$ converges weakly to $X$ for the functional $h$, see for example Kloeden & Platen (1992). For $k \in \{0, \ldots, N-1\}$ define

$$M_k^\Delta = E \left( h \left( X_N^\Delta \right) \right | \mathcal{F}_k).$$

As in Chapter 1, see (1.1.13), we let $u : \{t_0, \ldots, t_N\} \times \mathbb{R}^d \rightarrow \mathbb{R}$, be a valuation function of the form

$$u(t_k, X_k^\Delta) = M_k^\Delta.$$

This expression and the law of iterated conditional expectations applied to the discrete time martingale $M^\Delta$ means that

$$u(t_k, X_k^\Delta) = E \left( u(t_{k+1}, X_{k+1}^\Delta) \right | \mathcal{F}_k).$$

We will now estimate the function $u$ at time $t_k$, $0 \leq k \leq N$ using a backward numerical technique. The basic idea is as follows: From (3.30) we can evaluate $u(t_k, X_k^\Delta)$ if we know or have previously estimated $u(t_{k+1}, X_{k+1}^\Delta)$. The functional $h$ determines the values of $u$ at time $t_N = T$. Consequently we can evaluate $u$ at the earlier times $t_{N-1}, t_{N-2}, \ldots, t_0$.

In practice we cannot evaluate $u$ at all nodes of the tree or lattice formed from the two point random variables $\Delta W_k^j$, $j \in \{1, \ldots, m\}$, $k \in \{0, \ldots, N-1\}$. For example if $N = 256$ and $m = 2$ then the total number of outcomes or paths for the increments $\Delta W_k^j$ equals $2^{512} > 10^{153}$. However we can estimate $u$ at certain points and use interpolation techniques to compute values between the points. For $n \geq 2$ multidimensional interpolation methods are needed. We may also require these interpolation procedures to produce $C^{1,2}$ smooth functions, if these estimates of $u$ are to be used together with some other variance reduction techniques.

This method for calculating $u$ using neighbouring points obtained from the two point random variables $\Delta W_k^j$ at the next time step is related to a more general Markov chain approximation technique proposed by Platen (1992).
2.4. CONTROL VARIATES AND INTEGRAL REPRESENTATIONS

Suppose \( k \) and \( X^\Delta_k^i \) are fixed with \( k \in \{0, \ldots, N-1\} \). Let \( \Delta W_k = (\Delta W^1_k, \ldots, \Delta W^m_k) \) and \( X^\Delta_{k+1}(\Delta W_k) \) be the value of \( X^\Delta_{k+1} \) corresponding to the outcomes for \( \Delta W_k \) and the \( i \)-th component of \( X^\Delta \). By (3.30) and the property, obtained from (3.5), that the vector of increments \( \Delta W_k \) takes values in the set \( Q = \{-\sqrt{\Delta}, \sqrt{\Delta}\}^{1,\ldots,m} \) we have

\[
u(t_k, X^\Delta_k) = \frac{1}{2^m} \sum_{\Delta W_k \in Q} u\left(t_{k+1}, X^\Delta_{k+1}(\Delta W_k)\right). \tag{3.31}
\]

We can now attempt to compute the integrands in an approximate representation of \( M^\Delta_{k+1} = u(t_{k+1}, X^\Delta_{k+1}) \) of the form

\[
u(t_{k+1}, X^\Delta_{k+1}) \approx \nu(t_k, X^\Delta_k) + \sum_{j=1}^m \zeta^j_k \Delta W^j_k. \tag{3.32}
\]

This representation can be obtained from the Euler approximation (3.2) with \( \nu(t_k, X^\Delta_k) \) replacing \( M^\Delta_{k+1} \). In the case \( m \geq 2 \) we may not be able to find processes \( \zeta^j_k \) which solves (3.32) exactly. This is because the vector of increments \( \Delta W_k \) has \( 2^m \) outcomes, hence equation (3.32), involves solving a system of \( 2^m \) linear equations with \( m \) unknown variables \( \zeta^1_k, \ldots, \zeta^m_k \). In fact, the discrete framework here leads to a form of incompleteness and several possible choices for \( \zeta^j_k \). For a discussion on these issues the interested reader is referred to Hofmann, Platen & Schweizer (1992).

However, if we use the criteria of minimization of the squares

\[
\left( \nu(t_{k+1}, X^\Delta_{k+1}) - \nu(t_k, X^\Delta_k) - \sum_{j=1}^m \zeta^j_k \Delta W^j_k \right)^2,
\]

then we can find the optimal vector of coefficients \( \zeta^1_k, \ldots, \zeta^m_k \) using least-squares analysis. Some additional information on the application of least-squares is given in the next section on control variates. We remark that, as in the case for the valuation function \( u \), we can estimate the integrands \( \zeta^j_k, j \in \{1, \ldots, m\}, k \in \{0, \ldots, N-1\} \) at certain points and use multidimensional interpolation methods to determine values at intermediate points. This method provides an effective mechanism for approximating the integrand \( \xi = (\xi^1, \ldots, \xi^m) \) for a general valuation martingale \( M \) with representation (1.1) and which is of the form \( M_t = E(h(X^T_t) \mid F_t), \quad t_0 \leq t \leq T \), where \( X^{t_0, \bar{x}} \) is a \( d \)-dimensional diffusion process which satisfies (1.1.1).

2.4 Control Variates and Integral Representations

In this section we consider again a general \( d \)-dimensional diffusion process \( X^{t_0, \bar{x}} = \{X^{t_0, \bar{x}}_t = (X^{t_0, \bar{x}}_t, \ldots, X^{d, t_0, \bar{x}}_t), t_0 \leq t \leq T\} \) with initial value \( \bar{x} = (\bar{x}_1, \ldots, \bar{x}_d) \in \mathbb{R}^d \) satisfying (1.1.1). We assume as given by (1.1.13) with \( h = \tilde{h} \) and \( \tau = T \) that for a payoff functional \( h : \mathbb{R}^d \rightarrow \mathbb{R} \) there is a corresponding valuation martingale \( M \) and function \( u : ([t_0, T], \mathbb{R}^d) \rightarrow \mathbb{R} \) with

\[
M_t = u\left(t, X^{t_0, \bar{x}}_t\right) = E\left(h\left(X^{t_0, \bar{x}}_T\right) \mid F_t\right) \tag{4.1}
\]
for $t_0 \leq t \leq T$. Our aim will be to find control variate formulations that will enable us to construct an accurate and fast estimate of $E(h(X_{T}^{0,\xi})) = u(t_0, \xi)$.

The classical control variate method, see for example Ross (1991) and Law & Kelton (1991), is based on finding a random variable $Y$ with known mean $E(Y)$ and estimating $E(h(X_{T}^{0,\xi}))$ by using the variate $Z = h(X_{T}^{0,\xi}) - \alpha(Y - E(Y))$ for some suitable choice of $\alpha \in \mathbb{R}$, rather than $h(X_{T}^{0,\xi})$ directly. The parameter $\alpha$ is chosen to minimize the variance of $Z$. Because $E(Z) = E(h(X_{T}^{0,\xi}))$, $Z$ will be an unbiased estimator for $E(h(X_{T}^{0,\xi}))$. We assume of course that both $h(X_{T}^{0,\xi})$ and $Y$ can be evaluated for any realization $\omega \in \Omega$. With this type of formulation the random variable $Y$ is called a control variate for the estimation of $E(X_{T}^{0,\xi})$.

The basic control variate method is simple but very powerful. Suppose $\hat{X}_{T}^{0,\xi} = \{\hat{X}_{t}^{0,\xi} = (\hat{X}_{t}^{1,0,\xi}, \ldots, \hat{X}_{t}^{d,0,\xi}), \ t_0 \leq t \leq T\}$ is some $d$-dimensional diffusion process which also satisfies an equation of the form (1.1.1) and which approximates $X_{T}^{0,\xi}$, and the valuation function $\tilde{u} : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}$ for $h(\hat{X}_{T}^{0,\xi})$ is known and satisfies

$$\tilde{u}(t, \hat{X}_{t}^{0,\xi}) = E\left(h\left(\hat{X}_{T}^{0,\xi}\right) \mid \mathcal{F}_t\right)$$

for $t_0 \leq t \leq T$, where $h$ is the same payoff functional as used in (4.1). Then the random variable

$$\hat{Z}_T = h\left(X_{T}^{0,\xi}\right) - \alpha\left(h\left(\hat{X}_{T}^{0,\xi}\right) - E\left(h\left(\hat{X}_{T}^{0,\xi}\right)\right)\right)$$

$$= u\left(T, X_{T}^{0,\xi}\right) - \alpha\left(\tilde{u}\left(T, \hat{X}_{T}^{0,\xi}\right) - \tilde{u}(t_0, \xi)\right)$$

(4.3)

will be an unbiased estimator for $E(h(X_{T}^{0,\xi}))$ and will usually be a variance reduced estimator if $\hat{X}_{T}^{0,\xi}$ is close to $X_{T}^{0,\xi}$. Since

$$\text{Var}(\hat{Z}_T) = \text{Var}\left(h\left(X_{T}^{0,\xi}\right)\right) + \alpha^2 \text{Var}\left(h\left(\hat{X}_{T}^{0,\xi}\right)\right)$$

$$- 2\alpha \text{Cov}\left(h\left(X_{T}^{0,\xi}\right), h\left(\hat{X}_{T}^{0,\xi}\right)\right)$$

the value of $\alpha$ which will minimize the variance of $\hat{Z}_T$ is

$$\alpha = \frac{\text{Cov}\left(h\left(X_{T}^{0,\xi}\right), h\left(\hat{X}_{T}^{0,\xi}\right)\right)}{\text{Var}\left(h\left(\hat{X}_{T}^{0,\xi}\right)\right)}.$$ 

(4.4)

Note that if a Monte Carlo simulation using the variate $\hat{Z}_T$ is being performed we can estimate simultaneously the best value of $\alpha$ which will minimize the variance of $\hat{Z}_T$ as given by (4.4). This type of calculation can be performed as the simulation proceeds, an observation which is explained by Clewlow & Carverhill (1992, 1994). We do not need to store the values of the variates $h(X_{T}^{0,\xi})$ and $h(\hat{X}_{T}^{0,\xi})$ for each path or realization of the simulation. This is a simple but useful form of least-squares analysis which can be extended in a straightforward manner to include linear combinations of control variates of the form $\alpha_1 Y_1 + \ldots + \alpha_n Y_n$. 

As an example of this method we will consider a stochastic volatility model with a vector process \( X = (S, \sigma) = \{(S_t, \sigma_t); \ t_0 \leq t \leq T\} \) that satisfies the two-dimensional stochastic differential equation

\[
\begin{align*}
    dS_t &= \sigma_t S_t dW_t^1 \\
    d\sigma_t &= (\kappa - \sigma_t) \, dt + \xi \sigma_t dW_t^2
\end{align*}
\]

for \( t_0 \leq t \leq T \) with initial values \( S_{t_0} = s, \ \sigma_{t_0} = \sigma \). We take \((W^1, W^2)\) to be a two-dimensional Wiener process defined on the probability space \((\Omega, \mathcal{F}, P)\), where \( h(S_T, \sigma_T) = (S_T - K)^+ \) for some constant \( K > 0 \). These type of models have been considered for instance by Hofmann, Platen & Schweizer (1992) and Heston (1993).

Let \( \tilde{X} = (\hat{S}, \hat{\sigma}) = \{(\hat{S}_t, \hat{\sigma}_t), \ t_0 \leq t \leq T\} \) be an adjusted process which evolves according to the equation

\[
\begin{align*}
    d\hat{S}_t &= \hat{\sigma}_t \hat{S}_t dW_t^1 \\
    d\hat{\sigma}_t &= (\kappa - \hat{\sigma}_t) \, dt
\end{align*}
\]

for \( t_0 \leq t \leq T \) with the same initial conditions. That is \( \hat{S}_{t_0} = s \) and \( \hat{\sigma}_{t_0} = \sigma \). The valuation function \( \check{u} \) corresponding to the payoff structure \( h(\hat{S}_T, \hat{\sigma}_T) = (\hat{S}_T - K)^+ \) and given by

\[
\check{u}(t, \hat{S}_t, \hat{\sigma}_t) = E \left( (\hat{S}_t - K)^+ \, \bigg| \mathcal{F}_t \right)
\]

for \( t_0 \leq t \leq T \) can be computed explicitly for the process \( \tilde{X} = (\hat{S}, \hat{\sigma}) \). This result shows that the variate

\[
\hat{Z}_T = (S_T - K)^+ - \alpha((\hat{S}_T - K)^+ - E(\hat{S}_T - K)^+))
\]

will have outcomes that are easily computed, and, for this particular valuation problem, constitutes a powerful unbiased variance reduced estimator for \( E((S_T - K)^+) \).

This control variate technique can be extended to include approximations \( \bar{h} : \mathbb{R}^d \rightarrow \mathbb{R} \) of the payoff function \( h \). Thus, if the valuation function \( \check{u} : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R} \) given by

\[
\check{u}(t, \hat{X}_t^{t_0, z}) = E \left( \bar{h}(\hat{X}_T^{t_0, z}) \, \bigg| \mathcal{F}_t \right)
\]

(4.5)

for \( t_0 \leq t \leq T \) is known, then a control variate can be constructed from \( \bar{h}(\hat{X}_T^{t_0, z}) \) in a similar manner to that outlined for \( h(\hat{X}_T^{t_0, z}) \). That is, we use the unbiased estimator \( h(\hat{X}_T^{t_0, z}) - \alpha(\bar{h}(\hat{X}_T^{t_0, z}) - E(\bar{h}(\hat{X}_T^{t_0, z})) \) as is suggested by (4.3). We assume the expectation \( E(\bar{h}(\hat{X}_T^{t_0, z})) \) is known explicitly or can be accurately approximated. Consequently this control variate method extends to include changes both in the underlying diffusion \( X^{t_0, z} \) and the payoff function \( h \).
Using the martingale property of the process $\tilde{M} = \{\tilde{u}(t, \tilde{X}^{t_0}_t), t_0 \leq t \leq T\}$ as given by (4.5) we can write, subject to certain integrability conditions applying for $\tilde{M}$,

$$
\tilde{M}_t = \tilde{u} \left( t, \tilde{X}^{t_0}_t \right) = \hat{u}(t_0, \omega) + \sum_{j=1}^{m} \int_{t_0}^{t} \xi_j^i \ dW^j_s
$$

(4.6)

where $\xi = (\xi^1, \ldots, \xi^m)$ is a vector of $\mathcal{F}$-predictable processes satisfying (1.2).

The following formulation is useful for certain types of path dependent options, for example, American options or barrier options with stochastic volatility. Let $\tau : \Omega \rightarrow \mathbb{R}$ be a stopping time as given by (1.1.4) with continuation region $\Gamma_0$ and exercise boundary $\Gamma_1$. We take $h : \Gamma_1 \rightarrow \mathbb{R}$ to be some payoff function and we would like to approximate the expectation $E(h(X^{t_0}_T))$ assuming $\hat{u}$ is known. Using the Ito formula for semimartingales we have from (4.6) the relation

$$
\tilde{M}_\tau = \hat{u} \left( \tau, \tilde{X}^{t_0}_\tau \right) = \hat{u}(t_0, \omega) + \sum_{j=1}^{m} \int_{t_0}^{\tau} \xi_j^i \ dW^j_s.
$$

(4.7)

This means that $E(\hat{u}(\tau, \tilde{X}^{t_0}_\tau)) = \hat{u}(t_0, \omega)$ and therefore the random variable

$$
\tilde{Z}_\tau = h \left( X^{t_0}_\tau \right) - \alpha \left( \hat{u} \left( \tau, \tilde{X}^{t_0}_\tau \right) - \hat{u}(t_0, \omega) \right)
$$

(4.8)

is an unbiased estimator for $E(h(X^{t_0}_T))$. Note that $\hat{u}$ can be any valuation function of the form (4.6). For example it may correspond to a European style security with $M_t = E(h(X^{t_0}_t) | \mathcal{F}_t), \ t_0 \leq t \leq T$, for some payoff function $h'$ and have no direct relationship with the stopping time $\tau$, which is obtained for a different problem.

These control variate formulations can in addition be conveniently applied in a discrete time setting. For example, if we intend to use the control variate given by (4.3), we may replace $X$ and $\tilde{X}$ by discrete time approximations $X^\Delta$ and $\tilde{X}^\Delta$, respectively, say of the form (3.27). In fact there is often more flexibility with discrete time formulations. If for instance there is no natural choice for the process $\tilde{X}$ in (4.3) a control variate can be obtained from $X^\Delta$ itself since $E(X^\Delta)$ can usually be easily calculated for discrete numerical schemes.

For one-dimensional diffusions Clewlow & Carverhill (1992, 1994) have introduced and used linear combinations of discrete time martingale control variates. If we let $(t)_\Delta$ be a time discretization of the interval $[t_0, T]$ as given by (3.1), then these control variates in several dimensions take the form

$$
Y^\Delta_N = \sum_{k=0}^{N-1} \phi_k^i \left( \Delta X^\Delta_k - E \left( \Delta X^\Delta_k \right) \right)
$$

for $i \in \{1, \ldots, d\}$, $k \in \{0, \ldots, N - 1\}$, where $\Delta X^\Delta_k = X^\Delta_{k+1} - X^\Delta_k$ and $\phi_k^i$ is $\mathcal{F}_t_k$-measurable and is chosen as an approximation to a hedging parameter such as the delta or gamma for the $i$th component of the diffusion process $X$ given by (1.1.1). Since $E(Y^\Delta_k | \mathcal{F}_t_k) = Y_k$ for $\ell, k \in \{0, \ldots, N - 1\}$ with $\ell \leq k$ then $Y^\Delta$ is a discrete time martingale. The expectation $E(\Delta X^\Delta_{\Delta_k})$, $i \in \{1, \ldots, d\}$, $k \in \{0, \ldots, N - 1\}$ can usually
be calculated in a straightforward manner. For example with the Euler scheme (3.27)
\[ E(\Delta X_k^{\Delta_i}) = a^i(t_k, X_k^{\Delta}) \Delta. \]

If we consider the stochastic differential equation (1.1.1) which defines our underlying diffusion process a continuous time version of the above martingale control variate would be

\[ Y_T^i = \int_{t_0}^{T} \psi_i^j \left( dX_s^{t_0,x} - a^i(s, X_s^{t_0,x}) \right) ds \]

\[ = \int_{t_0}^{T} \psi_i^j \sum_{j=1}^{m} b^{ij} \left( s, X_s^{t_0,x} \right) dW_s^j \]

for \( i \in \{1, \ldots, d\} \), where \( \psi = (\psi^1, \ldots, \psi^d) \) is some vector of \( \mathcal{F} \)-predictable processes which satisfies appropriate integrability conditions. This representation shows that the control variate \( Y_T^i \) is expressible as an Itô integral and therefore its expectation is zero. It also follows that another natural choice for a control variate is simply an Itô integral of the form

\[ \bar{Y}_T = \sum_{j=1}^{m} \int_{t_0}^{T} \xi_j^j dW_s^j, \]

where \( \xi = (\xi^1, \ldots, \xi^d) \) is a vector of \( \mathcal{F} \)-predictable processes.

From the representation (1.1) we can now compute the variance under \( P \) of the unbiased estimator \( \bar{Z}_T \) of \( E(h(X_t^{t_0,x})) \) given by

\[ \bar{Z}_T = h\left(X_T^{t_0,x}\right) - \alpha \bar{Y}_T, \]

as

\[ \text{Var}(\bar{Z}_T) = \text{Var} \left( \sum_{j=1}^{m} \int_{t_0}^{T} \left( \xi_j^j - \alpha \xi_j^j \right) dW_s^j \right) \]

\[ = E \left[ \left( \sum_{j=1}^{m} \int_{t_0}^{T} \left( \xi_j^j - \alpha \xi_j^j \right) dW_s^j \right)^2 \right] \]

\[ = \sum_{j=1}^{m} \int_{t_0}^{T} E \left( \xi_j^j - \alpha \xi_j^j \right)^2 ds. \]

This shows that if we can find good approximations \( \bar{\xi}_j, j \in \{1, \ldots, m\} \) for the integrands \( \xi_j \), then with \( \alpha = 1 \) the variance of the unbiased estimator \( \bar{Z}_T \) will be small.

Note that to produce an arbitrarily small variance we require only an approximation to the integrand \( \xi = (\xi^1, \ldots, \xi^m) \) which is related to the deltas of the underlying security. The use of control variates based on gammas or other greeks is therefore, in a theoretical sense, not required.

Suppose the valuation function \( u \), and its approximation \( \hat{u} \), as given by (4.1) and (4.2) respectively, are of class \( C^{1,2} \). Then expanding both \( u(T, X_T^{t_0,x}) \) and \( \hat{u}(T, X_T^{t_0,x}) \)
using Ito’s rule and applying the Kolmogorov backward equation we obtain

\[
u(T, X_{T}^{t_0, x}) = u(t_0, x) + \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{T} \frac{\partial u}{\partial x_i} (s, X_{s}^{t_0, x}) b^{i,j} (s, X_{s}^{t_0, x}) \, dW^j_s
\]

and

\[
\hat{u}(T, \hat{X}_{T}^{t_0, x}) = \hat{u}(t_0, x) + \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{T} \frac{\partial \hat{u}}{\partial x_i} (s, \hat{X}_{s}^{t_0, x}) b^{i,j} (s, \hat{X}_{s}^{t_0, x}) \, dW^j_s.
\]  

(4.13)

The variance of the estimator \(\hat{Z}_T\) given by (4.3) is therefore

\[
\text{Var}(\hat{Z}_T) = \text{Var} \left( \sum_{i=1}^{d} \sum_{j=1}^{m} \int_{t_0}^{T} \left( \frac{\partial u}{\partial x_i} (s, X_{s}^{t_0, x}) b^{i,j} (s, X_{s}^{t_0, x}) - \alpha \frac{\partial \hat{u}}{\partial x_i} (s, X_{s}^{t_0, x}) b^{i,j} (s, X_{s}^{t_0, x}) \right) \, dW^j_s \right)
\]

\[
= \sum_{j=1}^{m} \int_{t_0}^{T} E \left( \sum_{i=1}^{d} \left( \frac{\partial u}{\partial x_i} (s, X_{s}^{t_0, x}) b^{i,j} (s, X_{s}^{t_0, x}) - \alpha \frac{\partial \hat{u}}{\partial x_i} (s, X_{s}^{t_0, x}) b^{i,j} (s, X_{s}^{t_0, x}) \right)^2 \, ds \right). 
\]  

(4.14)

If we set

\[
\tilde{\xi}_s^j = \sum_{i=1}^{d} \frac{\partial \hat{u}}{\partial x_i} (s, X_{s}^{t_0, x}) b^{i,j} (s, X_{s}^{t_0, x}) 
\]

(4.15)
as used in (4.10) for \(j \in \{1, \ldots, m\}\), then the variance of the estimator \(\hat{Z}_T\) given by (4.11) can be calculated from (4.12). Thus

\[
\text{Var}(\hat{Z}_T) = \sum_{j=1}^{m} \int_{t_0}^{T} E \left( \sum_{i=1}^{d} \frac{\partial(u - \alpha \hat{u})}{\partial x_i} (s, X_{s}^{t_0, x}) b^{i,j} (s, X_{s}^{t_0, x}) \right)^2 \, ds .
\]  

(4.16)

We observe that the variance obtained for \(\hat{Z}_T\) is similar to, but different from, that obtained for \(Z_T\) with this choice for \(\tilde{\xi} = (\tilde{\xi}_1^1, \ldots, \tilde{\xi}_m^m)\). In general we can expect a lower variance for \(\hat{Z}_T\) as can be seen by comparing (4.14) with (4.16). The computational loads corresponding to the two estimators can also vary. For example Monte Carlo estimations using the variate \(\hat{Z}_T\) involve only the calculation of payoff structures but require the simulation of two separate diffusion processes \(X_{t_0}^{t_0, x}\) and \(\hat{X}_{t_0}^{t_0, x}\). This can be compared with estimates of \(Z_T\) which require simulation of only one explicit diffusion process, \(X_{t_0}^{t_0, x}\). However evaluation of the Ito integrals (4.10) effectively involves the simulation of another diffusion process, namely the Ito integral itself. Choosing the best estimator for a given valuation problem often involves preliminary simulation studies and clearly depends on the specific structure of the underlying model and payoff functional.
2.4. CONTROL VARIATES AND INTEGRAL REPRESENTATIONS

We can extend the above analysis, for the estimator $\tilde{Z}_T$, to include linear combinations of control variates $\sum_{\ell=1}^{L} \alpha_{\ell} \tilde{Y}_{T}^{\ell}$, where each $\tilde{Y}_{T}^{\ell}$, $\ell \in \{1, \ldots, L\}$, is an Ito integral of the form (4.10). Using these Ito integrals the random variable

$$\tilde{Z}_T = h \left( X_T^{t_0, \xi} \right) - \sum_{\ell=1}^{L} \alpha_{\ell} \tilde{Y}_{T}^{\ell}$$  \hspace{1cm} (4.17)

will be an unbiased estimator for $E(h(X_T^{t_0, \xi}))$ with variance

$$\text{Var}(\tilde{Z}_T) = \text{Var} \left( h \left( X_T^{t_0, \xi} \right) \right) + \sum_{\ell=1}^{L} \sum_{\ell', k=1}^{L} \alpha_{\ell} \alpha_{k} \text{Cov}(\tilde{Y}_{T}^{\ell}, \tilde{Y}_{T}^{k})$$

$$- 2 \sum_{\ell=1}^{L} \alpha_{\ell} \text{Cov} \left( h \left( X_T^{t_0, \xi} \right), \tilde{Y}_{T}^{\ell} \right).$$  \hspace{1cm} (4.18)

This variance will be minimized if

$$\sum_{k=1}^{L} \alpha_{k} \text{Cov}(\tilde{Y}_{T}^{k}, \tilde{Y}_{T}^{k}) = \text{Cov} \left( h \left( X_T^{t_0, \xi} \right), \tilde{Y}_{T}^{k} \right)$$  \hspace{1cm} (4.19)

for each $\ell \in \{1, \ldots, L\}$. This system of linear equations will admit a unique solution if the set of control variates $\{\tilde{Y}_{T}^{\ell}, \ell \in \{1, \ldots, L\}\}$ is linearly independent. As is the case for a single control variate we can estimate the quantities Cov($\tilde{Y}_{T}^{\ell}, \tilde{Y}_{T}^{k}$) and Cov($h(X_T^{t_0, \xi})\tilde{Y}_{T}^{\ell}$), $\ell, k \in \{1, \ldots, L\}$, for a given simulation and simultaneously the optimal vector of coefficients $\alpha = (\alpha_1, \ldots, \alpha_L)$ to minimize the variance as given by (4.19).

Also, as has been noted for a single control variate we can calculate the optimal vector of coefficients $\alpha$ progressively during the simulation. We do not need to store output data for the variates $X_T^{t_0, \xi}, \tilde{Y}_{T}^{1}, \ldots, \tilde{Y}_{T}^{L}$ for each path of the simulation.

This formulation is simplified if we assume that Cov($\tilde{Y}_{T}^{\ell}, \tilde{Y}_{T}^{k}$) = 0 for $\ell \neq k, \ell, k \in \{1, \ldots, L\}$. That is, the random variables $\tilde{Y}_{T}^{\ell}$, $\ell \in \{1, \ldots, L\}$, are mutually orthogonal when considered as elements of the Hilbert space $L^2(\Omega, \mathcal{F}_T, P)$ with inner product $(X, Y) = E(XY)$ for $X, Y \in L^2(\Omega, \mathcal{F}_T, P)$. In this case (4.19) reduces to

$$\alpha_{\ell} = \frac{\text{Cov} \left( h \left( X_T^{t_0, \xi} \right), \tilde{Y}_{T}^{\ell} \right)}{\text{Var}(\tilde{Y}_{T}^{\ell})}$$

$$= \frac{E \left( \left( h \left( X_T^{t_0, \xi} \right) - E \left( h \left( X_T^{t_0, \xi} \right) \right) \right) \tilde{Y}_{T}^{\ell} \right)}{||\tilde{Y}_{T}^{\ell}||_2^2}$$  \hspace{1cm} (4.20)

for $\ell \in \{1, \ldots, L\}$, where $|| \cdot ||_2$ denotes the norm in $L^2(\Omega, \mathcal{F}_T, P)$.

A mutually orthogonal set of control variates $\tilde{Y}_{T}^{\ell}$, $\ell \in \{1, \ldots, L\}$, can be constructed using the Gram-Schmidt orthogonalization process. For example if $\tilde{Y}_{T}^{1}$ and $\tilde{Y}_{T}^{2}$ are not orthogonal we can replace $\tilde{Y}_{T}^{2}$ with

$$\tilde{Y}_{T}^{2} - \frac{\text{Cov}(\tilde{Y}_{T}^{1}, \tilde{Y}_{T}^{2})}{\text{Var}(\tilde{Y}_{T}^{1})} \tilde{Y}_{T}^{1}.$$
CHAPTER 2. VARIANCE REDUCTION TECHNIQUES

There are many ways of constructing the control variate process \( \tilde{Y} = (\tilde{Y}^1, \ldots, \tilde{Y}^L) \). If we have an equi-spaced time discretization \((t)_{\Delta}\) of the interval \([t_0, T] \) of the form (3.1) a good choice for \( \tilde{Y} \) using control variates of the form (4.10), for certain types of valuation problems, is given by

\[
\tilde{Y}^\ell_T = \sum_{j=1}^{m} \int_{t_0}^{T} 1_{\{s \in [t_{\ell-1}, t_{\ell})\}} \xi^\ell_s dW^s_s
\]

for \( \ell \in \{1, \ldots, L\} \). For \( \ell_1 \neq \ell_2 \)

\[
(Y^\ell_1, Y^\ell_2) = \sum_{j=1}^{m} \int_{t_0}^{T} E \left( 1_{\{s \in [t_{\ell_1-1}, t_{\ell_1})\}} 1_{\{s \in [t_{\ell_2-1}, t_{\ell_2})\}} \left( \xi^\ell_s \right)^2 \right) ds
\]

so any two control variates \( \tilde{Y}^{\ell_1}_T \) and \( \tilde{Y}^{\ell_2}_T, \ell_1 \neq \ell_2 \) are orthogonal.

These control variate methods can be applied iteratively or in combination with other techniques. For example we may find an initial approximation \( u_1 : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R} \) for \( u \) from finite-difference methods. The integrands given by (4.15) together with appropriate interpolation routines could then be used to construct an Ito integral control variate of the form (4.10) and from this a new approximation \( u_2 \) for \( u \) would be obtained. Clearly this procedure can be repeated until sufficient accuracy is achieved.

2.5 Other Variance Reduction Methods

In this section we will consider some supplementary variance reduction techniques. We will describe some extensions of existing methods and new formulations of certain classical techniques. These include applications of the classical conditional variance formula, stratified sampling, use of antithetic variates and quasi Monte Carlo. As in the previous section we assume a payoff structure of the form \( h(X^0_T, \Xi) \), where \( X^0_T, \Xi \) is a \( d \)-dimensional diffusion process which satisfies (1.1.1).

The classical conditional variance formula, see for example Ross (1991), takes the form

\[
\text{Var}(Y) = E(\text{Var}(Y \mid Z)) + \text{Var}(E(Y \mid Z))
\]

for any two random variables defined on some sample space \( \Omega' \) with probability measure \( P' \).

To extend this result to stochastic processes we first use the concept of conditional variance with respect to sub-\( \sigma \)-algebras of \( \mathcal{F}_T \), see for example Shiryaev (1984). Let \( X \) be an integrable random variable and \( \mathcal{G} \subseteq \mathcal{F}_T \) a sub-\( \sigma \)-algebra. The conditional variance of \( X \) with respect to \( \mathcal{G} \), denoted \( \text{Var}(X \mid \mathcal{G}) \) is defined by

\[
\text{Var}(X \mid \mathcal{G}) = E \left( (X - E(X \mid \mathcal{G}))^2 \mid \mathcal{G} \right)
\]

\[
= E \left( X^2 \mid \mathcal{G} \right) - (E(X \mid \mathcal{G}))^2.
\]
2.5. OTHER VARIANCE REDUCTION METHODS

Let $\mathcal{G}_1, \mathcal{G}_2$ be sub-$\sigma$-algebras of $\mathcal{F}_T$ with $\mathcal{G}_1 \subseteq \mathcal{G}_2$. Using the law of iterated conditional expectations and the definition of conditional variance we have for any integrable random variable $X$ the relations

\[
\text{Var}(X | \mathcal{G}_1) = E \left( X^2 | \mathcal{G}_1 \right) - \left( E(X | \mathcal{G}_1) \right)^2.
\]

\[
= E \left( E \left( X^2 | \mathcal{G}_2 \right) | \mathcal{G}_1 \right) - \left( E \left( E(X | \mathcal{G}_2) | \mathcal{G}_1 \right) \right)^2
\]

\[
= E \left( E \left( X^2 | \mathcal{G}_2 \right) | \mathcal{G}_1 \right) - E \left( \left( E(X | \mathcal{G}_2) \right)^2 | \mathcal{G}_1 \right) + E \left( \left( E(X | \mathcal{G}_2) \right)^2 | \mathcal{G}_1 \right) - \left( E \left( E(X | \mathcal{G}_2) | \mathcal{G}_1 \right) \right)^2
\]

\[
= E \left( \text{Var} \left( X | \mathcal{G}_2 \right) | \mathcal{G}_1 \right) + \text{Var} \left( E(X | \mathcal{G}_2) | \mathcal{G}_1 \right).
\]

(5.3)

In particular if we let $\mathcal{G}_1 = \{\phi, \Omega\}$ we obtain

\[
\text{Var}(X) = E \left( \text{Var} \left( X | \mathcal{G}_2 \right) \right) + \text{Var} \left( E(X | \mathcal{G}_2) \right) \tag{5.4}
\]

for $\mathcal{G}_2$ a sub-$\sigma$-algebra of $\mathcal{F}_T$. This is, for stochastic processes, the analogue of the classical conditional variance formula (5.1).

Let $t_0 < t \leq T$ and consider the random variable $E(h(X_{t_0}^t) | \mathcal{F}_t)$. Since

\[
E \left( E \left( h \left( X_{t_0}^t \right) | \mathcal{F}_t \right) \right) = E \left( h \left( X_{t_0}^t \right) \right)
\]

this variate is an unbiased estimator for $E(h(X_{t_0}^t))$. Also, applying the conditional variance formula (5.4) we have the inequality

\[
\text{Var} \left( E \left( h \left( X_{t_0}^t \right) | \mathcal{F}_t \right) \right) \leq \text{Var} \left( h \left( X_{t_0}^t \right) \right) \tag{5.5}
\]

for $t_0 \leq t \leq T$. In general the inequality in (5.5) will be strict. In this case $E(h(X_{t_0}^t) | \mathcal{F}_t)$ will be a variance reduced unbiased estimator for $E(h(X_{t_0}^t))$. This type of error minimization is known in its classical formulation as variance reduction by conditioning, see Law & Kelton (1991).

We will now turn our attention to the variance reduction technique of stratified sampling. This is an error minimization technique of long standing which has been widely used in general simulation, see for example Ross (1991). Surprisingly the method seems to have been underutilized for financial modelling problems, however recently Curran (1994) has described a simple form of stratified sampling for Asian options and geometric Brownian motion. We will describe another version of stratified sampling which can be applied to a wide class of diffusion models and valuation functionals.

Let $A_i \subseteq \mathcal{F}_T, i \in \{1, \ldots, N\}$ be a set of events satisfying

\[
\bigcup_{i=1}^N A_i = \Omega, \quad A_i \cap A_j = \emptyset
\]
for \( i, j \in \{1, \ldots, N\} \). We assume \( P(A_i) = \frac{1}{N} \) for all \( i, i \in \{1, \ldots, N\} \). Let \( \mathcal{A} = \sigma \{A_i, i \in \{1, \ldots, N\}\} \) be the \( \sigma \)-algebra consisting of finite unions of these sets.

Let \( Z : \Omega \to \mathbb{R} \) be a random variable. We denote by \( Z_{A_i} : A_i \to \mathbb{R} \) the restriction of \( Z \) to \( A_i \) given by \( Z_{A_i}(\omega) = Z(\omega) \) for \( \omega \in A_i \). \( Z_{A_i} \) is a random variable defined on the probability space \((A_i, \mathcal{F}_i, P_i)\), where \( \mathcal{F}_i = \{A_i \cap F : F \in \mathcal{F}_T\} \) and \( P_i : \mathcal{F}_i \to [0,1] \) is given by \( P_i(A_i \cap F) = P(A_i \cap F)/P(A_i) \) for \( F \in \mathcal{F}_T \).

Let \( \bar{Z} = \sum_{i=1}^{N} Z_{A_i}/N \), where \( Z_{A_i}, i \in \{1, \ldots, N\} \), are assumed to be independent. Since

\[
E(\bar{Z}) = \sum_{i=1}^{N} \frac{1}{N} E(Z_{A_i}) = \sum_{i=1}^{N} \int_{A_i} Z \, dP = \int_{\Omega} Z \, dP = E(Z),
\]

(5.6)

\( \bar{Z} \) will be an unbiased estimator for \( E(Z) \). Furthermore, using the independence property of \( Z_{A_i}, i \in \{1, \ldots, N\} \), we have from (5.4)

\[
\text{Var}(\bar{Z}) = \sum_{i=1}^{N} \frac{\text{Var}(Z_{A_i})}{N^2} = \frac{1}{N} \text{Var}(Z),
\]

(5.7)

This inequality will be strict if \( \text{Var}(E(Z/A)) > 0 \). Consequently if we set \( Z = h(X^{t_0,\bar{x}}) \) we obtain from (5.6) an unbiased estimator for \( E(h(X^{t_0,\bar{x}})) \) which from (5.7) will usually be a variance reduced estimator.

As an example let \( X^A \) be an Euler approximation of the one-dimensional diffusion \( X^{t_0,\bar{x}} \) of the form (3.27), where \( \Delta W_k, k \in \{0, \ldots, N-1\} \), are two point random variables given by (3.5). Since we are using two point variates with \( N \) time steps, for practical purposes, we can replace the underlying sample space \( n \) with \( \Omega_N \).

\[
\Omega_N = \{-1, 1\}^{\{0, \ldots, N-1\}}.
\]

(5.8)

For each \( \omega \in \Omega_N \) we denote by \( \omega_k, k \in \{0, \ldots, N-1\} \), the value of \( \omega \) at the \( k \)th index point. For \( \omega \in \Omega_N \) the corresponding 'path' for \( \Delta W_k, k \in \{0, \ldots, N-1\} \) is given by \( \Delta W_k(\omega) = \omega_k \sqrt{\Delta} \) with probabilities \( P_N(\omega) = \frac{1}{2^N} \). Although the number of states \( \omega \in \Omega_N \) is finite, nevertheless, with current technology and values of \( N \) say greater than 30, usually only a tiny fraction of these paths can be sampled.

Let \( \alpha \in \Omega_{N'} \) with \( N' \in \{1, \ldots, N\} \), \( N' < N \) and define \( A_{\alpha} = \{\omega \in \Omega_N : \omega_k = a_k \text{ for } 0 \leq k < N'\} \). The sets \( A_{\alpha}, \alpha \in \Omega_{N'} \) form a collection of \( 2^{N'} \) sets with probabilities \( P(A_\alpha) = \frac{1}{2^{N'}} \). It is apparent that \( \bigcup_{\alpha \in \Omega_{N'}} = \Omega_N \) and that for \( a_1, a_2 \in \Omega_{N'} \), with \( a_1 \neq a_2 \), then \( A_{a_1} \cap A_{a_2} = \emptyset \). Consequently the partitioning rules required for the relations (5.6) and (5.7) are valid with \( Z = h(X^A) \).

With this method our discrete sample space \( \Omega_N \) forms a binary lattice or tree which is divided into sub-lattices \( A_{\alpha}, \alpha \in \Omega_{N'} \) starting at time \( t_{N'} < t_N = T \). A
2.5. OTHER VARIANCE REDUCTION METHODS

stratified Monte Carlo estimation of $E(h(X_T^N))$ using the sets $A_a$, $a \in \Omega_{N'}$, would mean exhausting all paths $\omega \in \Omega_N$ up to time $t_{N'}$ and then sampling, randomly and independently, within the sub-lattices $A_a$, $a \in \Omega_{N'}$. This method not only reduces the variance it also reduces the computational load because the raw or naïve Monte Carlo estimation procedure would involve, usually, many duplicate traversals of the early nodes of the lattice.

We will now describe a variance reduction procedure based on the use of antithetic variates. This technique has been applied for example by Duffie & Glynn (1992), Hull & White (1987, 1988), Clewlow & Carverhill (1992, 1994) and Barraquand (1993). A description of its use in general simulation is given by Ross (1991) and Law & Kelton (1991). Here we describe a version of antithetic variance reduction which together with other variance reduction procedures has been used in the applications considered in Part II of this thesis.

The construction of antithetic variates will be illustrated using the probability space $(\Omega, \mathcal{F}, P)$ where $\Omega = C([t_0, T], \mathbb{R}^2)$. For $\omega \in \Omega$ we denote by $\omega_1(t)$ and $\omega_2(t)$ the components of $\omega(t) \in \mathbb{R}^2$ so that $\omega(t) = (\omega_1(t), \omega_2(t)) \in \mathbb{R}^2$ for $t \in [t_0, T]$. We assume $P$ is the two-dimensional Wiener measure under which the coordinate mappings $W_1^1(\omega) = \omega_1(t)$ and $W_1^2(\omega) = \omega_2(t)$ define a two-dimensional Wiener process on $(\Omega, \mathcal{F}, P)$.

For $\omega \in \Omega$, define $\bar{\omega} \in \Omega$ by $\bar{\omega}(t) = (-\omega_1(t), -\omega_2(t))$, $t_0 \leq t \leq T$ and the random variables $\bar{h}(X_{T}\mid \bar{\omega})$ by

$$
\bar{h} \left( X_{T}^{t_0} \mid \bar{\omega} \right) = h \left( X_{T}^{t_0} \right) \left( \bar{\omega} \right)
$$

for $\omega \in \Omega$.

The variate

$$
\bar{h} \left( X_{T}^{t_0} \right) = \frac{1}{2} \left( h \left( X_{T}^{t_0} \right) + h \left( X_{T}^{t_0} \right) \right)
$$

is an unbiassed estimator for $E(h(X_{T}^{t_0}))$ since $E(\bar{h}(X_{T}^{t_0})) = E(h(X_{T}^{t_0}))$. In addition

$$
\text{Var} \left( \bar{h} \left( X_{T}^{t_0} \right) \right) = \frac{1}{4} \left( \text{Var} \left( h \left( X_{T}^{t_0} \right) \right) + \text{Var} \left( h \left( X_{T}^{t_0} \right) \right) \right)
$$

$$
+ 2 \text{Cov} \left( h \left( X_{T}^{t_0} \right), h \left( X_{T}^{t_0} \right) \right)
$$

and, because the random variables $h(X_{T}^{t_0})$ and $\bar{h}(X_{T}^{t_0})$ will often be negatively correlated, $\bar{h}(X_{T}^{t_0})$ will then be a variance reduced estimator for $E(h(X_{T}^{t_0}))$. We remark that other combinations are possible for $\bar{\omega}$, for example, reflection for only one component or partial reflections over time. These alternatives can then be assembled in various ways to produce new variance reduced unbiassed estimators for $E(h(X_{T}^{t_0}))$. The use of antithetic variates is a general variance reduction method which can be combined with other techniques considered in this chapter.

We will now consider briefly the application of quasi Monte Carlo methods. There is an extensive literature on quasi Monte Carlo techniques with overviews provided by Ripley (1983) and Niederreiter (1992). Applications to financial modelling problems
have been considered by Barraquand (1993), Paskov & Traub (1994) and Joy, Boyle & Tan (1995).

To illustrate the most commonly used procedure, consider a $d$-dimensional diffusion process $X^{t_0,x}$ which satisfies (1.1.1) with payoff functional $h(X^{t_0,x})$. We assume the joint density function $p_X : \mathbb{R}^d \to \mathbb{R}$ of $X^{t_0,x}$ is known so that

$$u(t_0,x) = E \left( h \left( X^{t_0,x}_T \right) \right) = \int_{\mathbb{R}^d} h(x) p_X(x) \, dx. \quad (5.10)$$

Consequently, estimation of $u(t_0,x)$ can be considered as a numerical integration problem over $\mathbb{R}^d$. For a one-dimensional diffusion process $X^{t_0,x}$ if we let $F_X : \mathbb{R} \to [0,1]$ be the distribution function for $X^{t_0,x}_T$ with $F'_X(x) = p_X(x)$, $x \in \mathbb{R}$, then (5.10) can be expressed in the form

$$u(t_0,x) = \int_0^1 h(F_X^{-1}(u)) \, du. \quad (5.11)$$

This equation means we can transform the valuation problem further into one involving the computation of a Riemann integral over the unit interval $[0,1]$. For a $d$-dimensional diffusion process $X^{t_0,x}$, and subject to certain conditions holding for the joint distribution function $F_X$ of $X^{t_0,x}_T$, the right-hand side of (5.11) can often be written as a standard Riemann integral over the $d$-dimensional unit cube. Note that this version of quasi Monte Carlo requires the density or distribution function for $X^{t_0,x}$ to be known, a requirement which is often difficult to satisfy for many types of valuation problems.

A Monte Carlo estimation of (5.11) would usually involve simulation of the variate $\frac{1}{N} \sum_{i=1}^N h(F_X^{-1}(U_i))$, where $U_i$, $1 \leq i \leq N$, are independent, uniformly distributed random variables, using random or pseudo-random numbers. If these pseudo-random points are replaced by so-called low discrepancy points, see for example Niederreiter (1992), the corresponding procedure is referred to as a quasi Monte Carlo method. Low discrepancy point sets such as Sobol or Halton sequences exhibit less deviations from uniformity compared to pseudo random point sets. This property generally leads to faster rates of convergence compared to pseudo random numbers. Improved versions of certain classes of low discrepancy sequences have recently been developed by Tezuka (1993, 1994) and Tezuka & Tokuyama (1994). Some preliminary investigations on the application of quasi random numbers to the simulation of stochastic differential equations have been conducted by Hofmann & Mathe (1995).

We will now describe a simple, but effective form of quasi Monte Carlo which is related to the work of Barraquand (1993). The method involves a type of regular systematic sampling of the underlying sample space. We consider a $d$-dimensional diffusion process $X^{t_0,x}$ with $m$ driving Wiener processes $W^j$, $1 \leq j \leq m$. Let $X^\Delta$ be a discrete time weak approximation for $X^{t_0,x}$ using two point random variables given by (3.27) and (3.5), respectively. The use of two point random variables, means that for practical purposes we can replace the sample space $\Omega$ with

$$\Omega_{N,m} = \{-1,1\}^{\{0,\ldots,N-1\} \times \{1,\ldots,m\}} \quad (5.12)$$
2.5. OTHER VARIANCE REDUCTION METHODS

which is the multidimensional version of the discrete sample space $\Omega_N$ given in (5.8). For $\omega \in \Omega_{N,m}$ let $\omega_{i,j}$, $i \in \{0, \ldots, N - 1\}$, $j \in \{1, \ldots, m\}$ be the value of $\omega_{i,j}$ at the $(i,j)$th index point. Let $\psi : \{-1,1\} \to \{0,1\}$ be given by $\psi(-1) = 0$, $\psi(1) = 1$ and $Q_{N,m} = \{0, \ldots, 2^{Nm} - 1\}$. Define the mappings $\phi_1, \phi_2 : \Omega_{N,m} \to Q_{N,m}$ by

$$
\phi_1(\omega) = \sum_{i=1}^{N} \sum_{j=1}^{m} \psi(\omega_{i,j}) 2^{m(i-1)+j-1}
$$

$$
\phi_2(\omega) = \sum_{i=1}^{N} \sum_{j=1}^{m} \psi(\omega_{i,j}) 2^{N(j-1)+i-1}
$$

(5.13)

for $\omega \in \Omega_{N,m}$.

These functions define one to one mappings of $\Omega_{N,m}$ onto the set $Q_{N,m}$ and consequently the inverse functions $\phi_1^{-1}, \phi_2^{-1} : Q_{N,m} \to \Omega_{N,m}$ exist. For each integer $k \in Q_{N,m}$ there will be a corresponding path, either $\phi_1^{-1}(k)$ or $\phi_2^{-1}(k)$ which is an element of $\Omega_{N,m}$. The inverse functions $\phi_1^{-1}$ and $\phi_2^{-1}$ determine different scanning orders of $\Omega_{N,m}$. To build a set of say $M$ samples from $\Omega_{N,m}$ we generate a set of $M$ integers from $Q_{N,m}$ and use for example one of the inverse functions $\phi_1^{-1}$ or $\phi_2^{-1}$ to obtain the corresponding paths $\omega \in \Omega_{N,m}$. A number of methods can be used to select the $M$ integers from the set $Q_{N,m}$, see for example Stroud (1971). Barraquand (1993) uses the following algorithm: Let $K$ denote the integer $\lfloor 2^{N/m} M \rfloor$, where $[a]$ is the largest integer not exceeding $a$. We assume $M$ is of the form $M = \gamma p^\gamma$, where $p$ is a prime integer not equal to 2 and $\gamma$ is some positive integer. We construct the samples $\omega_\ell$, $\ell \in \{1, \ldots, M\}$, using say the inverse function $\phi_1^{-1}$ by setting

$$
\omega_\ell = \phi_1^{-1}(\ell K).
$$

(5.14)

Since $MK < 2^{N/m}$, then $\ell K \in Q_{N,m}$ for all $\ell \in \{1, \ldots, M\}$ and consequently this set of samples is well-defined. Other variations of this approach are also possible. We may for instance apply quasi Monte Carlo to each component of the driving Wiener processes separately. This allows the sampling intensity to be adjusted for each component. As an example let $m = 2$, with $\Omega_N$ given by (5.8) and $Q_N = \{0, \ldots, N - 1\}$. We define the mappings $\varphi : \Omega_N \to Q_N$ by

$$
\varphi(\omega) = \sum_{i=1}^{N} \psi(\omega_i) 2^{i-1}
$$

(5.15)

for $\omega \in \Omega_N$. Evidently $\varphi$ is a one to one mapping of $\Omega_N$ onto $Q_N$ and therefore has an inverse $\varphi^{-1}$. We now choose $M_1 = p_1^{K_1}$ and $M_2 = p_2^{K_2}$ as suggested above with $K_1 = \lfloor N/M_1 \rfloor$ and $K_2 = \lfloor N/M_2 \rfloor$ and construct $M = M_1 M_2$ samples $\omega_{\ell,k} \in \Omega_{N,2}$ by the rule

$$
\omega_{\ell,k}^{i,j} = \begin{cases} 
\varphi^{-1}(\ell K_1)_i & : j = 1 \\
\varphi^{-1}(k K_2)_i & : j = 2 
\end{cases}
$$

(5.16)

for $(\ell, k) \in \{1, \ldots, M_1\} \times \{1, \ldots, M_2\}$ and $(i, j) \in Q_N \times \{1, 2\}$.
We remark that the formulation of quasi Monte Carlo described by Barraquand requires the densities of the underlying diffusion to be known as is also the case for the version based on Riemann integration over the $d$-dimensional unit cube, see (5.11). The procedure, outlined above, using the relations (5.12)–(5.16), does not require these densities to be known and therefore represents an important extension to these other methods. In fact this procedure using (5.12)–(5.16) can be applied to any discrete time approximation to a diffusion process as long as the Wiener increments are replaced by multipoint random variables.

Summarizing the different variance reduction techniques developed and discussed in this chapter it is apparent that nearly all of them are very general and can be combined with each other. These features turn out to be important in providing the flexibility needed in the construction and engineering of efficient valuation and hedging systems for a large range of financial modelling problems.
Valuation of Barrier Options under Stochastic Volatility

There is now considerable interest in the valuation and hedging of a range of exotic, and other customized, derivative securities. In this chapter we consider an application of some of the stochastic analytic and numerical methods described in the previous chapters. This application deals with the pricing of foreign exchange barrier options under stochastic volatility with particular reference to the Heston (1993) model.

A barrier option is one the payoff structure of which depends not only on the final price of the underlying security but also on whether the price of the security has hit a pre-determined level or barrier. These are path dependent options since their value depends on the past history of security prices.

Some barrier options return a fixed payoff if the barrier is reached, for example a call or put CAP. Other types called knockout options disappear or become valueless if the barrier is touched. We will examine a class of knockout options called down-and-out call options. These are derivative securities which become valueless or are knocked out if at any time prior to maturity the underlying asset reaches or falls below the barrier. If the barrier is not reached the option returns the standard European call payoff structure. A knockin option is a barrier option which only has some value or comes into exercise if the barrier is hit. Knockout or knockin barrier options are of interest mainly because the possibility of hitting or not hitting the barrier means that they are cheaper than the corresponding standard options.

The observation frequency of a barrier option refers to how often the barrier condition is checked. Clearly this is an important feature of a barrier option since a more frequently observed option will usually be cheaper than a less frequently observed one. For continuously observed barriers, and where the underlying security is assumed to evolve according to the Black-Scholes model, analytic valuations are now available for several types of instruments. For example closed-form solutions for various products have been provided by Merton (1973), Kentwell (1992), Rubinstein & Reiner (1993) and Rich (1993).

Unfortunately these formulas do not in general hold in cases where departures from the Black-Scholes model are permitted. In particular, in recent years researchers have focussed much attention on the merits and effects of allowing for stochastic volatility. However to our knowledge the barrier option pricing problem has not been solved analytically in this environment. Some examples of stochastic volatility models include those proposed by Hull & White (1987), Johnson & Shanno (1987), Scott (1987), Wig-

The main aim of this chapter is to show that fast and accurate numerical valuations are now possible for knockout or knockin barrier options even in a stochastic volatility setting. We demonstrate these methods by computing the prices of down-and-out call options for the Heston model.

We choose the Heston model because analytic valuations are available for standard European options and because this model assumes that volatility movements are random and can be correlated with the returns of the underlying security. These are features which seem to be desirable in a stochastic volatility model, see for example Gatheral (1992).

We remark that the methods developed and used here can also be applied to the valuation of both standard European and barrier options for many other types of stochastic volatility models. The standard European component of these valuation procedures should therefore be of independent interest since many stochastic volatility models have proved to be analytically intractable even for the valuation of these standard instruments.

3.1 The Black-Scholes Framework

Let \( W = \{ W_t, t \geq t_0 \} \) be a one-dimensional Brownian motion defined on the probability space \((\Omega, \mathcal{F}, P)\). As in Section 1.1 we take the filtration \( \mathcal{F} = (\mathcal{F}_t)_{t \geq t_0} \) to be the \( P \)-augmentation of the natural filtration of \( W \). We assume \( P \) is the risk-neutral measure and that we have two deterministic bond price processes \( B^d = \{ B^d_t, t_0 \leq t \leq T \} \) and \( B^f = \{ B^f_t, t_0 \leq t \leq T \} \) for the domestic and foreign markets respectively, together with an exchange rate process \( X = \{ X_t, t_0 \leq t \leq T \} \). The arbitrage free model which describes the dynamics of the bonds and exchange rate processes is given by the following system of stochastic differential equations.

\[
\begin{align*}
    dB^f_t &= r_f B^f_t dt, \\
    dB^d_t &= r_d B^d_t dt, \\
    d(B^f_t X_t) &= r_d B^f_t X_t dt + \sigma B^f_t X_t dW_t, \quad (1.1)
\end{align*}
\]

for \( t_0 \leq t \leq T \) with initial values at time \( t_0 \) of \( B^f_{t_0} = b^f, \ B^d_{t_0} = b^d \) and \( X_{t_0} = x \) and final values at time \( T \) of \( B^f_T = B^d_T = 1 \). With this model the random variable \( B^f_t X_t, t_0 \leq t \leq T \) can be considered as a price adjusted foreign bond (adjusted by the exchange rate) which represents the value of the foreign bond in the domestic economy at time \( t \). Consequently the process \( (B^f X) \) replaces the risky asset in the standard Black-Scholes formulation for stock dynamics. To simplify the notation in what follows we will not include the initial conditions in the symbols used to denote the bond price.
3.1. THE BLACK-SCHOLES FRAMEWORK

and exchange rate processes, as has been done in Chapters 1 and 2. For example we will use $X_t$ rather than $X_t^{t_0,t}$ to denote the value of the exchange rate at time $t$, $t_0 \leq t \leq T$.

The constant values $r_f$ and $r_d$ represent the foreign and domestic interest rates, respectively. The parameter $\sigma$ denotes the volatility of the price adjusted foreign bond $B^f X$.

Using Ito’s formula it follows from the last equation in the system of equations (1.1) that

$$dX_t = (r_d - r_f)X_t \, dt + \sigma X_t \, dW_t,$$  \hspace{1cm} (1.2)

for $t_0 \leq t \leq T$. Note that the bond price processes $B^f$ and $B^d$ can be solved explicitly using the relations

$$B_t^f = e^{-r_f(T-t)}$$

and

$$B_t^d = e^{-r_d(T-t)}$$

for $t_0 \leq t \leq T$. Let us now consider a down-and-out call option on the price adjusted foreign bond process $B^f X$. First we restate the definition of this option - it gives the holder the right to buy units in the foreign currency at time $T$ at the fixed exchange rate $K$; but only if the exchange rate process $X$ has not hit or fallen below the barrier level $H$ before maturity at time $T$.

To model the payoff structure of this type of instrument we will consider the region $\Gamma_0$ defined by

$$\Gamma_0 = [0,T) \times (H, +\infty).$$  \hspace{1cm} (1.4)

Let $\tau : \Omega \to \mathbb{R}^+$ be the stopping time given by (1.1.4) using the process $X$ with corresponding exercise boundary

$$\Gamma_1 = \left\{ (\tau(\omega), X_{\tau(\omega)}) \in [t_0,T] \times \mathbb{R} : \omega \in \Omega \right\}.$$  \hspace{1cm} (1.5)

With these definitions established the payoff structure for a European style down-and-out barrier option denoted by $G : \Omega \to \mathbb{R}^+$ can now be expressed in the form

$$G(\omega) = h(\tau(\omega), X_{\tau(\omega)}(\omega))$$

for $\omega \in \Omega$, where $h : \Gamma_2 \to \mathbb{R}^+$ is some payoff function and $\Gamma_2 = \{(t,H) : t \in [t_0,T] \cup \{(T,y) : y \in (H, +\infty)\}).$

From the continuity of the sample paths of $X$ we see that $\Gamma_1 \subseteq \Gamma_2$ and consequently the payoff structure $G$ as given by (1.6) is well-defined.

For a down-and-out call option we take the function $h : \Gamma_2 \to \mathbb{R}$ to be given by

$$h(t,y) = \begin{cases} 0 & : \text{for } t_0 \leq t \leq T, \, y = H \\ (y-K)^+ & : \text{for } t = T, \, y > H \end{cases}$$

with $K > H$. This definition and (1.6) means that we can express $G$ in the form

$$G = (X_T - K)^+ 1_{(\tau=\infty)}.$$
Let the valuation function \( u : \Gamma_0 \cup \Gamma_1 \rightarrow \mathbb{R} \) be given by (1.1.5). That is, using (1.3), \( u \) can be written in the form

\[
 u(t, x) = E \left( B_t^d h(\tau, X_\tau) \mid X_t = x \right)
\]

\[= B_t^d E (h(\tau, X_\tau) \mid X_t = x) \quad (1.8)\]

for \((t, x) \in \Gamma_0 \cup \Gamma_1\). Define \( \bar{X} = \{ \bar{X}_t, t_0 \leq t \leq T \} \) and \( \bar{u} : \Gamma_0 \cup \Gamma_1 \rightarrow \mathbb{R} \) to be the \( B^d \)-discounted process and valuation function given by

\[
 \bar{X}_t = B_t^f X_t / B_t^d
\]

(1.9)

for \( t_0 \leq t \leq T \), and

\[
 \bar{u}(t, \bar{x}) = E \left( h(\tau, \bar{X}_\tau) \mid \bar{X}_t = \bar{x} \right) \quad (1.10)
\]

for \((t, \bar{x}) \in \Gamma_0 \cup \Gamma_1\), where \( \tau \) is the same stopping time used in (1.5) and (1.6) and corresponds to the process \( X \) (not \( \bar{X} \)).

Expanding \( \bar{X}_t \) using Ito’s rule, (1.9) and the first two equations in (1.1) we see that

\[
d\bar{X}_t = \sigma \bar{X}_t dW_t. \quad (1.11)
\]

Also, applying Ito’s formula for semimartingales, the Kolmogorov backward equation for \( \bar{u} \), which holds because of the form of (1.10), and equation (1.11) we can infer that

\[
 \bar{u}_t = \bar{u}_{t_0} + \int_{t_0}^{t \wedge \tau} \frac{\partial}{\partial \bar{x}} \bar{u}_s d\bar{X}_s \quad (1.12)
\]

for \( t_0 \leq t \leq T \), where \( \bar{u}_t = \bar{u}(t \wedge \tau, \bar{X}_{t \wedge \tau}) \) for \( t_0 \leq t \leq T \).

Let us now apply a dynamic portfolio strategy \( \Phi = (\xi_t, \eta_t)_{t \in [t_0, T]} \), where at time \( t \), \( t_0 \leq t \leq T \) we hold \( \eta_t \) units in the domestic bond \( B_t^d \), and \( \xi_t \) units of the foreign bond with each unit valued at \( B_t^f X_t \) in the domestic economy.

Choose \( \xi_t \) and \( \eta_t \) by

\[
 \xi_t = \frac{\partial}{\partial \bar{x}} \bar{u}_t \quad (1.13)
\]

and

\[
 \eta_t = \bar{u}_t - \xi_t \bar{X}_t. \quad (1.14)
\]

Using the relations (1.3) and (1.9) we see that if \( \tau = T \), then \( \bar{X}_\tau = X_\tau \) and if \( \tau < T \), then from the definition of \( h \) given by (1.7), \( h(\tau, \bar{X}_\tau) = h(\tau, X_\tau) = 0 \), so that

\[
 E(h(\tau, \bar{X}_\tau)) = E(h(\tau, X_\tau)).
\]

This result together with the relations (1.8) and (1.1.9), applied to the processes \( \bar{X} \) and \( X \) means that

\[
 \bar{u}(t \wedge \tau, \bar{X}_{t \wedge \tau}) = u(t \wedge \tau, X_{t \wedge \tau}) / B_t^d \quad (1.15)
\]

for \( t_0 \leq t \leq T \). Note that from (1.15) we can write

\[
 \bar{u}(t, \bar{x}) = u(t, B_t^d \bar{x} / B_t^d) / B_t^d
\]
3.2. A Model with Stochastic Volatility

For practical reasons we would like to use more general classes of models other than the Black-Scholes formulation described by (1.1) above. In particular the assumption of constant volatility is regarded by many individuals as being too restrictive. Consequently we now consider a more general process \( Z = (B^f, B^d, X, v) \), which allows for stochastic volatility and which is defined by the following system of stochastic differential equations:

\[
dB^f_t = r_f B^f_t \, dt,
\]

for

\[
(t, B^d_t x/B^f_t) \in \Gamma_0 \cup \Gamma_1
\]

so that

\[
\frac{\partial}{\partial x} \tilde{u} = \frac{\partial}{\partial x} u/B^f_t
\]

and consequently the hedge ratio \( \xi_t, \ t_0 \leq t \leq T \), can equivalently be expressed in the form

\[
\xi_t = \frac{\partial}{\partial x} u_t/B^f_t. \tag{1.16}
\]

Using (1.14) the value \( u_t \) of the portfolio with strategy \( \Phi \) in the domestic economy (but before discounting by the domestic bond) at time \( t, \ t_0 \leq t \leq T \), satisfies the relation

\[
u_t = \xi_t B^f_t X_t + \eta_t B^d_t.
\]

Moreover using equation (1.8) and the condition \( B^d_t = 1 \) we have

\[
u_T(\omega) = h(\tau(\omega), X_{\tau(\omega)})
\]

for any \( \omega \in \Omega \). However \( G(\omega) = h(\tau(\omega), X_{\tau(\omega)}(\omega)) \) represents the payoff structure for our option for \( \omega \in \Omega \). Consequently our portfolio process fully replicates this payoff structure for any scenario \( \omega \in \Omega \).

In addition, the Ito integral \( \int_{t_0}^{t \wedge \tau} \xi_s \, d\tilde{X}_s \) can be interpreted as the discounted gain from trade resulting from the rate movements of \( X_t, \ t_0 \leq t \leq T \). Consequently from (1.12) and (1.13) our portfolio process is self-financing following an initial cost of \( u_{t_0} = \tilde{u}_{t_0} B^d_{t_0} \). The fair price for the option at time \( t_0 \) is therefore, from (1.15), the value \( u_{t_0} = u(t_0 \wedge \tau, X_{t_0 \wedge \tau}) \). Thus by continuously hedging the portfolio using the strategy \( \Phi \) we can fully replicate the payoff structure \( G \) of the option.

The equations (1.8) together with (1.13) and (1.14) provide a mechanism for determining both the fair price of the option and the corresponding hedge ratios needed to replicate the underlying payoff structure. For the Black-Scholes model described by the system of equations (1.1), these prices and hedge ratios can be computed explicitly, see for example Kentwell (1992).
CHAPTER 3. VALUATION OF BARRIER OPTIONS

\[ dB_t^d = r_d B_t^d dt, \]
\[ dX_t = \mu_X dt + (k_1 v_t + k_2 \sqrt{\nu_t}) X_t dW_t^1, \]
\[ dv_t = \kappa(v_t - \bar{v}) dt + (p_1 v_t + p_2 \sqrt{\nu_t}) \left( \rho dW_t^1 + \sqrt{1 - \rho^2} dW_t^2 \right). \]  

(2.1)

for \( t_0 \leq t \leq T \) with \( k_1, k_2, p_1, p_2 \geq 0, \kappa \leq 0, \rho \in [0, 1] \), initial values, at time \( t_0 \), of \( B_{t_0}^f = b_f, B_{t_0}^d = b_d, X_{t_0} = \bar{x} \) and \( v_{t_0} = \bar{v} \) and final values, at time \( T \), of \( B_T^f = B_T^d = 1 \).

With this system of equations \( W^1 \) and \( W^2 \) represent independent Wiener processes defined on the probability space \((\Omega, \mathcal{F}, P)\).

Let us explain some of the main features of this model. As in the previous section the bond price processes \( B^f \) and \( B^d \) are both deterministic with constant interest rates \( r_f \) and \( r_d \), respectively. The exchange rate process \( X \) follows a generalized geometric Brownian motion with a stochastic diffusion coefficient.

The process \( v \) is closely related to the instantaneous variance of the exchange rate process \( X \). In our model this process is disturbed by some external noise, where \( \rho \) accounts for the correlation between this noise source and the noise of the exchange rate process \( X \). Note that \( v \) is continuously pulled back towards a long term value \( \bar{v} \). The parameter \( \kappa \) measures the strength of the restoring force and is referred to as the mean reversion factor or speed of adjustment.

For parameter values \( k_1 = p_1 = 1, k_2 = p_2 = 0 \) and \( \bar{v} = 0 \) the process \( v \) follows a geometric Brownian motion and can be interpreted as the volatility process of the exchange rate \( X \). For parameter values \( k_1 = p_1 = 0 \) and \( k_2 = p_2 = 1 \) the system of equations corresponds to the Heston model (1993).

We will now briefly consider pricing and hedging procedures for a European style down-and-out barrier option for the system of equations (2.1) with a continuation region \( \Gamma_0 \), exercise boundary \( \Gamma_1 \) and payoff structure \( G \) given by (1.4), (1.5) and (1.6), respectively.

For this type of valuation problem the stochastic volatility in our model creates an intrinsic risk which in general does not allow for the full replication of the underlying payoff structure without extra cost. Following the approach of Foellmer & Schweizer (1991) and Hofmann, Platen & Schweizer (1992), and using (1.1.11) we obtain for a contingent claim with payoff structure \( h(\tau, X_\tau) \) an option pricing formula of the form

\[ u'_t = u'(t \wedge \tau, X_{t \wedge \tau}, v_{t \wedge \tau}) = B_{t \wedge \tau}^d \tilde{E}(h(\tau, X_\tau) | \mathcal{F}_t) \]

\[ = B_t^d \tilde{E}(h(\tau, X_\tau) | \mathcal{F}_t). \]

(2.2)

for \( t_0 \leq t \leq T \), where the expectation is chosen with respect to an appropriately defined probability measure \( \tilde{P} \).

For incomplete markets, for example, if \( p_1 \neq 0 \) or \( p_2 \neq 0 \) in the system of equations (2.1), there is still no general agreement on how to choose this measure. Based on
the arguments presented by Hofmann, Platen & Schweizer (1992) we will choose this measure $\hat{P}$ as the *minimal equivalent martingale* measure.

An *equivalent martingale measure* $\hat{P}$ for the given exchange rate process $X$ is one for which the $B^d$-discounted process $\tilde{X}$ given by (1.9) is an $\mathcal{(\mathcal{F}, \hat{P})}$-martingale and the measures $P$ and $\hat{P}$ have the same nullsets. Thus, an equivalent martingale measure can be interpreted as one which induces a price system which is consistent with having $\tilde{X}$ as an equilibrium exchange rate.

An equivalent martingale measure $\hat{P}$ for $X$ is called *minimal* if any local $P$-martingale $M$ which is orthogonal to $X$ remains a local martingale under $\hat{P}$. Intuitively, $\hat{P}$ is that equivalent martingale measure which is closest to $P$ in a certain sense.

In practical terms, using the minimal equivalent martingale measure $\hat{P}$ has the effect that the actual expected growth rates for all traded stocks change to $r_d - r_f$, and all other nontraded assets are left completely unchanged under the new measure. Thus for our model (2.1), the stochastic differential equations for the components $X$ and $v$ become

$$dX_t = (r_d - r_f)X_t \, dt + (k_1v_t + k_2\sqrt{v_t}) \, X_t \, d\tilde{W}_t^1$$

and

$$dv_t = \kappa(v_t - \bar{v}) \, dt + (p_1v_t + p_2\sqrt{v_t}) \left( \theta \, d\tilde{W}_t^1 + \sqrt{1 - \theta^2} \, d\tilde{W}_t^2 \right)$$

respectively, where $\tilde{W}_1$ and $\tilde{W}_2$ are independent Wiener processes under $\hat{P}$.

With this measure $\hat{P}$ the hedging strategy $\Phi' = (\xi'_t, \eta'_t)_{t \in [0, T]}$ has components which are similar to (1.13) and (1.14) and can be written in the form

$$\xi'_t = \frac{\partial}{\partial x} \bar{u}'_t$$

and

$$\eta'_t = \bar{u}'_t - \xi'_t \bar{X}_t$$

for $t_0 \leq t \leq T$, where $\bar{X}_t$ is given by (1.9) and

$$\bar{u}'_t = \bar{u}'(t \wedge \tau, \bar{X}_{t \wedge \tau}, v_{t \wedge \tau}) = \tilde{E}(h(\tau, \bar{X}_\tau) | \mathcal{F}_t).$$

As in the case for the Black-Scholes model considered in the previous section we can show that the time $t$ value of the portfolio with hedging strategy $\Phi'$ is $u'_t$. Also from (2.2) and the condition $B^d_F = 1$, we can show that $u'_T = h(\tau, X_\tau)$. This means that the hedging strategy $\Phi'$ replicates the claim’s payoff at the terminal time $T$. However, the strategy $\Phi' = (\xi'_t, \eta'_t)_{t \in [0, T]}$ will not in general be riskless. This means the process

$$\bar{C}_t = \bar{u}'_t - \int_{t_0}^{t \wedge \tau} \xi'_s \, d\bar{X}_s$$

of cumulative $B^d$-discounted costs, is not a constant as it is in the classical Black-Scholes case. In fact applying Ito’s formula for semimartingales, (2.3) and (1.11), together with the Kolmogorov backward equation which holds for $\bar{u}'_t$ by (2.6), we have

$$\bar{u}'_t = \bar{u}'_{t_0} + \int_{t_0}^{t \wedge \tau} \xi'_s \, \bar{X}_s + \int_{t_0}^{t \wedge \tau} \frac{\partial}{\partial v} \bar{u}'_s \left( p_1v_s + p_2\sqrt{v_s} \right) \left( \theta \, d\tilde{W}_s^1 + \sqrt{1 - \theta^2} \, d\tilde{W}_s^2 \right),$$
so that from (2.7),

$$C_t = \bar{u}'_t + \int_{t_0}^{t \wedge \tau} \frac{\partial}{\partial u} \bar{u}_s' (p_1 v_s + p_2 \sqrt{v_s}) \left( \sigma d\tilde{W}_s^1 + \sqrt{1 - \sigma^2} d\tilde{W}_s^2 \right)$$  \hspace{1cm} (2.8)

for \( t_0 \leq t \leq T \). This means that variance of \( \bar{C}_t \) under \( \tilde{P} \) denoted by \( \bar{\text{Var}}(\bar{C}_t) \) can be calculated using the relation

$$\bar{\text{Var}}(\bar{C}_t) = \mathbb{E} \left( \left( \int_{t_0}^{t \wedge \tau} \frac{\partial}{\partial u} \bar{u}_s' (p_1 v_s + p_2 \sqrt{v_s}) \left( \sigma d\tilde{W}_s^1 + \sqrt{1 - \sigma^2} d\tilde{W}_s^2 \right) \right)^2 \right)$$

$$= \mathbb{E} \left( \int_{t_0}^{t \wedge \tau} \left( \frac{\partial}{\partial u} \bar{u}_s' (p_1 v_s + p_2 \sqrt{v_s}) \right)^2 ds \right).$$

This result shows that in general \( \bar{\text{Var}}(\bar{C}_t) > 0 \) for \( t_0 \leq t \leq T \) so that \( \bar{C}_t \) fluctuates randomly. Consequently the strategy \( \Phi' \) is not self-financing in these incomplete market circumstances. But the choice of the probability measure \( \tilde{P} \) will be such that the \( B^d \)-discounted cost process \( \bar{C} \) becomes an \( (\mathcal{F}, \tilde{P}) \)-martingale as can be seen from (2.8). This makes the strategy \( \Phi' \) mean-self-financing, that is

$$\mathbb{E} \left\{ \bar{C}_T - \bar{C}_t \mid \mathcal{F}_t \right\} = 0.$$  \hspace{1cm} (2.9)

Moreover, it can be shown that \( \Phi' \) minimizes the remaining risk

$$R_t(\Phi') = \mathbb{E} \left\{ (\bar{C}_T - \bar{C}_t)^2 \mid \mathcal{F}_t \right\}.$$  \hspace{1cm} (2.10)

We remark that the hedging strategy \( \Phi' \) given by (2.4) and (2.5) using the minimal equivalent martingale measure, is mean self-financing also in the case where both \( r_d \) and \( r_f \) are stochastic, see Heath & Platen (1992).

### 3.3 Numerical Procedures for Barrier Options

The problem with general systems of stochastic differential equations of the type described in (2.1) and (2.3) is that there is usually no explicit solution for the option pricing formula (2.2) or hedging strategy (2.4) and (2.5). In these cases we usually require the application of stochastic numerical and other related approximation methods to estimate the solution.

As indicated by (2.2) computation of an option price for the stochastic volatility model (2.1) with the adjusted equations (2.3) requires an estimate of the expectation, under the measure \( \tilde{P} \), of functionals of the underlying diffusion process. For this type of problem we do not require strong or pathwise approximations; rather it is sufficient to approximate the underlying probability law of the diffusion process.

The class of numerical schemes called weak approximations are designed to approximate these probability laws and are therefore suitable for option pricing estimates. These schemes are classified according to their weak order of convergence, which is defined as follows:
3.3. NUMERICAL PROCEDURES FOR BARRIER OPTIONS

Let \((t)_\Delta\) be an equi-spaced discretization grid of the form (2.3.1) with step size \(\Delta = (T - t_0)/N\). We say that an approximation \(Y_\Delta = \{Y_k^\Delta, k \in \{0, \ldots, N\}\}\) for the \(d\)-dimensional diffusion process \(Y\) converges with weak order \(\beta > 0\) as the step size \(\Delta\) tends to 0 if there exist constants \(K > 0\) and \(\delta_0 < T\) such that for every function \(g : \mathbb{R}^d \to \mathbb{R}\) from a given class \(C_p\) of test functions we have for all \(\Delta \in (0, \delta_0)\) the inequality

\[
|E(g(Y_T)) - E\left(g\left(Y_T^{\Delta}\right)\right)| \leq K \Delta^\beta.
\]

For the class \(C_p\) of test functions we may use for instance the polynomials. This choice allows a clear classification of a wide range of numerical schemes and also includes the convergence of all moments of \(Y_N\) and \(Y_\Delta\). For example, the Euler scheme (2.3.27) converges under sufficient regularity conditions, applied to the drift \(a\) and diffusion \(b\) coefficients, with weak order \(\beta = 1.0\). A more complete coverage of stochastic numerical procedures and their applications, including issues relating to strong and weak orders of convergence, is provided by Kloeden & Platen (1992).

With reference to the numerical experiments described in the next section we used a derivative free method of weak order \(\beta = 2.0\) due to Platen (1984) which as an approximation for the \(d\)-dimensional diffusion process \(Z = (Z^1, \ldots, Z^d)\) given by

\[
dZ_t = a(t, Z_t) dt + \sum_{j=1}^{m} b^j(t, Z_t) dW_t
\]

has the form

\[
Y_{k+1}^\Delta = Y_k^\Delta + \frac{1}{2} \left( a(Y) + a(Y_k^\Delta) \right) \Delta
\]

\[
+ \frac{1}{4} \sum_{j=1}^{n} \left[ \left( b^j(\bar{R}_k^j) + b^j(\bar{R}_k^j) + 2b^j(\bar{Y}_k^\Delta) \right) \Delta \bar{W}_k^j
\]

\[
+ \sum_{j=r \neq j}^{n} \left( b^j(\bar{U}_k^j) + b^j(\bar{U}_k^j) - 2b^j(\bar{Y}_k^\Delta) \right) \Delta \bar{W}_k^j \Delta^{-\frac{1}{2}} \right]
\]

\[
+ \frac{1}{4} \sum_{j=1}^{n} \left[ \left( b^j(\bar{R}_k^j) - b^j(\bar{R}_k^j) \right) \left( (\Delta \bar{W}_k^j)^2 - \Delta \right)
\]

\[
+ \sum_{j=1}^{n} \left( b^j(\bar{U}_k^j) - b^j(\bar{U}_k^j) \right) \left( \Delta \bar{W}_k^j \Delta \bar{W}_k^j + \bar{V}_{r,s} \right) \right] \Delta^{-\frac{1}{2}}
\]

for \(k \in \{0, \ldots, N - 1\}\) with supporting points

\[
\bar{Y} = Y_k^\Delta + a(Y_k^\Delta) \Delta + \sum_{j=1}^{n} b^j(Y_k^\Delta) \Delta \bar{W}_k^j,
\]

\[
\bar{R}_k^j = Y_k^\Delta + a(Y_k^\Delta) \Delta \pm b^j(Y_k^\Delta) \sqrt{\Delta},
\]

\[
\bar{U}_k^j = Y_k^\Delta \pm b^j(Y_k^\Delta) \sqrt{\Delta}.
\]
CHAPTER 3. VALUATION OF BARRIER OPTIONS

where \( \Delta \hat{W}_k^j \), \( j \in \{1, \ldots, m\} \), \( k \in \{0, \ldots, N - 1\} \) are chosen as independent \( N(0, \Delta) \) Gaussian distributed random variables under the measure \( \tilde{P} \). These random variables correspond to the \( m \) independent driving Wiener processes in the underlying diffusion process \( Z \).

In this scheme we also choose the variates \( V_{j_1,j_2} \), \( j_1, j_2 \in \{1, \ldots, m\} \) as two-point random variables with

\[
\tilde{P}(V_{j_1,j_2} = \pm \Delta) = \frac{1}{2} \quad \text{for} \quad j_2 = 1, \ldots, j_1 - 1,
\]

\[
V_{j_1,j_1} = -\Delta,
\]

and

\[
V_{j_1,j_2} = -V_{j_2,j_1} \quad \text{for} \quad j_2 = j_1 + 1, \ldots, m.
\]

For the Heston model under consideration we use the value \( m = 2 \) because of the form of (2.3).

Let us now consider the problem of using Monte Carlo simulation to approximate the option price \( u'_{t_0} \) at time \( t_0 \), given by equation (2.2). If we use the discrete time weak approximation \( Y^\Delta \) given by (3.2) for the vector diffusion process \( Z(X,v) \) given by (2.3) we can estimate the corresponding option price \( u'(t_0, \bar{x}) \) at time \( t_0 \) with the conditional expectation

\[
B_{t_0}^d \tilde{E} \left( h \left( \tau^\Delta, Y^\Delta_{\tau^\Delta} \right) \bigg| Y_0 = \bar{x} \right), \tag{3.3}
\]

where the stopping time \( \tau^\Delta : \Omega \to \mathbb{R} \) is given by

\[
\tau^\Delta(\omega) = \inf \left\{ t_i : \left( t_i, Y^\Delta_{t_i} \right) \not\in \Gamma_0, \ i \in \{1, \ldots, N\} \right\} \tag{3.4}
\]

and the function \( \pi^\Delta : \Omega \to \{1, \ldots, N\} \) is defined by

\[
\pi^\Delta(\omega) = \inf \left\{ i : \left( t_i, Y^\Delta_{t_i} \right) \not\in \Gamma_0, \ i \in \{1, \ldots, N\} \right\}.
\]

for \( \omega \in \Omega \).

A Monte Carlo estimation of (3.3) would involve generating the outcomes \( Y^\Delta(\omega_i) \) for say \( M \) paths \( \omega_i, \ i \in \{1, \ldots, M\} \), using the numerical scheme (3.2) and then computing the sample mean given by

\[
\frac{1}{M} B_{t_0}^d \sum_{i=1}^{M} h \left( \tau^\Delta(\omega_i), Y^\Delta_{\tau^\Delta(\omega_i)}(\omega_i) \right) \tag{3.5}
\]

which would be the estimate of the option price, where each \( h(\tau^\Delta(\omega_i), Y^\Delta_{\tau^\Delta(\omega_i)}(\omega_i)) \), \( i \in \{1, \ldots, M\} \) represents an independent realization of the random variable \( h(\tau^\Delta, Y^\Delta_{\tau^\Delta}) \).

We can ensure that the expectation (3.3) is close to the option price \( u'(t_0, \bar{x}) \) by the use of appropriate numerically stable higher order schemes, however the closeness of the two estimates (3.3) and (3.5) depends ultimately on the variance of \( h(\tau, X_{\tau}) \). Increasing the sample size \( M \) of our simulation, generally reduces the variance but only with order \( M^{-\frac{1}{2}} \) as \( M \to \infty \).
3.3. NUMERICAL PROCEDURES FOR BARRIER OPTIONS

We therefore require other estimators; ones which have the same or nearly the same expectation but smaller variance. The variance reduction procedures described in Chapter 2 can be used to construct these estimators.

We will now describe a variance reduction technique closely related to the control variate formulations (2.4.3) and (2.4.7) and which was found to be effective for the computation of down-and-out call prices for stochastic volatility models of the form (2.3). This technique was incorporated in the numerical procedures whose results, for the Heston model, are described in the next section. The main idea with this method is to simulate only the difference between the Heston model and another which is close to the Heston formulation and for which a known explicit formula exists for the option price. The Black-Scholes framework is clearly a reasonable choice as a generator of control variates for the Heston model.

To be more explicit we consider two vector valued processes 

\[
Z = Z_t = (B^f_t, B^d_t, X_t, v_t), \quad t_0 \leq t \leq T \] and \( \hat{Z} = \hat{Z}_t = (\hat{B}^f_t, B^d_t, \hat{X}_t, \hat{v}_t), \quad t_0 \leq t \leq T \]

defined on the same probability space \((\Omega, \mathcal{F}, \mathbb{P})\) by

\[
dB^f_t = r_f B^f_t \, dt \\
dB^d_t = r_d B^d_t \, dt \\
dx_t = (r_d - r_f) X_t \, dt + \sqrt{v_t} X_t \, d\hat{W}^1_t \\
dv_t = \kappa (v_t - \bar{v}) \, dt + \sigma \sqrt{v_t} \left( \varphi \, dW^1_t + \sqrt{1 - \varphi^2} \, dW^2_t \right) \\
d\hat{x}_t = (r_d - r_f) \hat{x}_t \, dt + \hat{v} \hat{x}_t \, d\hat{W}^1_t \\
d\hat{v}_t = 0
\]

for \( t_0 \leq t \leq T \) with initial values \( B^f_{t_0} = b^f, \ B^d_{t_0} = b^d, \ X_{t_0} = \bar{x} \) and \( v_{t_0} = \underline{v} \), at time \( t_0 \), and final values \( B^f_T = B^d_T = 1, \) at time \( T \), where \( \hat{W}^1 \) and \( \hat{W}^2 \) are independent Wiener processes under the measure \( \hat{\mathbb{P}} \). Here the processes \( Z \) and \( \hat{Z} \) correspond to the Heston and Black-Scholes models respectively. The initial value \( \hat{v}_{t_0} \) can be chosen so that the process \( Z \) and \( \hat{Z} \) are close in some reasonable sense. For example, one possible choice for \( \hat{v}_{t_0} \) is to let it be equal to the square root of the average value of \( \{ v_t, \ t_0 \leq t \leq T \} \) as it would evolve according to the Heston model but with no noise component, that is with \( \sigma = 0 \).

In this case \( v_t \) can be solved explicitly with

\[
v_t = \bar{v} + (\underline{v} - \bar{v}) e^{\kappa (t-t_0)}
\]

for \( t_0 \leq t \leq T \), so that

\[
\hat{v}_{t_0} = \sqrt{\frac{1}{T-t_0} \int_{t_0}^{T} v_t \, dt}
\]
CHAPTER 3. VALUATION OF BARRIER OPTIONS

\[
\sqrt{\bar{v} + \frac{(\bar{v} - \bar{u})}{\kappa (T - t_0)}} (e^{\kappa T} - 1)
\]  

Let the stopping time \( \tau : \Omega \rightarrow \mathbb{R} \) be given by (1.1.4) using the process \( X \) and the region \( \Gamma_0 \) given by (1.4). Define the stopping time \( \tilde{\tau} : \omega \rightarrow \mathbb{R} \) by

\[
\tilde{\tau}(\omega) = \inf\{t_0 : (t, X_t) \notin \Gamma_0\}.
\]

The option price at time \( t_0 \) for the process \( \tilde{Z} \), denoted by \( \tilde{u}'(t_0, \bar{x}) \) can be computed from the formula (2.2). Thus

\[
\tilde{u}'(t_0, \bar{x}) = B_{t_0} d E_h(T, X_t) - \tilde{u}'(t_0, \bar{x})/B_{t_0} d),
\]

where we recall that \( \bar{X} \) has the initial value \( \bar{X}_{t_0} = \bar{x} \) at time \( t_0 \). This option price corresponds to the case of a continuously observed down-and-out call for a Black-Scholes model with constant volatility and is known explicitly. This fact will be used in the control variate formulation described below.

Consider the random variable

\[
\tilde{Z}_{\tau, \bar{x}} = B_{t_0} d \tilde{E}(h(T, X_t) - \tilde{u}'(t_0, \bar{x}))
\]

which using (3.10) can be written in the form

\[
\tilde{Z}_{\tau, \bar{x}} = B_{t_0} d \tilde{E}(h(T, X_t) - \tilde{u}'(t_0, \bar{x}))/B_{t_0} d)
\]

Since

\[
\tilde{E}(\tilde{Z}_{\tau, \bar{x}}) = B_{t_0} d \tilde{E}(h(T, X_t)) = u'(t_0, \bar{x}),
\]

by equation (2.2), \( \tilde{Z}_{\tau, \bar{x}} \) is an unbiased estimator for \( u'(t_0, \bar{x}) \) which is the option price we want to compute. If \( \tilde{Z} \) is close to \( Z \), which is the case for reasonable choices of the parameters \( r_f, r_d, \bar{v}, \sigma, \rho, \bar{x} \) and \( \bar{u} \), with \( \bar{u}_{t_0} \) chosen according to (3.8), then the variance of the estimator \( \tilde{Z}_{\tau, \bar{x}} \) will be much smaller than the variance of \( B_{t_0} d h(T, X_t) \). Consequently the corresponding statistical error will be smaller than that obtained from a standard Monte Carlo simulation of \( u'(t_0, \bar{x}) \) using the variate \( B_{t_0} d h(T, X_t) \).

At this point we replace the diffusion processes \( Z \) and \( \tilde{Z} \) with corresponding discrete time weak approximations \( Y^\Delta \) and \( \tilde{Y}^\Delta \), respectively, using the numerical scheme (3.2). The discrete time representation of the estimator \( \tilde{Z}_{\tau, \bar{x}} \) denoted by \( \tilde{Z}_{\tau, \bar{x}}^\Delta \), now takes the form

\[
\tilde{Z}_{\tau, \bar{x}}^\Delta = B_{t_0} d \tilde{E}(h(T, X_t) - \tilde{u}'(t_0, \bar{x}))/B_{t_0} d)
\]

where \( \tau^\Delta \) and \( \pi^\Delta \) are given by

\[
\tau^\Delta = \inf \{t_i : (t_i, Y^\Delta_i) \notin \Gamma_0, 1 \leq i \leq N\},
\]

\[
\pi^\Delta = \inf \{i \in \{1, \ldots, N\} : (t_i, Y^\Delta_i) \notin \Gamma_0\},
\]
3.3. NUMERICAL PROCEDURES FOR BARRIER OPTIONS

and \( \hat{\tau}^\Delta \) and \( \hat{\tau}^\Delta \) are defined in a similar fashion except we replace the discrete time approximations \( Y_i^\Delta \) with \( \hat{Y}_i^\Delta, i \in \{1, \ldots, N\} \).

A Monte Carlo estimation of the option price \( \hat{u}'(t_0, \bar{z}) \) using (3.12) would be performed in a similar fashion to that for the estimate based on equation (3.3) and given in (3.5). That is we would obtain say \( M \) outcomes \( \hat{Z}_\delta^\Delta_t(\omega_i) \) and compute the sample mean \( \hat{u}^\Delta(t_0, \bar{z}) = \frac{1}{M} \sum_{i=1}^{M} \hat{Z}_\delta^\Delta_t(\omega_i) \). The optimal value of \( \alpha \) to minimize the sample variance \( \frac{1}{M-1} \sum_{i=1}^{M} (\hat{Z}_\delta^\Delta_t(\omega_i) - \hat{u}^\Delta(t_0, \bar{z}))^2 \) can be obtained as the simulation proceeds using (2.4.4).

It can be seen from the formulation of this variance reduction technique that it is very general and can be applied to a wide class of stochastic volatility models and other valuation problems. It can also be conveniently combined with the use of antithetic variates and stratified sampling as described in Section 2.5.

Note also that when the value of the parameter \( H \) is very low, the option price \( u'(t_0, \bar{z}) \) of a down-and-out call approaches that of a European call option. Consequently these procedures enable us to calculate standard European calls but in a stochastic volatility setting. This result is of independent interest since, for the Heston model under consideration, the closed-form valuation procedures provided by Heston (1993), which rely on the inversion of certain characteristic functions in the complex plane, are difficult to implement.

Furthermore, these procedures can be adapted to take into account the observation frequency of the option. This is of considerable practical value as the barrier condition, for all traded instruments of this kind, are in fact observed and tested only at discrete points in time. This is usually daily but sometimes can be less frequent. Clearly the observation frequency of a barrier option can have a significant effect on the price of the option. To see how these methods can be changed to suit the observation frequency of the option let \( \{t_{ij} \colon j \in \{0, \ldots, J\}\} \) for some integer \( J \leq N \) be a subset of time points from our discretization grid \( \{t_i \colon i \in \{0, \ldots, N\}\} \) with \( t_{i_0} = t_0 \) and which corresponds to the times or fixings at which the barrier condition is checked. Thus we assume our discretization grid \( t_{ij} \) is finer than the fixings for our barrier option.

We now replace the estimator \( \hat{Z}_\delta^\Delta_t \) with

\[
\hat{Z}_\delta^\Delta_t = B^d_{t_0} \left( h(\hat{\tau}^\Delta, Y^\Delta_{\delta, \Delta}) - \alpha \left( h(\hat{\tau}^\Delta, \hat{Y}^\Delta_{\delta, \Delta}) - \hat{u}'(t_0, \bar{z}) \right) / B^d_{t_0} \right),
\]

where

\[
\hat{\tau}^\Delta = \inf \left\{ t_{ij} \colon (t_{ij}, Y_{ij}^\Delta) \notin \Gamma_0, j \in \{1, \ldots, J\} \right\}
\]

and

\[
\hat{\tau}^\Delta = \inf \left\{ t_{ij} \colon (t_{ij}, Y_{ij}^\Delta) \notin \Gamma_0, j \in \{1, \ldots, J\} \right\}.
\]

That is, for the component \( h(\hat{\tau}^\Delta, Y^\Delta_{\delta, \Delta}) \) of the estimator \( \hat{Z}_\delta^\Delta_t \) we use the fixings of the barrier option as the times to stop the approximation \( Y^\Delta_{ij}, j \in \{1, \ldots, \tau\} \). However for the control variate \( h(\hat{\tau}^\Delta, \hat{Y}^\Delta_{\delta, \Delta}) - \hat{u}'(t_0, \bar{z}) \) we use the whole grid \( t_{ij} \), to determine
the times at which the approximation $Y^\Delta_i$, $i \in \{1, \ldots, N\}$ should be stopped. This is necessary as we want to ensure, see (3.10), that

$$E \left( h \left( t^\Delta, Y^\Delta \right) - \tilde{u}'(t_0, \mathbf{z}) / B_{t_0}^d \right) \approx 0.$$ 

Since the option price $\tilde{u}(t_0, \mathbf{z})$ is obtained from a continuously observed barrier we use the whole grid $(t)_\Delta$.

We remark finally that all of the methods and results described in this chapter can be adapted to other types of barrier options, such as down-and-out puts or up-and-in calls, in a stochastic volatility setting. In addition even extra features such as double or partial barriers can still be accommodated with these methods. For example a partial down-and-out call for the Heston model would be similar to the usual down-and-out call except the barrier condition would only be applied for a subset of the interval $[t_0, T]$. In this case we could reasonably expect that a good control variate would be a linear combination $\alpha_1 Y_1 + \alpha_2 Y_2$ of the form (2.4.17), where $Y_1$ is obtained from a standard European call and $Y_2$ is obtained from a down-and-out call for a corresponding Black-Scholes model. As is the case for the estimator $\tilde{z}^\Delta_{\tau, \tilde{r}}$ or $\tilde{z}^\Delta_{\tau, \tilde{r}, \tilde{t}}$, we can compute the optimal values of the coefficient vector $(\alpha_1, \alpha_2)$ using the multidimensional analogue of (2.4.4), namely (2.4.19).

### 3.4 Simulation Results for Barrier Options

As has been previously mentioned there is no explicit solution for the option price or hedging strategy for the model described by the system of equations (2.3). However using the numerical techniques described in the previous section we can obtain fast and accurate valuations. For the numerical experiments described in this section we employed the higher order approximation (3.2) to reduce the systematic error, that is the difference between $\tilde{Z}_{\tau, \tilde{r}}$ given by (3.11) and $\tilde{Z}_{\tau, \tilde{r}, \tilde{t}}$ given by (3.12), see Kloeden & Platen (1992). We also incorporated the three variance reduction methods of control and antithetic variates and stratified sampling as outlined in the previous section and Section 2.5 to minimize the statistical error, see again Kloeden & Platen (1992).

For these simulation experiments we used the Heston and control variate models corresponding to the vector process $Z$ and $\tilde{Z}$, respectively, given by (3.6). For simplicity we used the values $r_d = r_f = 0$. This means according to the first equation in (2.3) that there is no drift component in the stochastic differential equation for the exchange rate $X$ under the minimal equivalent martingale measure $\tilde{P}$. The other parameters were assigned the following default values: $H = 95.0$, $K = 100.0$, $\kappa = -2.0$, $\bar{\sigma} = 0.01$, $\sigma = 0.2$, $\varrho = 0.0$, $T = 0.5$ with initial values $X_0 = z = 100.0$ and $V_0 = \nu = 0.01$ at time $t_0 = 0$.

The statistical errors and associated confidence intervals were estimated by dividing the total number of outcomes into say $L$ batches. The sample means were taken within each batch to form asymptotically Gaussian statistics. The means $\tilde{\mu}_L$ and sample
3.4. SIMULATION RESULTS FOR BARRIER OPTIONS

variance $\hat{\sigma}_L^2$ of these statistics were then taken over the batches. We obtain statistical error bounds at a 99% confidence level by forming the interval $(\hat{\mu}_L - a_L, \hat{\mu}_L + a_L)$, where $a_L = t_{0.99,L-1} \sqrt{\hat{\sigma}_L^2 / L}$ and $t_{0.99,L-1}$ is the value of the Student t-distribution with $L - 1$ degrees of freedom evaluated at a confidence level of 99%.

For the numerical scheme (3.2) we used $N = 16$ discretization points with $M = 20$ batches, each with 256 trajectories. The paths in each batch were themselves divided into 64 groups of 4, constructed by means of an antithetic variate generation procedure as follows:

For $N$ discretization points, define $\chi : \{0, \ldots, N - 1\} \rightarrow \{-1, 1\}$ by

$$\chi(i) = \begin{cases} +1 & : i \leq (N - 1)/2 \\ -1 & : i > (N - 1)/2 \end{cases}$$

for $i \in \{0, \ldots, N - 1\}$. Let $(\Delta \hat{W}_k^1, \Delta \hat{W}_k^2)$, $k \in \{0, \ldots, N - 1\}$, be the Wiener increment approximations used in the numerical scheme (3.2). As has been noted previously we use the value $m = 2$ because there are two independent driving Wiener processes in the formulation of the Heston model given by (3.6) using the process $Z$. A single realization for the control variate estimator $\hat{Z}_{r,T}^\Delta(\omega_1)$, $\omega_1 \in \Omega$ given by (3.12) is obtained by determining the $2N$ outcomes $(\Delta \hat{W}_k^1, \Delta \hat{W}_k^2)$, $k \in \{0, \ldots, N - 1\}$. With these outcomes we compute simultaneously the additional outcomes

$$(-\Delta \hat{W}_k^1, -\Delta \hat{W}_k^2),$$

$$\left(\chi(k) \Delta \hat{W}_k^1, \chi(k) \Delta \hat{W}_k^2\right),$$

$$(-\chi(k) \Delta \hat{W}_k^1, -\chi(k) \Delta \hat{W}_k^2)$$

for $k \in \{0, \ldots, N - 1\}$. These three, antithetically produced, sets of outcomes are then substituted into the numerical scheme (3.2) to produce three additional realizations for the estimator $\hat{Z}_{r,T}^\Delta$, say $\hat{Z}_{r,T}^\Delta(\omega_2)$, $\hat{Z}_{r,T}^\Delta(\omega_3)$ and $\hat{Z}_{r,T}^\Delta(\omega_4)$, $\omega_2, \omega_3, \omega_4 \in \Omega$. This method thus combines full reflection of both independent Wiener components and partial reflections for approximately half of the time interval $[0, T]$. The procedure is computationally efficient since we require only one original set of $2N$ pseudo or quasi random numbers to produce the four realizations for the estimator $\hat{Z}_{r,T}^\Delta$.

The stratified sampling method used in these simulation experiments was based on a two-dimensional version of the example described in Section 2.5. With this technique we replace the Gaussian increments used in (3.2) by corresponding two point random variables with probabilities given by (2.3.5). Multipoint approximations for the Gaussian increments can also be used however this extension of the basic method was not tested in the simulation experiments outlined in this section.

If we let $\Omega_{N,2}$ be the discrete sample space given by (2.5.12) the partitioning sets $A_a$, $a \in \Omega_{2,2}$ given by (2.5.8) take the form

$$A_a = \{\omega \in \Omega_{N,2}: \omega_{k,1} = a_{k,1}, \omega_{k,2} = a_{k,2} \text{ for } k \in \{0, 1\} \}$$

(4.2)
CHAPTER 3. VALUATION OF BARRIER OPTIONS

With this version of stratified sampling we obtain $2^4 = 16$ partitioning sets $A_a$, $a \in \Omega_{2,2}$ and sample independently and separately within each set. As stated in Section 2.5 this method not only reduces the variance for the estimator $\hat{Z}_{r,t}^{\Delta}$, it is also computationally efficient in that it avoids duplicate traversals of the early nodes of the lattice which results from the use of two point variates.

Further reductions in the variance could probably be obtained by using the third quasi Monte Carlo technique outlined in Section 2.5, see in particular the equations (2.5.12)-(2.5.16). Note that this form of quasi Monte Carlo requires the use of multipoint approximations for the increments $\Delta \hat{W}_k^j$, $j \in \{1, \ldots, m\}$, $k \in \{0, \ldots, N - 1\}$ used in the numerical scheme (3.1). This error reduction procedure can be conveniently combined with the control variate and stratified sampling methods previously described. However the use of quasi Monte Carlo in this form would tend to reduce the effectiveness of antithetic variates and would, most likely, partially or fully replace these variates in any practical implementation of the method. Note that this additional technique was not tested in the simulation results presented here.

Using a 486, 33 MHz personal computer, with 16 discretization points and 5120 (= $20 \times 256$) sample paths, option prices can typically be computed within 10 seconds. For all of the numerical results presented in this section a relative statistical error, based on the criteria given above, of 0.1% at a 99% confidence level was achieved.

The instantaneous variance $\nu_t$ of the exchange rate evaluated at time $t$, $t_0 \leq t \leq T$, has a stationary distribution with $P$-a.s. positive values, whenever $-\kappa \bar{v} \geq \frac{1}{2} \sigma^2$. Consequently for these default parameter settings the value for $\sigma$ is the maximum possible value and produces the most pronounced stochastic volatility effects. These choices for the model parameters also represent a worst case scenario for the valuation procedures and software, as they generate the largest corresponding error terms.

Figure 3.4.1 shows a typical pattern of prices for down-and-out calls for both the Heston and Black-Scholes models using different values of the barrier level $H$. For the Black Scholes model we used the process $\hat{Z}$ defined in (3.6) together with the initial value $\hat{v}_0$, at time 0, given by (3.8). For the Heston model using the process $Z$, again defined in (3.6), we used the default value, $\sigma = 0.2$. As previously mentioned this means that a strong stochastic volatility effect is incorporated and that relatively large price differences between the two models of the order of $5 - 7\%$ result for barrier levels below 95% of the spot exchange rate $X_0 = \bar{x} = 100.0$.

Note that for low values of the barrier level $H$, the barrier effect is reduced and we obtain corresponding European call prices for the Heston and Black-Scholes models, respectively. Clearly for the default parameters used the Heston model returns lower prices, however for other settings higher prices can be obtained. A three dimensional representation of these results for the Heston model using different values for the barrier level $H$ and times to maturity $T$ is given in Figure 3.4.2.

An important consideration for financial institutions dealing with exotic options are the risks associated with trading in these instruments. One of the reasons for the
3.4. SIMULATION RESULTS FOR BARRIER OPTIONS

Figure 3.4.1: Option prices for the Heston and Black-Scholes models for different levels of the barrier level $H$.

interest in the Heston model is its potential to provide the basis for better hedging of the underlying security. Hedge ratios for both the Heston and Black-Scholes model are illustrated in Figure 3.4.3 using different values for the spot exchange rate $X_0$ and a barrier level $H = 80.0$. These hedge ratios were computed from central finite differences and the technique of common random number generation, see for example Ross (1991) or Law & Kelton (1991). For the default parameter settings and values of $X_0$ in the range $90.0 \leq X_0 \leq 95.0$ hedge ratio differences of the order of $5 - 10\%$ were observed. For the Black-Scholes model, we calculated the initial value $\hat{v}_0$ according to (3.8) as has been explained for the results shown in Figure 3.4.1.

Figure 3.4.4 displays price differences $(v'_0 - \hat{v}_0)$ between the Heston and Black-Scholes models using different values for the spot exchange rate $X_0$ and times to maturity $T$. The values for the other parameters used are as given in the default parameter set except for the value of $H$ which was set at $95\%$ of the level of $X_0$. The parameter $\hat{v}_0$ for the Black-Scholes model was again determined from (3.8). This figure clearly illustrates a version of the smile effect in prices which has been observed empirically for many instruments.

A different view of similar results showing the smile effect in prices can be obtained if we keep the time to maturity $T$ constant at the default value $T = 0.5$ and change the barrier level $H$ as a percentage of the spot exchange rate $X_0$. This view of price differences $(v'_0 - \hat{v}_0)$ is shown in Figure 3.4.5.
CHAPTER 3. VALUATION OF BARRIER OPTIONS

Figure 3.4.2: Option prices for the Heston model for different levels of the barrier and time to maturity.

Figure 3.4.3: Hedge ratios for Heston and Black-Scholes models for different values of the exchange rate.
3.4. SIMULATION RESULTS FOR BARRIER OPTIONS

Figure 3.4.4: Price differences between the Heston and Black-Scholes models using different exchange rates and times to maturity.

Figure 3.4.5: Price differences between the Heston and Black-Scholes models using different levels of the barrier and times to maturity.
Chapter 4

Valuation of European Bond Options for Two-Factor Interest Rate Models

In this chapter a general methodology is developed for pricing discount bonds and European style contingent claims for multifactor term structure models. We demonstrate this approach by efficiently computing the prices of discount bonds and European call options on bonds for a version of the Fong & Vasicek (1991a,b) model. This version is extended to include time dependent parameters and we compute these prices, using a combination of stochastic and deterministic numerical methods.

Fong and Vasicek provide analytic solutions for the pricing of discount bonds but not options and not in cases where key parameters are given by time dependent functions. We consider the use of time dependent parameters because they allow for greater flexibility in fitting the model to actual market data.

A large number of different term structure models have been considered in recent years. Some of the more popular models have been proposed by Black (1976), Vasicek (1977), Cox, Ingersoll & Ross (1985), Ho & Lee (1986), Longstaff (1989), Black, Derman & Toy (1990), Hull & White (1990), Fong & Vasicek (1991a,b) and Heath, Jarrow & Morton (1992).

The question regarding what factors to include in a multifactor term structure model is an important one. For the extended Fong and Vasicek model considered in this chapter the short rate and the instantaneous variance of the short rate are chosen as the two factors to be represented. These choices are supported by the empirical results obtained by Dybvig (1989). Other interesting multi-factor stochastic volatility models have been proposed for example by Longstaff & Schwartz (1992a,b), Duffie & Kan (1994), Ritchken & Sankarasubramanian (1995) and Brace & Musiela (1994).

Note that it will not be our goal to evaluate the merits of the Fong and Vasicek model or to compare it with other models. Also, the important task of parameter estimation and calibration is not considered here. In this sense we have considered only part of the work required to turn a specific model into a practical pricing tool. The Fong and Vasicek model is chosen because it incorporates some of the features that seem to be desirable in an interest rate term structure model and because it provides an effective vehicle to illustrate the pricing and numerical methods developed.

We emphasize that the methods developed in this chapter can also be applied to other multifactor models. However for simplicity we focus on the pricing of European style contingent claims for the extended Fong and Vasicek model. This makes the pricing of zero coupon bonds reasonably straightforward. The pricing of corresponding
American instruments and exotics is also possible but more complex and is discussed in the next chapter.

4.1 Two-Factor Stochastic Volatility Models

Let us begin our analysis with a consideration of two-factor models. We assume that the short rate \( r_t \) and the instantaneous variance of the short rate \( \nu_t \), are given by the following Markovian system of stochastic differential equations

\[
\begin{align*}
    dr_t &= a_t \, dt + b_t \, dW^1_t, \\
    d\nu_t &= g_t \, dt + h_t \left( \rho \, dW^1_t + \sqrt{1 - \rho^2} \, dW^2_t \right),
\end{align*}
\]

for \( t_0 \leq t \leq T \) with initial values \( r_{t_0} \) and \( \nu_{t_0} \), where \( W = (W^1, W^2) \) is a two-dimensional Wiener process defined on the probability space \( (\Omega, \mathcal{F}, P) \). As in Section 1.1 we assume the filtration \( \mathcal{F} = (\mathcal{F})_{t \geq t_0} \) is the \( P \)-augmentation of the natural filtration of \( W \). The drift and diffusion coefficients for each component may depend on \( t, r_t \) and \( \nu_t \) however we have indicated only the time dependence in the formulation given above. We assume that the drift and diffusion coefficients satisfy appropriate growth and Lipschitz continuity conditions so that (1.1) admits a unique strong solution. The parameter \( \rho \in [-1, 1] \) measures the correlation between the noise terms of the short rate \( r_t \) and \( \nu_t \). As in the previous chapter we will not include the initial conditions when we write \( r_t \) and \( \nu_t \), \( t_0 \leq t \leq T \), as the solution of the two equations in (1.1).

Let \( B_t = B(t, T, r_t, \nu_t) \) be the price at time \( t \) of a zero coupon discount bond that pays one monetary unit at maturity \( T \). Assuming \( P \), the underlying probability measure, is the corresponding risk-neutral measure or more precisely the minimal equivalent martingale measure, see Hofmann, Platen & Schweizer (1992), we can use general valuation arguments to express this price as

\[
B_t = E \left( \exp \left\{ - \int_t^T r_s \, ds \right\} \bigg| \mathcal{F}_t \right),
\]

for \( t_0 \leq t \leq T \), where \( E \) denotes expectation with respect to the measure \( P \). This expression for \( B_t \) can also be obtained from the general contingent claim valuation formula (1.1.5) with \( \tau = T, \bar{P} = P \) and \( h(T, r_T, \nu_T) = 1 \).

For the extended Fong and Vasicek model we choose the functions appearing in (1.1) as follows:

\[
\begin{align*}
    a_t &= \alpha_t (r_t - r^*_t) + \lambda \nu_t, \\
    b_t &= \sqrt{\nu_t}, \\
    g_t &= \gamma (\bar{\nu} - \nu_t) - \xi \eta \nu_t, \\
    h_t &= \xi \sqrt{\nu_t},
\end{align*}
\]
for $t_0 \leq t \leq T$, where $\alpha_t$ and $\bar{r}_t$ are time-dependent values and $\lambda$, $\gamma$, $\xi$ and $\eta$ are non-negative constants. These parameters whose significance is outlined below are the same as those proposed by Fong & Vasicek (1991a,b) except we allow $\alpha_t$ and $\bar{r}_t$ to be functions of time. We consider time dependent parameters here because they allow greater freedom in fitting the model to actual market data. Hull & White (1990) also consider and use time dependent parameters in the drift term of the short rate for the Vasicek model.

With these specifications the equations in (1.1) can be rewritten in the form

$$
\begin{align*}
    dr_t &= (\alpha_t (\bar{r}_t - r_t) + \lambda v_t) \, dt + \sqrt{v_t} \, dW_1^t, \\
    dv_t &= (\gamma (\bar{v} - v_t) - \xi \eta v_t) \, dt + \xi \sqrt{v_t} \left( \rho \, dW_1^t + \sqrt{1 - \rho^2} \, dW_2^t \right),
\end{align*}
$$

(1.4)

for $t_0 \leq t \leq T$. Thus both the short rate $r_t$ and its instantaneous variance $v_t$ follow mean reverting processes. The short rate $r_t$ is pulled towards the time dependent level $\alpha_t \bar{r}_t + \lambda v_t$ with back-driving intensity $\alpha_t$. The instantaneous variance $v_t$ is attracted towards the fixed level $\frac{\gamma \bar{v}}{\gamma + \xi \eta}$ with intensity $\gamma + \xi \eta$.

For the extended Fong and Vasicek model the partial differential equation for the bond price $B_t$ becomes

$$
\begin{align*}
    \frac{\partial B_t}{\partial t} &+ (\alpha_t (\bar{r}_t - r_t) + \lambda v_t) \frac{\partial B_t}{\partial r} + (\gamma (\bar{v} - v_t) - \xi \eta v_t) \frac{\partial B_t}{\partial v} \\
 &+ \frac{1}{2} \frac{\partial^2 B_t}{\partial r^2} + \rho \xi v_t \frac{\partial^2 B_t}{\partial r \partial v} + \frac{1}{2} \xi^2 v_t \frac{\partial^2 B_t}{\partial v^2} - r_t B_t = 0
\end{align*}
$$

(1.5)

for $t_0 \leq t \leq T$ with boundary condition $B(T, T, r_T, v_T) = 1$. This is identical to the partial differential equation obtained by Fong & Vasicek (1991a,b) in the special case where $\alpha_t$ and $\bar{r}_t$ are constants.

As is shown by Fong & Vasicek (1991b), the solution of (1.5) has the form

$$
B(t, T, r_t, v_t) = \exp \{ -r_t D(t, T) + v_t F(t, T) + G(t, T) \}
$$

(1.6)

for $t_0 \leq t \leq T$, where $D$, $F$ and $G$ are time dependent real valued functions that satisfy the boundary conditions

$$
D(T, T) = F(T, T) = G(T, T) = 0.
$$

The functions $D$, $F$ and $G$ are the solutions of the ordinary differential equations

$$
\begin{align*}
    \frac{dD(t, T)}{dt} &= \alpha_t D(t, T) - 1, \\
    \frac{dF(t, T)}{dt} &= (\gamma + \xi \eta - \frac{1}{2} \xi^2 F(t, T) + \xi \rho D(t, T))F(t, T) \\
    &+ (\lambda - \frac{1}{2} D(t, T))D(t, T), \\
    \frac{dG(t, T)}{dt} &= \alpha_t \bar{r}_t D(t, T) - \gamma \bar{v} F(t, T),
\end{align*}
$$

(1.7)
for $t_0 \leq t \leq T$. We remark that Fong and Vasicek reverse the direction of time in their formulation of these equations. A similar expression to (1.6) holds for the short bond price $B^*_t = B^*(t, T^*, r_t, v_t)$ at time $t$, $t_0 \leq t \leq T^*$. In this case the corresponding functions $D(\cdot, T^*)$, $F(\cdot, T^*)$ and $G(\cdot, T^*)$ satisfy ordinary differential equations of the form (1.7) and the boundary conditions

$$D(T^*, T^*) = F(T^*, T^*) = G(T^*, T^*) = 0.$$  

The extended Fong and Vasicek model, characterized by the system of equations (1.4), includes terms relating to the market price of risk. For this model the market price of risk due to interest rate changes $q_t$, and volatility changes $p_t$, $t_0 \leq t \leq T$ are given by

$$q_t = \lambda \sqrt{v_t},$$

$$p_t = \eta \sqrt{v_t},$$

respectively. To transform the equations in (1.4) to the original ones for $r_t$ and $v_t$ proposed by Fong and Vasicek we use Girsanov’s Theorem to change the underlying probability measure $P$ to a measure $\tilde{P}$ so that the process $\tilde{W} = (\tilde{W}^1, \tilde{W}^2)$ defined by the components

$$\tilde{W}^1_t = W^1_t - \int_{t_0}^{t} d^1(s, r_s, v_s) ds,$$

$$\tilde{W}^2_t = W^2_t - \int_{t_0}^{t} d^2(s, r_s, v_s) ds,$$

for $t_0 \leq t \leq T$, is a two-dimensional Wiener process with respect to the new probability measure $\tilde{P}$. This measure is defined by the Radon-Nikodym derivative

$$\frac{d\tilde{P}}{dP} = \exp \left\{ - \int_{t_0}^{T} \frac{1}{2} \left( d^{(1)}(s, r_s, v_s)^2 + d^{(2)}(s, r_s, v_s)^2 \right) ds ight\},$$

$$+ \int_{t_0}^{T} d^{(1)}(s, r_s, v_s) dW^1_s + \int_{t_0}^{T} d^{(2)}(s, r_s, v_s) dW^2_s \right\},$$

(1.10)

where

$$d^{(1)}(t, r_t, v_t) = -q_t,$$

$$d^{(2)}(t, r_t, v_t) = \frac{\sqrt{r_t^2 + q_t^2}}{\sqrt{1 - \theta^2}}.$$  

(1.11)

Using equations (1.8), (1.9) and (1.11) we obtain from (1.4) the relations

$$dr_t = \alpha_t (\bar{r}_t - r_t) dt + \sqrt{r_t} d\tilde{W}^1_t,$$

$$dv_t = \gamma (\bar{v}_t - v_t) dt + \xi \sqrt{v_t} \left( \rho d\tilde{W}^1_t + \sqrt{1 - \theta^2} d\tilde{W}^2_t \right).$$

(1.12)
4.2. PRICING OF CONTINGENT CLAIMS

for $t_0 \leq t \leq T$, which are the Fong and Vasicek equations for $r_t$ and $v_t$, respectively.

Expanding (1.6) using the Ito rule, Feynman-Kac formula (1.5), and (1.1), we obtain

$$\frac{dB_t}{B_t} = r_t dt + \sqrt{v_t} \frac{1}{B_t} \frac{\partial B_t}{\partial r} dW^1_t + \xi \sqrt{v_t} \frac{1}{B_t} \frac{\partial B_t}{\partial v} \left( \phi dW^1_t + \sqrt{1 - \phi^2} dW^2_t \right)$$

(1.13)

so that from (1.8) and (1.9) the stochastic differential equation for $B_t$, using the Wiener processes $\tilde{W}^1$ and $\tilde{W}^2$ becomes

$$\frac{dB_t}{B_t} = \left( r_t - \lambda v_t \frac{1}{B_t} \frac{\partial B_t}{\partial r} + \eta \xi v_t \frac{1}{B_t} \frac{\partial B_t}{\partial v} \right) dt$$

$$+ \sqrt{v_t} \frac{1}{B_t} \frac{\partial B_t}{\partial r} d\tilde{W}^1_t + \xi \sqrt{v_t} \frac{1}{B_t} \frac{\partial B_t}{\partial v} \left( \phi d\tilde{W}^1_t + \sqrt{1 - \phi^2} d\tilde{W}^2_t \right),$$

(1.14)

where $\tilde{W}^1$ and $\tilde{W}^2$ are independent Wiener processes under the transformed measure $\tilde{P}$. This measure can be interpreted as the objective ‘real world’ probability measure. Thus under the measure $\tilde{P}$ the drift component for $B_t$ is $r_t B_t - \lambda v_t \frac{\partial B_t}{\partial r} + \eta \xi v_t \frac{\partial B_t}{\partial v}$.

4.2 Pricing of Contingent Claims

We consider now the problem of pricing a general contingent claim on a discount bond $B_t$ which matures at time $T$. Our general contingent claim is characterized by a payoff structure of the form $G = h(B_{T^*})$, where $T^*$ is the expiration time for the claim, $t_0 < T^* < T$. An example would be a European call option with strike $K$ having a payoff structure of the form

$$G = h(B_{T^*}) = (B_{T^*} - K)^+ = \max(B_{T^*} - K, 0).$$

In the remaining part of this section we will synthesize a complete market by using two bonds maturing at different times as hedging instruments and via this mechanism evaluate the price of the contingent claim $H$ at time $t_0$, $0 \leq t_0 \leq T^*$. This analysis will be undertaken under the corresponding risk neutral pricing measure $P$. In the case of an incomplete market, for instance if only one bond can be used as a hedging instrument, we could choose as a suitable pricing measure the minimal equivalent martingale measure described in Section 3.2 and Hofmann, Platen & Schweizer (1992). This measure is the risk-neutral measure when the market is complete.

We assume that the stochastic differential equations, which describe the behaviour of the short rate $r_t$ and its instantaneous variance $v_t$, are of the general form given by (1.1) with the bond price $B_t$ given by equation (1.2). The expectation in equation (1.2) is taken with respect to the minimal equivalent martingale measure $P$.

Let $B^*_t = B(t, T^*, r_t, v_t)$ and $B_t = B(t, T, r_t, v_t)$ be the prices at time $t$, $t_0 \leq t \leq T$, of discount bonds maturing at times $T^*$ and $T$, respectively. We denote by $\beta_t$ the value of a savings account which accumulates interest continuously and evolves according to
the linear growth equation (1.1.2) with solution (1.1.3) for \( t_0 \leq t \leq T \). The process \( \beta \) can be interpreted as a portfolio of extremely short bonds that are continuously rolled over from one (short) period to the next.

The price \( u_t \) at time \( t, \ t_0 \leq t \leq T \), of a contingent claim with payoff structure \( G = h(B_{T^-}) \), is obtained from the time-independent versions of (1.1.5) and (1.1.11) with \( \tau = T^* \) and can be written in the form

\[
 u_t = u(t, \beta_t, r_t, v_t) = E \left( \exp \left\{ - \int_t^{T^*} r_s ds \right\} h(B_{T^*}) \left| \mathcal{F}_t \right. \right) = \beta_t E \left( \frac{h(B_{T^*})}{\beta_{T^*}} \left| \mathcal{F}_t \right. \right),
\]

for \( t_0 \leq t \leq T^* \), where, as is the case for the discount bond price \( B_t \) given by (1.2), the conditional expectation is taken under the minimal equivalent martingale measure \( P \).

We will now assume that the bond price formulas \( B_t = B(t, T, r_t, v_t) \) and \( B_t = B^*(t, T^*, r_t, v_t) \) can be solved for both \( r_t \) and \( v_t \) in terms of the state variables \( B_t \) and \( B^*_t, \ t_0 \leq t \leq T^* \). That is we can express \( r_t \) and \( v_t \) in the form \( r_t = r(t, B_t, B^*_t) \) and \( v_t = v(t, B_t, B^*_t) \). It follows that the price \( u_t \) can be expressed as a function depending on \( t, \beta_t, B_t, B^*_t \) rather than \( t, \beta_t, r_t, v_t \). These expressions for \( r_t \) and \( v_t \) can be obtained in a straightforward manner for the extended Fong and Vasicek model as will be shown at the end of Section 4.4.

Using these assumptions we will construct a dynamical hedging portfolio, comprised of units in the \( T \) maturing bond \( B \), the \( T^* \) maturing bond \( B^* \) and the savings account \( \beta \), that allows one to fully replicate the contingent claim \( G \).

Introducing the discounted values

\[
 \tilde{u}_t = \frac{u_t}{\beta_t}, \quad \tilde{B}_t = \frac{B_t}{\beta_t}, \quad \tilde{B}^*_t = \frac{B^*_t}{\beta_t}
\]

for \( t_0 \leq t \leq T^* \), we choose

\[
 \eta_t = \frac{\partial}{\partial B} \tilde{u}_t, \quad \eta^*_t = \frac{\partial}{\partial B^*} \tilde{u}_t,
\]

as the number of units to be held in the \( T \) maturing bond \( B_t \),

\[
 \zeta_t = \tilde{u}_t - \eta_t \tilde{B}_t - \eta^*_t \tilde{B}^*_t,
\]

as the number of units to be held in the \( T^* \) maturing bond \( B^*_t \) and

\[
 \eta_t = \frac{\partial}{\partial B} \tilde{u}_t = \frac{\partial}{\partial B} u_t.
\]
This means that the value of the corresponding portfolio at time \( t, \ t_0 \leq t \leq T^* \), satisfies the relation
\[
u_t = \eta_t B_t + \eta_t^* B_t^* + \zeta t \beta_t.
\]
Also, from (2.1) we see that
\[
u_{T^*} = h(B_{T^*}). \tag{2.6}
\]
Consequently, in the case where we use two hedging instruments and a savings account, corresponding to a complete market situation, the contingent claim is fully replicated and \( u_t \) as given by (2.1) represents the price of the contingent claim at time \( t_0 \leq t \leq T^* \).

Since \( B_{T^*}^* = 1 \), for a European call option with strike \( K \) the option pricing formula given by (2.1) takes the form
\[
u_t = \beta t E \left( \left( \frac{B_{T^*}}{B_{T^*}^*} \right)^+ \mid \mathcal{F}_t \right), \tag{2.7}
\]
for \( t_0 \leq t \leq T^* \).

If we consider a general multifactor model with \( m \) driving Wiener processes we would then need to include \( m \) risky assets in our hedging portfolio in order to synthesize a complete market and therefore to fully replicate our contingent claim.

Unfortunately, the valuation of the conditional expectation given by (2.7) is rather involved because we have to handle concurrently both the evolution of the savings account process \( \beta \) together with the discounted long bond process \( B \). To simplify this problem one can apply a separation principle which enables this valuation formula to be expressed in a more computationally convenient form. This principle which uses the so called forward measure is described by El Karoui & Rochet (1989) and Jamshidian (1989a) and has also been discussed by Brace & Musiela (1994) and Goldman et al. (1995a).

More precisely, it holds under rather general conditions that the price \( u_t \), given by equation (2.1), of the contingent claim can be written in the form
\[
u_t = B_t^* \tilde{E} \left( h(B_{T^*}) \mid \mathcal{F}_t \right), \tag{2.8}
\]
where \( \tilde{E} \) denotes expectation under a new probability measure \( \tilde{P} \), the forward measure. This probability measure is defined by the Radon-Nikodym derivative
\[
\frac{d\tilde{P}}{dP} = \exp \left\{ -\int_{t_0}^{T} \frac{1}{2} \left[ (\theta_{1,s}^*)^2 + (\theta_{2,s}^*)^2 \right] ds \right\}
+ \int_{t_0}^{T} \theta_{1,s}^* dW_s^1 + \int_{t_0}^{T} \theta_{2,s}^* dW_s^2
\]
\[
\tag{2.9}
\]
with
\[
\theta_{1,s}^* = \frac{b_t}{B_t^*} \frac{\partial B_t^*}{\partial r} + \varphi h_t \frac{1}{B_t^*} \frac{\partial B_t^*}{\partial v},
\]
\[
\theta_{2,s}^* = \sqrt{1 - \varphi^2} h_t \frac{1}{B_t^*} \frac{\partial B_t^*}{\partial v},
\]
for \( t_0 \leq t \leq T \). The functions \( \theta^*_1 \) and \( \theta^*_2 \) in general depend on \( t, r_t \) and \( v_t \). Using Girsanov’s Theorem we know that the process \( \tilde{W} = (\tilde{W}^1, \tilde{W}^2) \) with components given by

\[
\tilde{W}^1_t = W^1_t - \int_{t_0}^t \theta^*_{1,s} \, ds \\
\tilde{W}^2_t = W^2_t - \int_{t_0}^t \theta^*_{2,s} \, ds
\]

for \( t_0 \leq t \leq T^* \) is a two-dimensional Wiener process under the measure \( \tilde{P} \).

If we define the processes \( \theta_1 \) and \( \theta_2 \) given by

\[
\theta_{1,t} = b_t \frac{1}{B_t} \frac{\partial B_t}{\partial r} + \phi h_t \frac{1}{B_t} \frac{\partial B_t}{\partial \nu}
\]

and

\[
\theta_{2,t} = \sqrt{1 - \phi^2} h_t \frac{1}{B_t} \frac{\partial B_t}{\partial \nu}
\]

for \( t_0 \leq t \leq T \), then applying Ito’s formula to (1.6) and the corresponding representation for \( B^*_t \) using (1.1) we obtain

\[
dB_t = r_t B_t \, dt + \theta_{1,t} B_t \, dW^1_t + \theta_{2,t} B_t \, dW^2_t \\
dB^*_t = r_t B^*_t \, dt + \theta^*_{1,t} B^*_t \, dW^1_t + \theta^*_{2,t} B^*_t \, dW^2_t \quad (2.10)
\]

for \( t_0 \leq t \leq T^* \).

By using the Wiener process \( \tilde{W} = (\tilde{W}^1, \tilde{W}^2) \) rather than \( W = (W^1, W^2) \) the stochastic differential equations for \( r_t \) and \( v_t \) given by (1.1) become

\[
dr_t = \left( a_t + b_t \theta^*_{1,t} \right) \, dt + b_t \, d\tilde{W}^1_t, \quad (2.11) \\
dv_t = \left( g_t + h_t \left( \phi \theta^*_{1,t} + \sqrt{1 - \phi^2} \theta^*_{2,t} \right) \right) \, dt + h_t \left( \phi \, d\tilde{W}^1_t + \sqrt{1 - \phi^2} \, d\tilde{W}^2_t \right).
\]

Now let us define the \( B^*_t \)-discounted process \( X \) by

\[
X_t = \frac{B_t}{B^*_t} \quad (2.12)
\]

for \( t_0 \leq t \leq T^* \). We note that \( X_{T^*} = B_{T^*} \) since \( B^*_T = 1 \) and consequently the payoff structure \( h(B_{T^*}) \) used in (2.8) can be replaced by \( h(X_{T^*}) \). By applying Ito’s formula to the quotient \( X = \frac{B}{B^*} \) together with (2.10) we see that

\[
dx_t = \left( X_t \theta^*_{1} \left( \theta^*_{1} - \theta_1 \right) + X_t \theta^*_{2} \left( \theta^*_{2} - \theta_2 \right) \right) \, dt \\
+ X_t \left( \theta_{1,t} - \theta^*_{1,t} \right) \, dW^1_t + X_t \left( \theta_{2,t} - \theta^*_{2,t} \right) \, dW^2_t. \quad (2.13)
\]

Replacing the Wiener process \( W = (W^1, W^2) \) by \( \tilde{W} = (\tilde{W}^1, \tilde{W}^2) \) this stochastic differential equation becomes

\[
dx_t = X_t (\theta_{1,t} - \theta^*_{1,t}) \, d\tilde{W}^1_t + X_t (\theta_{2,t} - \theta^*_{2,t}) \, d\tilde{W}^2_t \quad (2.14)
\]
for $t_0 \leq t \leq T^*$. From (2.14) we can conclude that the discounted process $X$ is a martingale under the measure $\tilde{P}$. This observation can be used to simplify considerably the stochastic analytic and numerical methods which are needed to solve the system of stochastic differential equations (2.11) and (2.14) for $r_t$, $v_t$ and $X_t$, respectively, and which are used in the valuation formula (2.8) with $X_{T^*}$ replacing $B_{T^*}$. For the extended Fong and Vasicek model the stochastic differential equations for $v_t$ and $X_t$ are independent of $r_t$, a result that will be verified in Section 4.4. This means that computation of option prices from (2.8) can be obtained from the evolution of the processes $X$ and $v$ only rather than $X$, $v$ and $r$.

4.3 Stochastic Numerical and Variance Reduction Methods

In this section we will consider the problem of finding efficient, stochastic numerical and variance reduction procedures for approximating option prices using the valuation formula (2.8) based on solutions to the stochastic differential equations (2.11) and (2.14). We require numerical procedures to approximate the solution of (2.8) because in general there is no explicit solution for this formula. The methods developed here also allow for computation of the hedge ratios (2.3)-(2.5), either from finite-difference estimates or the exact representation results developed in Chapter 1. It should be noted that in principle these numerical approximations can be applied to other pricing measures besides the measure $\tilde{P}$ used in (2.8).

Using the formula (2.8) and the measure $\tilde{P}$, rather than (2.2) and the measure $P$, is computationally more convenient for two reasons. Firstly, it allows the option price $u_t$ to be expressed as the product of simpler components namely a discount bond price $B_t$ and a conditional expectation $\tilde{E}(h(B_{T^*}) \mid \mathcal{F}_t) = \tilde{E}(h(X_{T^*}) \mid \mathcal{F}_t)$. Secondly, this conditional expectation can often be computed using non-stochastic methods for the case where $v_t$ is deterministic and $r_t$ evolves according to the first equation in (2.11). As we will see shortly, this provides a mechanism for the construction of control variates which can be used to approximate option prices for the extended Fong and Vasicek model.

Here we apply Monte Carlo estimation techniques using the payoff structure $h(X_{T^*}) = h(B_{T^*})$, or related estimators obtained from control variate formulations of the type described in Section 2.5. Because the random variable $X_{T^*}$ given by (2.13) cannot be determined analytically we require an approximation to it which is obtained from an appropriate discrete time numerical procedure. As has been explained in Chapter 3, for option price calculations, where estimates of the expectation of functionals of Itô diffusion processes are needed, we use the class of numerical schemes called weak approximations.

For the simulation results described in Section 4.5 we used a class of stochastic approximations called weak Euler predictor-corrector methods. These methods exhibit
good convergence and stability properties, the latter following from their close association with corresponding implicit schemes.

The choice of what class of stochastic approximations is most appropriate is usually based on a consideration of the dynamics of the underlying system of stochastic differential equations. A class of methods which works for one application may need to be replaced by a new class for a different application. It should be emphasized that the stability properties of weak numerical schemes are much more delicate and subtle than those of corresponding deterministic schemes for ordinary differential equations. In fact when using stochastic numerical methods the situation can easily arise whereby even for very small step sizes a scheme becomes unstable. These difficulties can often be avoided by a thorough and systematic investigation of numerical stability.

Additional information on predictor-corrector methods is given by Kloeden & Platen (1992). In principle these schemes are derived from implicit Euler methods, however all implicit methods require in general at each time step the solution to an algebraic equation or system of equations. This overhead is avoided by use of a predictor-corrector method, where the approximate value at the right hand side of the scheme is substituted by a predicted value computed by a straightforward Euler approximation.

For the vector valued process $Z$ given by (3.3.1) and the equi-spaced time discretization $(t)$ of the interval $[t_0, T^\star]$ given by (2.3.1) with step size $\Delta = (T^\star - t_0)/N$ these schemes have the form, see Kloeden & Platen (1992),

$$
Y_n^\Delta = Y_n^\Delta + \left\{ \alpha \tilde{a}_n \left( \tau_{n+1}, \tilde{Y}_{n+1} \right) + (1 - \alpha) \bar{a}_n \left( \tau_n, Y_n^\Delta \right) \right\} \Delta 
+ \sum_{j=1}^m \left\{ \eta b^j \left( \tau_{n+1}, \tilde{Y}_{n+1} \right) + (1 - \eta) b^j \left( \tau_n, Y_n^\Delta \right) \right\} \Delta \tilde{W}_n^j
$$

(3.1)

where $\bar{a}_n$ is the modified drift vector given by

$$
\bar{a}_n = a - \eta \sum_{j=1}^m \sum_{k=1}^d b^j_k \frac{\partial b^j}{\partial z^k}
$$

and $\tilde{Y}_{n+1}$ is the predictor

$$
\tilde{Y}_{n+1} = Y_n^\Delta + a \left( \tau_n, Y_n^\Delta \right) \Delta + \sum_{j=1}^m b^j \left( \tau_n, Y_n^\Delta \right) \Delta \tilde{W}_n^j,
$$

for $n = 0, 1, \ldots, N - 1$, $\alpha, \eta \in [0, 1]$ and initial value $Y_0^\Delta = Z_{t_0}$. Here the increments $\Delta \tilde{W}_n^j$, $j = 1, \ldots, m$, can be chosen as independent Gaussian $N(0, \Delta)$ distributed random variables or as two-point distributed random variables with probabilities determined from (2.3.5). Note that (3.1) is expressed using vector notation for the process $Y^\Delta = (Y^\Delta,1, \ldots, Y^\Delta,d)$.

We remark that the parameters $\alpha$ and $\eta$ can be adjusted to give different degrees of implicitness in the drift and diffusion terms, respectively.
4.3. VARIANCE REDUCTION METHODS

As mentioned previously, in the case of the extended Fong and Vasicek model the components $X_t$ and $v_t$ evolve independently of $r_t$. It is therefore sufficient to approximate the functional in (2.8) by using only the components $X_t$ and $v_t$, $t_0 \leq t \leq T^*$. Consequently we use the value $m = 2$ in the scheme (3.1) for approximations of the diffusion process $(X, v)$.

Let us now consider the problem of using Monte Carlo simulation to approximate the option price $u_{t_0}$, at time $t_0$, given by equation (2.8). If we replace the random variable $X_t$ by the discrete time weak approximation $\tilde{Y}_{N^1}$, then an estimate $u_{t_0}^\Delta$ of this option price is given by

$$u_{t_0}^\Delta = B_{t_0}^* \tilde{E} \left( h \left( Y_{N^1}^\Delta \right) \bigg| Y_0^\Delta = (X_{t_0}, v_{t_0}) \right),$$

(3.2)

where $X_{t_0}$ and $v_{t_0}$ are the initial values for $X$ and $v$ at time $t_0$ and $Y^\Delta$ is an approximation for the component $X$. Note that $X_{t_0} = B_{t_0}/B_{t_0}^*$ can be computed from (1.6) and the corresponding formula for $B_{t_0}^*$ given by (1.6).

A Monte Carlo estimation of (3.2) would involve the generation of say $M$ outcomes $Y_N(\omega_i), i \in \{1, \ldots, M\}$ and computing the sample mean

$$u_{t_0}^{\Delta M} = B_{t_0}^* \frac{1}{M} \sum_{i=1}^{M} h \left( Y_{N^1}^\Delta(\omega_i) \right),$$

(3.3)

where each $Y_{N^1}^\Delta(\omega_i), i = 1, \ldots, M$, represents an independent realization of $Y_{N^1}^\Delta$ at time $t_N = T^*$ starting at time $t_0$. This formula is similar to that obtained for the computation of barrier options under stochastic volatility given by (3.3.5).

As explained in Chapter 3 we can ensure that $u_{t_0}^\Delta$ is close to $u_{t_0}$ by the use of appropriate numerically stable higher order schemes, however the closeness of $u_{t_0}^{\Delta M}$ to $u_{t_0}^\Delta$ depends ultimately on the variance of $h(X_t)$. Reducing this variance usually requires the application of a range of variance reduction techniques.

We will now describe a variance reduction technique based on the use of control variates, of the type described in Section 2.5, which was found to be effective in the simulation experiments outlined in Section 4.5.

Let $Z = (X, v)$ be the vector valued stochastic process which satisfies the first equation in (2.11) and (2.14) and $\hat{Z} = (\hat{X}, \hat{v})$ be another vector valued process which is close to $Z$ but which is more analytically tractable. For example, $\hat{Z}$ could be obtained from the extended Vasicek model considered by Hull & White (1990). Some additional information on how the process $\hat{Z}$ can be formulated is given in the next section. As explained in Section 2.4, the idea behind this technique is to simulate only the difference between $Z_t$ and $\hat{Z}_t$ and combine this with the more easily computed contingent claim price corresponding to $\hat{Z}_t$.

Let $u_{t_0}$ be the option price given by (2.8) and $\hat{u}_{t_0}$ the option price corresponding to the process $\hat{Z}$. That is

$$\hat{u}_{t_0} = B_{t_0}^* \hat{E} \left( h \left( \hat{X}_{T^*} \right) \bigg| \mathcal{F}_{t_0} \right),$$

(3.4)
for $t_0 \leq t \leq T^*$ with $\hat{X}_{t_0} = X_{t_0}$, and $\hat{v}_{t_0} = v_{t_0}$.

Consider the random variable
\[
\hat{Z}_T = B_{t_0}^* \left( h(\hat{X}_{T^*}) - \alpha \left( h(\hat{X}_{T^*}) - \hat{E} \left( h(\hat{X}_{T^*}) \right) \right) \right)
= B_{t_0}^* \left( h(X_{T^*}) - \alpha \left( h(X_{T^*}) - \hat{u}_{t_0} / B_{t_0}^* \right) \right)
\]
for $\alpha \in \mathbb{R}$. This estimator is obtained from equation (2.5.3). Since
\[
\hat{E} \left( \hat{Z}_T \right) = \hat{E} (B_{t_0}^* h(X_{T^*}))
= B_{t_0}^* \hat{E} (h(X_{T^*}))
\]
with initial conditions $\hat{X}_{t_0} = X_{t_0} = B_{t_0} / B_{t_0}^*$, $\hat{v}_{t_0} = v_{t_0}$, we see that $\hat{Z}_T$ is an unbiased estimator for $u_{t_0}$. The variance of $\hat{Z}_T$ under $\hat{P}$ denoted by $\text{Var}(\hat{Z}_T)$ can be calculated from (2.4.14). If $\hat{Z}$ is close to $Z$ this variance will be smaller than that obtained from the estimator $B_{t_0}^* h(X_{T^*})$. A discrete time version of this estimator denoted by $\hat{Z}_N^\Delta$ would be
\[
\hat{Z}_N^\Delta = B_{t_0}^* \left( h \left( Y_{N,1}^\Delta \right) - \alpha \left( h \left( Y_{T^*,1}^\Delta \right) - \hat{u}_{t_0} / B_{t_0}^* \right) \right),
\]
where $Y^\Delta$ and $\hat{Y}^\Delta$ are discrete time approximations corresponding to the processes $Z$ and $\hat{Z}$ respectively, using the scheme (3.1).

A Monte Carlo simulation of the option price $u_{t_0}$ using the estimator $\hat{Z}_N^\Delta$ would be conducted in a similar fashion to that for the estimator used in (3.2) and whose sample mean is calculated in (3.3). As explained in Section 2.5 the optimal value of $\alpha$ to minimize the variance of $\hat{Z}_N^\Delta$ can be computed as the simulation proceeds.

Another variance reduction technique, based on the martingale control variates proposed by Clewlow & Carverhill (1992, 1994), and which is also described and extended in Section 2.5 can be formulated as follows. We construct an estimator of the form
\[
\hat{Z}_N^\Delta = B_{t_0}^* \left( h \left( Y_{N,1}^\Delta \right) \right)
- \alpha \left( \sum_{k=1}^{N-1} \frac{\partial \hat{u}}{\partial \hat{x}} \left( t_k, Y_{k,1}^\Delta, Y_{k,2}^\Delta \right) \left[ \Delta Y_{k,1}^\Delta - \hat{E} \left( Y_{k,1}^\Delta \right) \right] \right)
- \beta \left( \sum_{k=1}^{N-1} \frac{\partial \hat{u}}{\partial \hat{v}} \left( t_k, Y_{k,1}^\Delta, Y_{k,2}^\Delta \right) \left[ \Delta Y_{k,2}^\Delta - \hat{E} \left( Y_{k,2}^\Delta \right) \right] \right)
\]
for $\alpha, \beta \in \mathbb{R}$, where $\Delta Y_{k,1}^\Delta = Y_{k+1,1}^\Delta - Y_{k,1}^\Delta$, $\Delta Y_{k,2}^\Delta = Y_{k+1,2}^\Delta - Y_{k,2}^\Delta$ and $Y_{1,1}^\Delta$ and $Y_{1,2}^\Delta$ are approximations for the components $X$ and $v$, respectively.

Here we take $\hat{u}_t = \hat{u}(t, \hat{X}_t, \hat{v}_t)$, $t_0 \leq t \leq T^*$ to be the option price at time $t$ for the process $\hat{Z}$. We assume $\hat{u}_t$, $t_0 \leq t \leq T^*$ can be computed analytically so that the partial derivatives $\frac{\partial \hat{u}}{\partial \hat{x}}$ and $\frac{\partial \hat{u}}{\partial \hat{v}}$ can also be efficiently evaluated. Note that for the numerical scheme (3.1) the expectations $\hat{E}(\Delta Y_{k,1}^\Delta)$ and $\hat{E}(\Delta Y_{k,2}^\Delta)$, $k \in \{1, \ldots, N - 1\}$, can be
4.4. EXTENDED FONG AND VASICEK MODEL

computed in a straightforward manner. Also, from (3.7) we see that this estimator is
unbiased and does not require outcomes to be generated for the approximation \( \hat{Y}^\Delta \)
corresponding to the process \( \hat{Z} = (\hat{X}, \hat{\nu}) \). As is the case for the estimator \( \hat{Z}_N^\Delta \), simulation
of the random variable \( \hat{Z}_N^\Delta \) can be used both to approximate the option price \( u_{t_0} \) and
simultaneously the optimal coefficient vector \( (\alpha, \beta) \) to minimize the variance of \( \hat{Z}_N^\Delta \).

These control variate methods can be conveniently combined with antithetic and
stratified sampling techniques of the type described in Section 2.5.

4.4 Application to the Extended Fong
and Vasicek Model

In this section we will consider the application of the stochastic numerical and variance
reduction procedures described in the previous section specifically with reference to the
extended Fong and Vasicek model, with drift and diffusion coefficients given by (1.3).
We will show how these techniques can be used to compute the prices of discount bonds
and European call options on these bonds.

Firstly, we will consider the task of efficiently computing the discount bond price
\( B_t, \; t_0 \leq t \leq T \) using (1.6) and the ordinary differential equations (1.7). We start
with the observation that the system of equations (1.7) that define the evolution of
the functions \( D, F \) and \( G \) fits within the framework given by the system of stochastic
differential equations (1.1) but with only deterministic coefficients. This means the
Euler predictor-corrector methods described in (3.1) can also be used to compute the
discount bond price \( B_t, \; t_0 \leq t \leq T \). If we set \( \alpha = \frac{1}{2} \) and remove the stochastic
components in the scheme (3.1), it becomes the well-known improved Euler or Heun
deterministic method, see Kloeden & Platen (1992). This method, applied backwards
in time starting from time \( T \), can be used to compute the functions \( D, F \) and \( G \) and
from these, see (1.6), the discount bond price \( B_t \) at time \( t, \; t_0 \leq t \leq T \). A similar
approximation procedure can be used for the short bond \( B^* \).

Let us now consider the structure of the system of stochastic differential equations
given by (2.11) and (2.14) for the extended Fong and Vasicek model with components
\( X \) and \( \nu \) and coefficient functions \( a_t, b_t, g_t \), and \( h_t \) given by (1.3). Simulated estimates
of the diffusion processes \( \nu \) and \( X \) involve evaluations of the processes \( \theta_1, \theta_1^*, \theta_2 \) and \( \theta_2^* \).
These processes require computation of the partial derivatives \( \frac{\partial B}{\partial r}, \frac{\partial B}{\partial v}, \frac{\partial B^*}{\partial r} \) and \( \frac{\partial B^*}{\partial v} \)
which can be evaluated explicitly for the extended Fong and Vasicek model by using
the discount bond price (1.6) for \( B_t \) and \( B^*_t \), respectively. Computing the expressions
for \( \frac{\partial B}{\partial r} \) and \( \frac{\partial B}{\partial v} \) we see that
\[
\frac{\partial B_t}{\partial r} = -D(t, T) B_t, \]
\[
\frac{\partial B_t}{\partial v} = F(t, T) B_t, \]
\[
\frac{\partial B_t^*}{\partial r} = -D(t,T^*) B_t^*, \quad (4.1)
\]
\[
\frac{\partial B_t^*}{\partial \theta} = F(t,T^*) B_t^*.
\]
Substituting these results in the relations for \( \theta_1, \theta_2, \theta_1^* \) and \( \theta_2^* \) used in (2.11) and (2.14) we have from (1.3) the identities
\[
\begin{align*}
\theta_{1,t} &= \sqrt{v_t} (\varrho \xi F(t,T) - D(t,T)), \\
\theta_{2,t} &= \sqrt{v_t} \sqrt{1 - \varrho^2 \xi F(t,T)}, \\
\theta_{1,t}^* &= \sqrt{v_t} (\varrho \xi F(t,T^*) - D(t,T^*)), \\
\theta_{2,t}^* &= \sqrt{v_t} \sqrt{1 - \varrho^2 \xi F(t,T^*)}.
\end{align*}
\]
From these equations it can be seen that the \( X_t \) component for the extended Fong and Vasicek model, given by (2.14), does not depend on the short rate \( r_t \). Consequently the \( r_t \) component can be ignored when we attempt to compute \( X_{T^*} \)-dependent contingent claims including European call options of the form (2.8). Of course the parameters which define the short rate process \( r \) still affect the functions \( D \) and \( F \) and naturally the bond prices \( B_t \) and \( B_t^* \). The evolution of the short rate therefore influences the initial value \( X_{t_0} = B_{t_0} / B_{t_0}^* \).

As was pointed out in the previous section the use of appropriate variance reduction techniques is crucial if we are to obtain fast and accurate valuations of option prices. General methods such as antithetic variates and stratified sampling can be incorporated in any approximation scheme of the form 3.1, however choosing a control variate of the type described in the previous section will depend on the structure of the underlying system of stochastic differential equations being modelled together with the parameter ranges being used.

Estimates of the option price via the control variate formulation (3.6) or (3.7) require the specific form of the stochastic differential equations for the state variables \( X_t \) and \( v_t \) to be computed. Using (4.2), (2.11) and (2.14) these stochastic differential equations become
\[
\begin{align*}
dX_t &= X_t \sqrt{v_t} \left[ \varrho \xi (F(t,T) - F(t,T^*)) + D(t,T) - D(t,T^*) \right] d\tilde{W}_t^1 \\
&\quad + X_t \sqrt{v_t} \sqrt{1 - \varrho^2 \xi (F(t,T) - F(t,T^*))} d\tilde{W}_t^2 \\
dv_t &= \left( \gamma \varrho + \nu_t \left[ - (\gamma + \xi \eta) + \xi^2 F(t,T^*) - \xi \varrho D(t,T^*) \right] \right) dt \\
&\quad + \xi \sqrt{v_t} \left( \varrho d\tilde{W}_t^1 + \sqrt{1 - \varrho^2} d\tilde{W}_t^2 \right). \quad (4.3)
\end{align*}
\]
Applying the Ito formula to the second of the above equations using the change of
variable $\sigma_t = \sqrt{\bar{v}_t}$ we obtain

$$
\begin{align*}
\text{d}\sigma_t &= \frac{1}{2} \left[ \left( \gamma \bar{v}_t - \frac{1}{4} \xi^2 \right) \sigma_t + \sigma_t \left( -\left( \gamma + \xi \eta \right) + \xi^2 F(t, T^*) - \xi \varrho D(t, T^*) \right) \right] \text{d}t \\
&\quad + \frac{1}{2} \xi \left( \varrho \text{d}\tilde{W}_t^1 + \sqrt{1-\varrho} \text{d}\tilde{W}_t^2 \right).
\end{align*}
$$

(4.4)

Thus if we use the instantaneous standard deviation $\sigma_t$ of the short rate $r_t$ instead of $\bar{v}_t$, then we obtain a stochastic differential equation for $\sigma_t$ with only an additive noise component. The stochastic differential equations for $(X, \sigma)$ suggest a natural choice for a control variate process $(\tilde{X}, \tilde{\sigma})$ for which option prices can be more easily computed. We take the stochastic differential equations for $(\tilde{X}, \tilde{\sigma})$ to be the same as for $(X, \sigma)$ except we set the noise term in the $\sigma$ component to be zero. These stochastic differential equations therefore become

$$
\begin{align*}
\text{d}\tilde{X}_t &= \tilde{X}_t \tilde{\sigma}_t \left[ \varrho \xi \left( F(t, T) - F(t, T^*) \right) + D(t, T) - D(t, T^*) \right] \text{d}\tilde{W}_t^1 \\
&\quad + \tilde{X}_t \tilde{\sigma}_t \sqrt{1-\varrho^2} \xi \left( F(t, T) - F(t, T^*) \right) \text{d}\tilde{W}_t^2 \\
\text{d}\tilde{\sigma}_t &= \frac{1}{2} \left[ \left( \gamma \bar{v}_t - \frac{1}{4} \xi^2 \right) \tilde{\sigma}_t + \tilde{\sigma}_t \left( -\left( \gamma + \xi \eta \right) + \xi^2 F(t, T^*) - \xi \varrho D(t, T^*) \right) \right] \text{d}t,
\end{align*}
$$

(4.5)

for $t_0 \leq t \leq T^*$ with initial value $(\tilde{X}_{t_0}, \tilde{\sigma}_{t_0})$.

Although a closed-form solution for $\tilde{\sigma}_t$ cannot be easily obtained, because of the presence of the time-dependent functions $F(\cdot, T^*)$ and $D(\cdot, T^*)$, a close approximation to the solution can be easily computed using deterministic numerical methods. With this formulation the state variable $\tilde{\sigma}$ can therefore be viewed as a non-stochastic time-dependent function. The diffusion coefficient for $\tilde{X}$ now appears in the form of a product of the state variable $\tilde{X}_t$ times a time-dependent function. Option prices for $(\tilde{X}, \tilde{\sigma})$ can therefore be evaluated using techniques which apply for one-factor models with multiplicative noise and time-dependent diffusion coefficients.

Other choices for the control variate are also possible. For example we could take $(\tilde{X}, \tilde{\varrho})$ to be the control variate process which is the same as $(X, \varrho)$ except we set the parameter $\xi$ equal to zero in all coefficients that include time-dependent parameters. Using this control variate the stochastic differential equations for $\tilde{X}_t$ and $\tilde{\varrho}_t$ given by (4.3) become

$$
\begin{align*}
\text{d}\tilde{X}_t &= \tilde{X}_t \sqrt{\tilde{\varrho}_t} \left( D(t, T^*) - D(t, T) \right) \text{d}\tilde{W}_t^1, \\
\text{d}\tilde{\varrho}_t &= \gamma \bar{v}_t - \left( \gamma + \xi \eta \right) \tilde{\varrho}_t \text{d}t
\end{align*}
$$

(4.6)

for $t_0 \leq t \leq T^*$ with initial value $(\tilde{X}_{t_0}, \tilde{\varrho}_{t_0})$.

With this formulation the $\tilde{\varrho}_t$ component can be solved explicitly on the interval $[t_0, T^*]$ yielding

$$
\tilde{\varrho}_t = e^{-(\gamma + \xi \eta)(t-t_0)} \left( \tilde{\varrho}_{t_0} - \frac{\gamma \bar{v}_t}{\gamma + \xi \eta} \left( 1 - e^{-(\gamma + \xi \eta)(t-t_0)} \right) \right)
$$

(4.7)
for \( t_0 \leq t \leq T^* \). This equation can be compared with (3.3.7) and (5.3.2).

The equation for \( \hat{X} \), given in (4.6), can also be solved explicitly on the interval \([t_0, T^*] \), see Kloeden & Platen (1992), with

\[
\hat{X}_t = X_{t_0} \exp \left\{ -\frac{1}{2} \int_{t_0}^t \hat{\sigma}_s (D^*(s, T^*) - D(s, T))^2 \, ds \right\} + \int_{t_0}^t \sqrt{\hat{\sigma}_s} (D^*(s, T^*) - D(s, T))^2 \, dW^1_s \tag{4.8}
\]

for \( t_0 \leq t \leq T^* \). For a European call option with strike \( K \) the option price formula (3.4) is given by

\[
\hat{u}_t = B_t^* \hat{E} \left( \left( \hat{X}_{T^*} - K \right)^+ \mid \mathcal{F}_t \right). \tag{4.9}
\]

Using the distributional properties of \( \hat{X}_{T^*} \) which can be determined from (4.8) with \( t = T^* \) we can now show, after some manipulation, that

\[
\hat{u}_t = B_t N \left( \hat{d} \left( t, \hat{X}_t \right) \right) - B_t^* K N \left( \hat{d} \left( t, \hat{X}_t \right) - \hat{\sigma}_t \right) \tag{4.10}
\]

for \( t_0 \leq t \leq T^* \) where

\[
\hat{\sigma}_t^2 = \int_{t_0}^{T^*} \hat{\sigma}_s \left( D^*(s, T^*) - D(s, T) \right)^2 \, ds,
\]

\[
\hat{d}(t, \hat{X}_t) = \frac{\ln(K/\hat{X}_t)}{\sigma_t} + \frac{\hat{\sigma}_t}{2},
\]

and \( N(\cdot) \) is the standard Gaussian distribution function.

With the above formulation we have used the bond price processes \( B \) and \( B^* \), corresponding to the extended Fong and Vasicek model, rather than the corresponding bond price processes for the extended Vasicek model as these choices generate better control variates for the two-factor valuation problem we are attempting to solve. The functions \( D(\cdot, T) \) and \( D(\cdot, T^*) \), as can be seen from (1.7), will be the same for both models.

We will now show that for the extended Fong and Vasicek model, both the short rate \( r_t \) and its instantaneous variance \( v_t \), can be expressed as functions of \( t, B_t \) and \( B_t^* \); that is in the form \( r_t = r(t, B_t, B_t^*) \) and \( v_t = v(t, B_t, B_t^*) \). This means in particular that the functions \( \theta_1, \theta_2, \theta_1^* \) and \( \theta_2^* \) given by (4.2) can be expressed as functions depending on \( t, B_t \) and \( B_t^* \).

Let us rewrite equation (1.6) for \( B_t \) and \( B_t^* \) in the following form

\[
\log B_t = -r_t D(t, T) + v_t F(t, T) + G(t, T)
\]

\[
\log B_t^* = -r_t D(t, T^*) + v_t F(t, T^*) + G(t, T^*). \tag{4.11}
\]

Solving the above equations for \( r_t \) and \( v_t \) we obtain

\[
r_t = \frac{F(t, T^*) \left[ \log B_t - G(t, T) \right] + F(t, T) \left[ G(t, T^*) - \log B_t^* \right]}{F(t, T) D(t, T^*) - F(t, T^*) D(t, T)}
\]

\[
v_t = \frac{G(t, T) - G(t, T^*)}{F(t, T) D(t, T^*) - F(t, T^*) D(t, T)}.
\]
4.5 EXPERIMENTAL RESULTS FOR EUROPEAN PRICING

\[ v_t = \frac{D(t, T^*) \left[ \log B_t - G(t, T) \right] + D(t, T) \left[ G(t, T^*) - \log B_t^* \right]}{D(t, T^*) F(t, T) - D(t, T) F(t, T^*)}, \] (4.12)

which are indeed functions only of \( t, B_t \) and \( B_t^* \). We realize that this procedure works only if the denominator in (4.12) for both \( r_t \) and \( v_t \) is non-zero. In fact if the denominators are close to zero both the short rate and instantaneous variance may explode which indicates a possible problem with the underlying model.

4.5 Experimental Results for European Pricing

The numerical procedures described in the previous sections provide powerful and flexible tools for the valuation of discount bonds and associated derivative securities for a class of two-factor term structure models that includes the extended Fong and Vasicek model. The main aim of this section is to demonstrate the power of the numerical and variance reduction methods employed rather than to explain in detail how those results were obtained.

For the numerical experiments described in this section we applied the predictor-corrector scheme (3.1) to compute outcomes for discrete time versions of the random variables \( X_{T^*} \) and \( \hat{X}_{T^*} \). These random variables appear in (3.5) and were used for the discrete time variance reduced estimator \( \hat{Z}_N \) given by (3.6). The Heun method was applied to approximate the functions \( D, F \) and \( G \) for both the long and short bond price processes \( B \) and \( B^* \), respectively. The variance reduction techniques of antithetic and stratified sampling as presented in Section 3.4, were also employed.

For the predictor-corrector method we used 16 discretization points with 20 batches each with 256 paths. Using the variance reduction techniques mentioned above, these choices for the numerical parameters mean that option prices can be computed typically with a relative statistical error of 0.2% at a 99% confidence level. For a 486, 33 MHz personal computer option prices at this level of accuracy can be calculated in about 10 seconds.

Simulation studies have also shown that discount bond prices for the extended Fong and Vasicek model, using equation (1.6), can be quickly and efficiently computed. Again for a 486, 33 MHz personal computer and for constant coefficients, prices accurate to 8 significant figures can be computed within one second. This degree of accuracy typically requires approximately 500-1000 discretization points. With time-dependent parameters and linear interpolation these prices are generally accurate to 6 significant figures. Several computations of discount bond prices per second can be made if less accuracy is needed.

Some examples of the type of yield curves, prices and hedge ratios that can be generated from the extended Fong and Vasicek model are illustrated in the figures below. For these experiments we used the extended Fong and Vasicek model given by (1.4). The default parameter values used were as follows: \( T = 4.0 \) (years), \( T^* = 1.0 \), \( \lambda = 0.0 \), \( \eta = 0.0 \), \( \gamma = 0.5 \), \( \bar{v} = 0.04 \), \( \xi = 0.04 \), \( \varphi = 0.0 \), with initial values \( r_0 = 0.055 \)
and $v_0 = 0.04$ at time $t_0 = 0$. The strike $K$ was set at 82.6 with both the long and short bond prices scaled with a face value of 100 monetary units. The time-dependent parameters $\alpha_t$ and $\bar{r}_t$, $0 \leq t \leq T$ were chosen as simple linear functions: $\alpha_t = 1.0 + 1.0 \left( t/T \right)$ and $\bar{r}_t = 0.08 - 0.04 \left( t/T \right)$ for $0 \leq t \leq T$. With these choices the market price of risk due to interest rate and volatility changes, see (1.8), are both set to zero.

Figure 4.5.1 shows a typical pattern of yield curves obtained from the Fong and Vasicek model using different values for the last reversion level $\bar{r}_T$. That is the slope of the linear function which defines the values of the reversion level $\bar{r}_t$ at time $t$, $0 \leq t \leq T$, is changed, and for each new value a yield curve is produced for the long bond with maturity $T$ extending out to 30 years. Note that we have obtained upward sloping and humped shaped curves. Other shapes such as downward sloping curves can be obtained by manipulating the time-dependent levels $\alpha_t$ and $\bar{r}_t$, $0 \leq t \leq T$ or other parameters.

Figure 4.5.2 depicts differences in yields between long bonds obtained from the extended Fong and Vasicek and corresponding extended Vasicek models. This diagram shows the differences in yields that arise from introducing stochastic volatility. The value $\xi = 0.0$ means that there is no stochastic volatility effect and no differences in yields result. Note that differences in yields even for high values of $\xi$ ($= 0.04$) and long maturities are relatively small, in fact less than 0.5%. However the impact of stochastic volatility, on option prices is much more pronounced as we will see shortly.
4.5. EXPERIMENTAL RESULTS FOR EUROPEAN PRICING

The calculation of hedge ratios is of considerable importance to risk managers. The
hedge ratios for $\eta_t = \frac{\partial}{\partial B} \bar{u}_t = \frac{\partial}{\partial B} u_t$, $0 \leq t \leq T^*$, given by (2.4) for both the extended
Fong and Vasicek and corresponding extended Vasicek models are shown in Figure 4.5.3
using different values of $\log(X_0/(K/100))$ which were obtained by keeping the model
parameters fixed and varying the strike $K$. The shapes of these two curves and the
differences between them are clearly similar to the corresponding curves obtained for
the Heston model with barriers, see Figure 3.4.3.

Illustrations of how stochastic volatility affects option prices are given in Figures
4.5.4 and 4.5.5. The first of these figures shows the differences in call option prices
between the extended Fong and Vasicek and corresponding extended Vasicek models
using different values of $\log(X_0/(K/100))$ and time to maturity $T^*$. We see, as is to be
expected, that the smile effect for prices diminishes as the time to maturity decreases.
Although not shown in this diagram for times to maturity of 1.0 years and nearly
at-the-money ($X_0 = (K/100)$) options, price differences of about 5% between the two
models were observed. However for deep out-of-the-money options price differences of
several hundred per cent can easily be obtained.

Figure 4.5.5 presents a different view of the stochastic volatility effect. Here we plot
price differences between the extended Fong and Vasicek and corresponding extended
Vasicek models using different values of $\log(X_0/(K/100))$ and the correlation parameter
$\varphi$. Note that the stochastic volatility effect is more pronounced for high (1.0) and low
(-1.0) values of $\varphi$. For $\varphi = 1.0$ and nearly at-the-money options, price differences of
15 – 25% were obtained.

The validation of results for the extended Fong and Vasicek model included the fol­
lowing checks and tests: Comparison with the explicit solution for the Vasicek model;
comparison with explicit solutions for certain special cases, with time-dependent param­
eters for the extended Vasicek model, see for example Hull & White (1990); estimation
of interpolation errors associated with time-dependent parameters based on compar­
isons with explicit solutions in cases where these could be obtained; investigation of
systematic errors based on the error propagation that occurs within the numerical
schemes and the use of different schemes; estimation of statistical errors based on dif­
ferent variance reduction techniques and long simulation runs, and stability analysis
for the underlying non-linear dynamics of the given model for a wide class of parameter
values and different numerical schemes.
Chapter 4. Valuation of European Bond Options

Figure 4.5.2: Differences in yields corresponding to different degrees of stochastic volatility.

Figure 4.5.3: Hedge ratios ($\eta_t$) for extended Fong and Vasicek and extended Vasicek models.
4.5. EXPERIMENTAL RESULTS FOR EUROPEAN PRICING

Figure 4.5.4: Price differences due to stochastic volatility for different values of the strike $K$ and time to maturity $T^*$.

Figure 4.5.5: Price differences curves due to stochastic volatility for different values of $\rho$ and the strike $K$. 
Chapter 5

Valuation of American Bond Options for Two-Factor Interest Rate Models

Most of the options traded on organized exchanges are American, which means the options can be exercised at any time up until the expiration date. For most types of American calls it can be shown that it is never optimal to exercise the option prior to maturity. However for American puts, early exercise of the option is optimal in some circumstances. This fact means that the difference in prices between American and European puts, referred to as the early exercise premium, is usually non-zero. Because of this, American puts are generally harder to value than their European cousins. This is particularly so in the case of complex multifactor models.

In the previous chapter we presented general stochastic approximation techniques for estimating the prices of European call options on discount bonds for a class of two-factor stochastic volatility term structure models. The purpose of this chapter is to extend this analysis to include American puts. The valuation of these instruments usually involves an analysis of the early exercise premium, however unfortunately this premium cannot be evaluated analytically. Even in the case of the Black-Scholes model no closed form expression currently exists for this premium. Consequently numerical procedures, in some form, are usually required to value the early exercise premium associated with these options.

An interesting and useful analysis of American puts is provided by Carr, Jarrow & Myneni (1992), which is related to the earlier work of McKean (1965) on a free boundary problem for the heat equation. They demonstrate that the early exercise premium for American puts can be expressed exactly using an integral expression which includes terms relating to the optimal early exercise boundary for the option. In the case of the Black and Scholes model this premium can be efficiently computed using a backward numerical technique.


In this chapter we propose an extension of the decomposition results obtained by
CHAPTER 5. VALUATION OF AMERICAN BOND OPTIONS

Carr, Jarrow & Myneni (1992) which is exact for an extended version of the Vasicek (1977) model. Using this decomposition together with appropriate stochastic and deterministic numerical methods we show how American puts can be efficiently computed for both the extended Vasicek (1977) model and the two-factor extended Fong & Vasicek (1991a,b) model considered in the previous chapter.

As is the case for European-style derivative securities the methods described in this chapter can be applied to more general multifactor term structure and stochastic volatility models such as those considered by Duffie & Kan (1994) and Ritchken & Sankarasubramanian (1995) which relate to a class of Heath, Jarrow & Morton (1992) models. We emphasize that the extended Fong and Vasicek model is used mainly as an example to illustrate the power and flexibility of the methods.

5.1 American Options for Two-Factor Bond Models

Let $W = (W^1, W^2)$ be a two-dimensional Brownian motion defined on the probability space $(\Omega, \mathcal{F}, P)$. As in Section 1.1 we assume the filtration $\mathcal{F} = (\mathcal{F}_t)_{t \geq t_0}$ is the $P$-augmentation of the natural filtration of $W$. A general methodology for pricing European-style contingent claims on discount bonds for two-factor models, extended to include time-dependent parameters, is given in Chapter 4 of this thesis.

As explained in Section 4.1, the bond prices $B_t$ and $B^*_t$ for discount bonds maturing at times $T$ and $T^*$, respectively, $t_0 \leq T^* < T$ are given by

$$B_t = E \left( \exp \left\{ - \int_t^T r_s \, ds \right\} \bigg| \mathcal{F}_t \right)$$

and

$$B^*_t = E \left( \exp \left\{ - \int_t^{T^*} r_s \, ds \right\} \bigg| \mathcal{F}_t \right)$$

for $t_0 \leq t \leq T^*$, where the expectation is taken with respect to the minimal equivalent martingale measure $P$. For complete markets this measure is the same as the risk neutral measure.

We will say that an interest rate term structure model is a two-factor model if the prices $B_t$ and $B^*_t$ of any two discount bonds at time $t$ with maturities $T$ and $T^*$, respectively, $t_0 \leq t \leq T^* \leq T$, follow a system of stochastic differential equations of the form

$$dB_t = r_t B_t \, dt + \theta_{1,t} B_t \, dW^1_t + \theta_{2,t} B_t \, dW^2_t$$

and

$$dB^*_t = r_t B^*_t \, dt + \theta^*_{1,t} B^*_t \, dW^1_t + \theta^*_{2,t} B^*_t \, dW^2_t,$$

for $t_0 \leq t \leq T^*$, where $\theta_{1,t}$, $\theta_{2,t}$, $\theta^*_{1,t}$ and $\theta^*_{2,t}$ are functions of the current time $t$, maturities $T$ and $T^*$, and the values, $B_t$ of the long bond, and $B^*_t$ of the short bond, at time $t$. 
5.1. AMERICAN OPTIONS FOR TWO-FACTOR BOND MODELS

We assume that the short rate \( r_t, \ t_0 \leq t \leq T \), can also be expressed as a function of these parameters, that is in the form

\[
r_t = r(t, B_t, T, B_t^*, T^*). \tag{1.3}
\]

In addition we assume that this function for \( r \) is sufficiently smooth to permit an application of Ito’s formula and therefore \( r_t \) can be expressed as the solution of a stochastic differential equation. These conditions for the extended Fong and Vasicek model given by (4.1.4), can be easily checked by using equations (4.4.12).

From the general pricing formula (1.1.5) with \( \tau = T \) the time \( t \) price of a European put \( P_t \) with strike \( K \) and maturity \( T^* \) on the long discount bond \( B \) maturing at time \( T \) is given by

\[
P_t = \mathbb{E} \left( \exp \left( - \int_t^{T^*} r_s \, ds \right) (K - B_{T^*})^+ \left| \mathcal{F}_t \right. \right), \tag{1.4}
\]

for \( t_0 \leq t \leq T^* \). This price \( P_t \) can also be computed from the European call option price \( C_t \), as given by (4.2.7) or (4.2.8), with \( C_t \) replacing \( U_t \), using put-call parity. In fact, if we use \( B_t^* \), the price of the short bond maturing at time \( T^* \), as the discount factor, \( p_t \) can be computed using the relation

\[
p_t = c_t + K B_t^* - B_t. \tag{1.5}
\]

Let us now consider the more delicate problem of computing the price \( P_t \) of an American put at time \( t \) with strike \( K \) and maturity \( T^* \) on the long discount bond \( B \). Using the work of Bensoussan (1984) and Karatzas (1988, 1989) this price can be written as

\[
P_t = \sup_{\tau \in [t, T^*]} \mathbb{E} \left( \exp \left\{ - \int_t^{T^*} r_s \, ds \right\} (K - B_{\tau})^+ \left| \mathcal{F}_t \right. \right), \tag{1.6}
\]

for \( t_0 \leq t \leq T^* \), where the supremum is taken over all stopping times \( \tau \) with values in \([t, T^*]\). It can be shown that under appropriate integrability and smoothness conditions there is a critical exercise function \( f : [t_0, T^*] \times \mathbb{R}^+ \to \mathbb{R}^+ \) with a corresponding stopping time \( \tau_f : \Omega \to [t_0, T^*] \) given by

\[
\tau_f(\omega) = \min \left( \inf \{ t \geq t_0 : B_t(\omega) \leq f(t, B_t^*(\omega)) \} , T^* \right) \tag{1.7}
\]

at which the supremum for the American price process is attained. This means that

\[
P_t = \mathbb{E} \left( \exp \left\{ - \int_t^{\tau_f} r_s \, ds \right\} (K - B_{\tau_f})^+ \left| \mathcal{F}_t \right. \right) \tag{1.8}
\]

for \( t_0 \leq t \leq T^* \). We assume that \( B_{t_0} > f(t_0, B_{t_0}^*) \), otherwise the option would be immediately exercised. Valuing an American put is therefore a problem in optimal stochastic control. That is, one has to find the critical exercise function \( f \) for which the supremum is attained.

Since our model contains the two stochastic processes \( B \) and \( B^* \), the corresponding critical exercise boundary

\[
\left\{ (t, B_t^*(\omega), f(t, B_t^*(\omega)) \in [t_0, T] \times \mathbb{R}^2 : \omega \in \Omega \right\}
\]
now becomes a surface below which the option should be exercised. With this formulation we have expressed the exercise condition in terms of critical values of $B_t$ varying according to different values of $B_t^*$. Other formulations are also possible. For example, critical values of $B_t$ varying according to different values of $r_t$ or, in the case of the extended Fong and Vasicek model, critical values of $B_t$ varying according to different values of $v_t$, the instantaneous variance of the short rate, see equations (4.2.4).

Comparing (1.4) and (1.6) we see that the price of an American put $P_t$ must be greater than or equal to the price of the corresponding European put $p_t$. The difference between $P_t$ and $p_t$, denoted by

$$e_t = P_t - p_t$$  \hspace{1cm} (1.9)

is therefore the early exercise premium of the American put.

Note that using the procedures described in Section 4.3 we can express the price of the European put $p_t$, equivalently in the form

$$p_t = B_t^* \hat{E} \left( (K - B_T)_{+} \mid \mathcal{F}_t \right)$$ \hspace{1cm} (1.10)

for $t \in [t_0, T^*]$, where $\hat{E}$ denotes expectation with respect to an adjusted measure $\hat{P}$, the so-called forward measure, see El Karoui & Rochet (1989), Jamshidian (1989b) or Brace & Musiela (1994). This formula should be compared to (1.4) which prices a European put under the minimal equivalent martingale measure $P$.

5.2 Representation of the Early Exercise Premium

We will now consider the problem of finding an exact representation of the early exercise premium for the two-factor model given by (1.2). We will find a random integral representation which is related to the work of McKean (1965), Kim (1990), Jacka (1991), Carr, Jarrow & Myneni (1992) and Chesney, Elliott & Gibson (1991).

The critical exercise function $f$, for the American put, see (2.7) and (2.8) can be used to divide the domain $\mathcal{D} = [t_0, T^*] \times \mathbb{R}^2$ into a continuation region $\Gamma^c_f$ and a stopping region $\Gamma^s_f$ defined by

$$\Gamma^c_f = \{(t, B_t, B_t^*) \in \mathcal{D} : B_t > f(t, B_t^*)\}$$ \hspace{1cm} (2.1)

and

$$\Gamma^s_f = \{(t, B_t, B_t^*) \in \mathcal{D} : B_t \leq f(t, B_t^*)\}.$$ \hspace{1cm} (2.2)

Consider now a savings account process $\beta$ which accumulates interest continuously at rate $r_t$, and which satisfies the linear growth equation (1.1.2) with solution (1.1.3) for $t_0 \leq t \leq T$.

Using the process $\beta$ we can express $B_t$, given by (1.1), and $P_t$, given by (1.8), equivalently in the forms

$$B_t = E \left( \frac{B_t}{\beta_t} \mid \mathcal{F}_t \right)$$
5.2. REPRESENTATION OF THE EARLY EXERCISE PREMIUM

\[ \beta_t E\left( \frac{1}{\beta_{t+T}} \left| \mathcal{F}_t \right. \right) \]  
\[ (2.3) \]

and

\[ P_t = E\left( \frac{\beta_{t\wedge T}}{\beta_{t+T}} \left( K - B_{t+T} \right)^+ \left| \mathcal{F}_t \right. \right) \]
\[ = \beta_{t\wedge T} E\left( \frac{1}{\beta_{t+T}} \left( K - B_{t+T} \right)^+ \left| \mathcal{F}_t \right. \right) \]
\[ (2.4) \]

for \( t_0 \leq t \leq T^* \).

The condition (1.3) means that we can write the price of the American put \( P_t \), \( t_0 \leq t \leq T^* \), in the form \( P_t = P(t, B_t, B_t^*) \) to express its functional dependence on the state variables \( t, B_t \) and \( B_t^* \) for \( (t, B_t, B_t^*) \in \Gamma_t^* \). We now define a scaled extension \( Z_t \) of the price \( P_t \) to the domain \( \mathcal{D} = \Gamma_t^* \cup \Gamma_{t+T}^* \) by

\[ Z_t = Z(t, B_t, B_t^*, \beta_t) = \begin{cases} \beta_t^{-1} P(t, B_t, B_t^*) & : (t, B_t, B_t^*) \in \Gamma_t^* \\ \beta_t^{-1} (K - B_t)^+ & : (t, B_t, B_t^*) \in \Gamma_{t+T}^* \end{cases} \]
\[ (2.5) \]

In this form we are considering \( Z_t \) as being dependent on the state variables \( t, B_t, B_t^* \) and \( \beta_t \).

Using the so-called 'smooth-fit' conditions for the American put (see Myneni (1992)) we know that \( Z_t \) is continuous and has a continuous partial derivative \( \frac{\partial Z}{\partial B} \) on the whole of the domain \( \mathcal{D} \).

If we denote by \( X = \{(X_t^1, X_t^2, X_t^3) : t_0 \leq t \leq T\} \) the vector diffusion process with components \( X^1 = B, X^2 = B^* \) and \( X^3 = \beta \) we can apply an extension of Ito's formula, see Theorem 2.10.1 in Krylov (1980), to obtain

\[ Z(T^*, B_{T^*}, B_{T^*}^*, \beta_{T^*}) = Z(t, B_t, B_t^*, \beta_t) + \int_t^{T^*} L^0 Z(s, B_s, B_s^*, \beta_s) \, ds + \sum_{j=1}^{2} \int_t^{T^*} L^j Z(s, B_s, B_s^*, \beta_s) \, dW_s^j, \]
\[ (2.6) \]

where \( L^0, L^1 \) and \( L^2 \) are the operators

\[ L^0 = \frac{\partial}{\partial s} + \sum_{i=1}^{3} a^i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,k=1}^{3} \sum_{j=1}^{2} b^{i,j} b^{k,j} \frac{\partial^2}{\partial x_i \partial x_k}, \]

\[ L^j = \sum_{i=1}^{3} b^{i,j} \frac{\partial}{\partial x_i}, \quad \text{for } j \in \{1, 2\}, \]

and \( a^i, b^{i,j}, i \in \{1, 2, 3\}, j \in \{1, 2\} \) are the drift and diffusion coefficients for the component processes \( B \) and \( B^* \) as used in (1.2), and \( \beta \) as given by (1.1.2).

This theorem requires that the function \( Z \) be an element of certain Sobolev or related spaces and that the generalized spatial derivatives of \( Z \) up to and including the
second order exist. A similar representation to (2.6) which relies on a version of Ito’s lemma for piecewise convex functions, has been employed by Carr, Jarrow & Myneni (1992) for American pricing, where the underlying security follows a one-dimensional geometric Brownian motion. We have used the theorem by Krylov because it is more easily applied to multidimensional American pricing problems.

Taking conditional expectations under $P$, of both sides of (2.6) at time $t$, we see that

$$E(Z(T^*, B_{T^*}, B_{T^*}^*, \beta_{T^*}) \mid \mathcal{F}_t) = Z(t, B_t, B_t^*, \beta_t) + E \left( \int_t^{T^*} L^0 Z(s, B_s, B_s^*, \beta_s) ds \mid \mathcal{F}_t \right).$$

(2.7)

The relation (2.3) shows that $\beta_t^{-1} B_t = E(1/\beta_T \mid \mathcal{F}_t)$ and therefore applying the Kolmogorov backward equation using the diffusion operator $L^0$ we have

$$L^0 \beta_t^{-1} B_t = 0$$

(2.8)

for $t_0 \leq t \leq T^*$. Also, according to (2.4) we can write

$$\beta_{t \wedge \tau_f}^{-1} P(t, B_t, B_t^*) = E(1/\beta_{\tau_f} (K - B_{\tau_f})^+ \mid \mathcal{F}_t)$$

for $t \leq \tau_f$. Consequently applying the Kolmogorov backward equation once again within the continuation region $\Gamma_f$ we can infer that

$$L^0 \beta_{t \wedge \tau_f}^{-1} P(t, B_t, B_t^*) = 0$$

(2.9)

for $(t, B_t, B_t^*) \in \Gamma_f$. We observe that if $(t, B_t, B_t^*) \in \Gamma_f$, then $t \leq \tau_f$. These results together with (1.1.2) show that

$$L^0 Z_t = L^0 Z(t, r_t, B_t, \beta_t) = \begin{cases} 0 & : (t, B_t, B_t^*) \in \Gamma_f^c \\ -K r_t \beta_t^{-1} & : (t, B_t, B_t^*) \in \Gamma_f^s \end{cases}$$

(2.10)

for $t_0 \leq t \leq T^*$. The expression for $L^0 Z_t$ in the stopping region $\Gamma_f^s$ can also be obtained directly from (2.5) using an $L^0$ expansion based on (1.2) and (1.1.2).

Consequently (2.7) can now be expressed in the form

$$E(Z(T^*, B_{T^*}, B_{T^*}^*, \beta_{T^*}) \mid \mathcal{F}_t) = Z(t, B_t, B_t^*, \beta_t) - K E \left( \int_t^{T^*} r_s \beta_s^{-1} 1_{(B_s \leq f(s, B_s^*))} ds \mid \mathcal{F}_t \right),$$

(2.11)

for $t_0 \leq t \leq T^*$. Note also from the definition of $Z_t$ given by (2.5) that

$$Z_{T^*} = Z(T^*, B_{T^*}, B_{T^*}^*, \beta_{T^*}) = \beta_{T^*}^{-1} (K - B_{T^*})^+.$$
Combining this relation, (1.4) and (2.11) with the definitions of $Z_t$ given by (2.5), we can infer that

$$P_t = p(t, B_t, B^*_t)$$

$$= E \left( \exp \left( - \int_t^{T^*} r_s \, ds \right) (K - B_{T^*})^+ \big| \mathcal{F}_t \right)$$

$$= E \left( \frac{\beta_t}{\beta_{T^*}} (K - B_{T^*})^+ \big| \mathcal{F}_t \right)$$

$$= E (\beta_t Z (T^*, B_{T^*}, B^*_{T^*}, \beta_{T^*}) \big| \mathcal{F}_t)$$

$$= \beta_t E (Z (T^*, B_{T^*}, B^*_{T^*}, \beta_{T^*}) \big| \mathcal{F}_t)$$

$$= \beta_t Z(t, B_t, B^*_t, \beta_t) - \beta_t K E \left( \int_t^{T^*} r_s \beta_s^{-1} \mathbf{1}_{\{B_s \leq f(s, B^*_t)\}} \, ds \big| \mathcal{F}_t \right)$$

$$= P(t, B_t, B^*_t) - K \left( \int_t^{T^*} \beta_t r_s \beta_s^{-1} \mathbf{1}_{\{B_s \leq f(s, B^*_t)\}} \, ds \big| \mathcal{F}_t \right)$$

$$= P(t, B_t, B^*_t) - K \left( \int_t^{T^*} r_s \exp \left( - \int_t^{s} r_u \, du \right) \mathbf{1}_{\{B_s \leq f(s, B^*_t)\}} \, ds \big| \mathcal{F}_t \right)$$

(2.12)

for $(t, B_t, B^*_t) \in \Gamma^*_j$. The second term on the right hand side of (2.12) is thus, according to (1.9), an expression for the early exercise premium $e_t$, which by Fubini's Theorem can be written in the form

$$e_t = e(t, B_t, B^*_t) = K \int_t^{T^*} E \left( \mathbf{1}_{\{B_s \leq f(s, B^*_t)\}} \big| \mathcal{F}_t \right) \, ds. \quad (2.13)$$

We remark that the relation (2.7), which is the key result we need for the proof of (2.13), can be derived under weaker conditions compared to (2.6), see Krylov (1980) Chapter 2.

We will now consider the problem of finding an integral representation of the early exercise premium in a form which uses $B^*$-discounting and the forward measure $\hat{P}$. We will provide only an outline of the required steps as this result, although of theoretical interest, is not further used in the remaining part of this chapter.

To do this we define $\tilde{Z}_t$ by the relation

$$\tilde{Z}_t = \tilde{Z}(t, B_t, B^*_t) = \begin{cases} 1/B^*_t P(t, B_t, B^*_t) & : (t, B_t, B^*_t) \in \Gamma^*_j \\ 1/B^*_t (K - B_t)^+ & : (t, B_t, B^*_t) \in \Gamma^*_j \end{cases}$$

Using the forward measure $\hat{P}$ we can write, similar to (2.4), (see also (1.10)) the relation

$$P_t = B^*_{t\wedge \tau} \hat{E} \left( (K - B_{\tau^*})^+ \big| \mathcal{F}_t \right). \quad (2.14)$$
The diffusion operators $L^0$, $L^1$ and $L^2$ are now replaced by similar ones denoted by $\tilde{L}^0$, $\tilde{L}^1$ and $\tilde{L}^2$. These have the same form as given in (2.6) except the drift and diffusions coefficients $\alpha^i$ and $\beta^i,j$, $i,j \in \{1,2\}$ now appear in a different form to account for the new measure $\tilde{P}$. We also have only two state variables $B_t$ and $B^*_t$, $t_0 \leq t \leq T^*$ for the function $\tilde{Z}$ as $1/\theta_t$, the discount factor, has been replaced with $1/B^*_t$.

As in (2.9) and using (2.14) we can infer that

$$\tilde{L}^0 \left( P(t,B^*_t,B_t) / B^{*}_{t_{t\wedge r_f}} \right) = 0$$

for $(t,B_t,B^*_t) \in \Gamma^*_f$.

These results together with (1.2) can be used to show, similar to (2.10), that

$$\tilde{L}^0 \tilde{Z}_t = \left\{ \begin{array}{ll}
0 & : (t,B_t,B^*_t) \in \Gamma^*_f \\
-K r_t/B^*_t & : (t,B_t,B^*_t) \in \Gamma^*_f
\end{array} \right. \quad (2.15)$$

for $t_0 \leq t \leq T^*$. Applying Ito's formula as in (2.7), taking conditional expectations, and using (2.15) we can prove that

$$\mathbb{E} \left( \tilde{Z} (T^*,B_{T^*},B^*_{T^*}) \mid \mathcal{F}_t \right) = \tilde{Z}(t,B_t,B^*_t)$$

$$-K \mathbb{E} \left( \int_t^{T^*} r_s/B^*_t 1_{\{B_s \leq f(s,B^*_s)\}} \, ds \mid \mathcal{F}_t \right)$$

for $t_0 \leq t \leq T^*$. From this result, the definition of $\tilde{Z}_t$, (1.10), (2.14) and the maturity condition $B^*_{T^*} = 1$ we now obtain the relation

$$p(t,B_t,B^*_t) = B^*_t \mathbb{E} \left( (K - B_{T^*})^+ \mid \mathcal{F}_t \right)$$

$$= B^*_t \mathbb{E} \left( \tilde{Z} (T^*,B_{T^*},B^*_{T^*}) \mid \mathcal{F}_t \right)$$

$$= B^*_t \tilde{Z} (t,B_t,B^*_t) - B^*_t K \mathbb{E} \left( \int_t^{T^*} r_s/B^*_t 1_{\{B_s \leq f(s,B^*_s)\}} \, ds \mid \mathcal{F}_t \right)$$

$$= P(t,B_t,B^*_t) - K B^*_t \mathbb{E} \left( \int_t^{T^*} r_s/B^*_t 1_{\{B_s \leq f(s,B^*_s)\}} \, ds \mid \mathcal{F}_t \right)$$

for $(t,B_t,B^*_t) \in \Gamma^*_f$. Consequently using Fubini's Theorem an alternate representation for the early exercise premium $e_t$ using $B^*$-discounting and the forward measure $\tilde{P}$ can be expressed in the form

$$e_t = K B^*_t \mathbb{E} \left( \int_t^{T^*} \tilde{Z} (T^*,B_{T^*},B^*_{T^*}) \mid \mathcal{F}_t \right) \, ds,$$

for $(t,B_t,B^*_t) \in \Gamma^*_f$. 
5.3 Explicit Representation of the Early Exercise Premium for the Extended Vasicek Model

We will now consider the valuation of American puts for an extended version of the Vasicek (1977) model which can be regarded as a one-factor version of the Fong & Vasicek (1991a,b) model. We take \( \hat{r}_t \), the short rate, and \( \hat{v}_t \), its instantaneous variance, to be given by the following pair of differential equations

\[
\begin{align*}
    d\hat{r}_t &= (\alpha_t (\hat{r}_t - \bar{r}) + \lambda \hat{v}_t) \ dt + \sqrt{\hat{v}_t} \ dW_t^1 \\
    d\hat{v}_t &= \gamma (\bar{v} - \hat{v}_t) \ dt
\end{align*}
\]

for \( t_0 \leq t \leq T^* \). In this model \( \alpha = \{\alpha_t, \ t_0 \leq t < \infty\} \) and \( \bar{r} = \{\bar{r}_t, \ t_0 \leq t < \infty\} \) are time-dependent deterministic functions and \( \lambda, \gamma, \bar{v} \) are constants. This formulation corresponds to the extended Fong and Vasicek model (4.1.4) with \( \xi = 0 \) and is closely related to the control variate formulations used in Section 4.4.

With this model the \( \hat{v}_t \) component can be solved explicitly on the interval \([t_0, T^*]\) to yield

\[
\hat{v}_t = \bar{v} + (\hat{v}_{t_0} - \bar{v}) e^{-\gamma(t-t_0)}
\]

for \( t_0 \leq t \leq T^* \) with initial value \( \hat{v}_{t_0} \). A similar formula was obtained for the instantaneous variance component given by (3.3.7), see also (4.4.7).

We therefore have only one source of noise for the short rate \( \hat{r}_t \) with a time-dependent diffusion coefficient for \( \hat{r}_t \) and mean-reverting drift coefficients for both \( \hat{r}_t \) and \( \hat{v}_t \). This type of structure which extends the original Vasicek (1977) model, has also been examined by Hull & White (1990).

The pricing of American puts for this extended Vasicek model is of independent interest. In addition, we will later use this model to build control variates for the more complex two-factor extended Fong and Vasicek model. As is explained in Section 2.4 these control variates are used to construct variance reduced estimators which combined with appropriate stochastic numerical methods enable us to perform fast and accurate valuations of American options and corresponding hedge ratios.

Let us denote by \( \hat{B}^*_t \) and \( \hat{B}_t \), \( t_0 \leq t \leq T^* \) the price at time \( t \) of discount bonds maturing at times \( T^* \) and \( T \), respectively, where the short rate process \( \hat{r} \) satisfies (3.1). \( \hat{B}^*_t \) and \( \hat{B}_t \) are given by equations of the form (1.1) with \( \hat{r} \) replacing \( r \).

Using (1.10) and the \( B^*_t \)-discounted process \( \hat{X} = \{\hat{X}_t = \hat{B}_t/B^*_t : t_0 \leq t \leq T^*\} \) we can express the price \( \hat{p}_t = \hat{p}(t, \hat{r}_t) \) of the European put in the form

\[
\hat{p}_t = \hat{B}^*_t \hat{E}\left( (K - \hat{X}_{T^*})^+ \mid \mathcal{F}_t \right)
\]

for \( t_0 \leq t \leq T^* \), where \( \hat{E} \) denotes expectation with respect to the forward measure \( \hat{P} \) corresponding to the system (3.1). This forward measure is defined by the Radon-Nikodym derivative given by (4.2.9). Because the coefficients \( \theta^*_1, \theta^*_2, \theta_1 \) and \( \theta_2 \) given
by (4.4.2) depend on the structure of the underlying system of stochastic differential equations, the forward measure \( \hat{P} \) may be different from the forward measure \( \tilde{P} \) used in Sections 5.1 and 5.2.

Furthermore, since \( \hat{t} \) is a time-dependent function, the stochastic differential equation for \( \hat{X} \) reduces to the case of the first equation in (4.4.6), where the functions \( D(\cdot, T^*) : [t_0, T^*] \to \mathbb{R} \) and \( D(\cdot, T) : [t_0, T] \to \mathbb{R} \) satisfy an ordinary differential equation of the form (4.1.7) with boundary condition \( D(T^*, T^*) = D(T, T) = 0 \). As mentioned in Chapter 4 these functions are the same for both the extended Fong and Vasicek model and the extended Vasicek model given by (4.1.4) and (3.1), respectively.

This equation for \( \hat{X} \) can be solved explicitly on the interval \([t_0, T^*]\), and is given by (4.4.8). The distributional properties of the random variable \( \hat{X}_{T^*} \) are known and consequently the option price given in (3.3) can also be computed explicitly in a manner similar to (4.4.10). In fact applying (3.3) and these distributional properties we can show, after some manipulation, that

\[
\hat{p}_t = \hat{B}_t^* K N(d(t, \hat{X}_t)) - \hat{B}_t N(d(t, \hat{X}_t) - \hat{\sigma}_t) \tag{3.4}
\]

for \( t_0 \leq t \leq T^* \), where \( d(t, \hat{X}_t) \) and \( \hat{\sigma}_t \) are as given in (4.4.8) and \( N(\cdot) \) is the standard Gaussian distribution function.

From the relation \( N(x) + N(-x) = 1, \ x \in \mathbb{R} \), equation (3.4) can be rewritten in the form

\[
\hat{p}_t = \hat{B}_t^* K - \hat{B}_t + \hat{B}_t N\left(\sigma_t - d(t, \hat{X}_t)\right) - \hat{B}_t^* K N\left(-d(t, \hat{X}_t)\right)
\]

\[
= \hat{B}_t^* K - \hat{B}_t + \hat{c}_t, \tag{3.5}
\]

where, \( \hat{c}_t \) denotes the corresponding price for a European call as given by (4.4.10) with \( \hat{u}_t \) replacing \( \tilde{u}_t \). This result can also be obtained by put-call parity arguments with \( \hat{B}_t^* \) as the discount factor.

Let us now consider the problem of pricing an American put for the simplified model given by (3.1). As is shown in (2.12) the price of an American put \( \hat{P}_t = \hat{P}(t, \hat{r}_t) \) can be decomposed into the corresponding European put price \( \hat{p}_t \) plus the early exercise premium \( \hat{\epsilon}_t = \hat{\epsilon}(t, \hat{r}_t) \) with

\[
\hat{P}_t = \hat{p}_t + \hat{\epsilon}_t, \tag{3.6}
\]

where

\[
\hat{\epsilon}_t = K E\left(\int_t^{T^*} \exp\left(-\int_t^s \hat{r}_u \, du\right) \hat{r}_s 1_{\{B_s \leq f(s)\}} \, ds \bigg| \mathcal{F}_t\right). \tag{3.7}
\]

Here \( E \) denotes expectation with respect to the corresponding minimal equivalent martingale measure. Because we are dealing with only a one-factor model of the short rate this becomes the risk neutral measure. Note that the critical exercise function \( f : [t_0, T] \to \mathbb{R} \) is one-dimensional and does not depend on the short maturing bond price as is the case for the more complex two-factor model.
We will now attempt to find a more explicit representation of the early exercise premium given by (3.7) by using the separation principle described in Goldman et al. (1995a). In fact applying this principle it can be shown that the early exercise premium $\hat{\epsilon}_t$ can be expressed in the form

$$\hat{\epsilon}_t = K \int_t^T E \left( \exp \left( - \int_t^s \hat{r}_u \, du \right) \mid \mathcal{F}_t \right) E \left( \hat{r}_s \mathbf{1}_{\{ B_s \leq f(s) \}} \mid \mathcal{F}_t \right) \, ds,$$

where $\hat{E}$ denotes expectation with respect to the forward measure $\hat{P}$, which for the system of equations (3.1) is defined by the Radon-Nikodym derivative

$$\frac{d\hat{P}}{dP} = \exp \left( - \int_{t_0}^T \frac{1}{2} \left( \hat{\theta}_s^* \right)^2 \, ds + \int_{t_0}^T \hat{\theta}_s^* \, dW^1_s \right)$$

with

$$\hat{\theta}_s^* = \sqrt{\hat{v}_s} \frac{1}{\hat{B}_s} \frac{\partial \hat{B}_s}{\partial r}.$$  

This specification for the derivative $d\hat{P}/dP$ follows from (4.2.9) and (4.1.3) and setting $\xi = 0$ as is required from (3.1). If we change the underlying probability measure to $\hat{P}$ the Girsanov’s Theorem states that the process $\hat{W}^1_{t_*} = W^1_{t_*} - \int_{t_0}^{t_*} \hat{\theta}_u^* \, du$, $t_0 \leq t \leq T$, will be a Wiener process under $\hat{P}$. Using this Wiener process the stochastic differential equation for the $\hat{r}$ component in (3.1) becomes

$$d\hat{r}_s = \left( \alpha_s (\hat{r}_s - \hat{r}_s) + \hat{v}_s \lambda + \sqrt{\hat{v}_s} \hat{\theta}_s^* \right) \, ds + \sqrt{\hat{v}_s} \, d\hat{W}^1_s$$

for $t_0 \leq t \leq T^*$.  

This stochastic differential equation can be solved explicitly on the interval $[t, s]$, $t_0 \leq t \leq s \leq T^*$, see Kloeden & Platen (1992) with

$$\hat{r}_s = \phi_{t,s} \left( \hat{r}_t + \int_t^s \phi_{t,u}^{-1} \left( \alpha_u \hat{r}_u + \hat{v}_u \lambda + \sqrt{\hat{v}_u} \hat{\theta}_u^* \right) \, du \right)$$

$$\quad + \phi_{t,s} \int_t^s \phi_{t,u}^{-1} \sqrt{\hat{v}_u} \, d\hat{W}_u^1,$$

where

$$\phi_{t,s} = \exp \left( - \int_t^s \alpha_u \, du \right),$$

for $t_0 \leq t \leq s \leq T^*$ with initial value $\hat{r}_t$ at time $t$.

It is clear from the structure of (3.12) that the random variable $\hat{r}_s$ is Gaussian with mean given by

$$\mu_{t,s} = \phi_{t,s} \left( \hat{r}_t + \int_t^s \phi_{t,u}^{-1} \left( \alpha_u \hat{r}_u + \hat{v}_u \lambda + \sqrt{\hat{v}_u} \hat{\theta}_u^* \right) \, du \right),$$

and variance

$$\sigma^2_{t,s} = \phi_{t,s}^2 \int_t^s \frac{\hat{v}_u}{\phi_{t,u}^2} \, du$$

for $t_0 \leq t \leq s \leq T^*$.  

The expression for the early exercise premium $\hat{e}_t$ at time $t$, $t_0 \leq t \leq T$ given by (3.8) is formulated using critical values of the long bond $\hat{B}$. We will now show that the critical boundary formulation can be expressed in terms of critical values of the short rate $\hat{r}$. For the extended Vasicek model given by (3.1) the formula for the long bond price $\hat{B}_t$, at time $t$, given by (4.1.6) can be simplified and written in the form

$$\hat{B}_t = \hat{B}(t,T,\hat{r}_t) = \exp\{-r_t D(t,T) + \hat{H}(t,T)\}$$

(3.15)

for $t_0 \leq t \leq T$, where the function $D$ is as specified in (4.1.7) and $\hat{H}$ is some other time-dependent function. This equation can be expressed in the form

$$\hat{r}_t = \left(\hat{H}(t,T) - \log(\hat{B}_t)\right) / D(t,T)$$

for $t_0 \leq t \leq T$. The equation for $D(\cdot, T)$ can be solved on the interval $[t_0, T]$ with

$$D(t,T) = \phi_{t,T} \int_t^T \phi_{u,T}^{-1} du$$

for $t_0 \leq t \leq T$, where $\phi$ is given in (3.12). We now assume that $\alpha_t > 0$ for $t_0 \leq t \leq T$ so that from the definition of $\phi$ we can infer that $D(t,T) > 0$ for $t_0 \leq t \leq T^*$ with $T^* < T$. Consequently the above expression for $\hat{r}_t$ is well-defined. This means that critical values for $\hat{B}_t$ can be translated into critical values for the short rate $\hat{r}_t$ at time $t$. If we denote by $f_\hat{r} : [t_0, T] \to \mathbb{R}$ the critical exercise function for $\hat{r}$, then using the above equation for $\hat{r}_t$ we can write

$$f_\hat{r}(x) = \left(\hat{H}(t,T) - \log(f(x))\right) / D(t,T)$$

for $x \in \mathbb{R}$, where $f$ is the critical exercise function for the long bond $\hat{B}$. Using the critical boundary function $f_\hat{r}$ the early exercise premium $\hat{e}_t$ given by (3.8) becomes

$$\hat{e}_t = K \int_t^T E \left( \exp\left(-\int_t^s \hat{r}_u du\right) \right) \hat{E} \left( \hat{r}_s 1_{\{\hat{r}_s \geq f_\hat{r}(s)\}} \big| \mathcal{F}_t \right) ds.$$  

(3.16)

Note that for points $(t, \hat{B}_t)$ in the stopping region $\Gamma^*_{f_\hat{r}} \hat{B}_t \leq f(t)$ so that from (3.15) $\hat{r}_t \geq f_\hat{r}(t)$ as is indicated in (3.16). In the next section we will find more convenient representations for the time-dependent functions $D$ and $\hat{H}$.

With these results established, we are now in a position to evaluate the second term in the integrand of (3.16). In particular using the distributional properties of the random variable $\hat{r}_s$, $t_0 \leq t \leq s \leq T^*$, with initial value $\hat{r}_t$ at time $t$ we can show that

$$\hat{E} \left( \hat{r}_s 1_{\{\hat{r}_s \geq f_\hat{r}(s)\}} \big| \mathcal{F}_t \right) = \int_{f_\hat{r}(s)}^\infty \frac{x}{\sigma_{t,s}} N' \left( \frac{x - \mu_{t,s}}{\sigma_{t,s}} \right) dx,$$

(3.17)

where $N'(y) = \frac{1}{\sqrt{2\pi}} \exp(-y^2/2)$ is the standard Gaussian density function and $\mu_{t,s}$ and $\sigma_{t,s}$ are given by (3.13) and (3.14), respectively.

The integral term on the right hand side of (3.17) can be solved explicitly in the form

$$\hat{E} \left( \hat{r}_s 1_{\{\hat{r}_s \geq f_\hat{r}(s)\}} \big| \mathcal{F}_t \right) = \mu_{t,s} N \left( \frac{\mu_{t,s} - f_\hat{r}(s)}{\sigma_{t,s}} \right) + \sigma_{t,s} N' \left( \frac{f_\hat{r}(s) - \mu_{t,s}}{\sigma_{t,s}} \right),$$

(3.18)
where $N(\cdot)$ is the standard Gaussian distribution function.

Let $\bar{B}_{t,s} = \bar{B}(t,s,\bar{r}_t)$ be the price of a discount bond maturing at time $s$ which is evaluated at time $t$, $t_0 \leq t \leq s \leq T$, and which evolves according to the short rate $\bar{r}$. As in (1.1) we can write this price in the form

$$\bar{B}_{t,s} = \bar{B}(t,s,\bar{r}_t) = E\left(\exp\left(-\int_t^s \bar{r}_u \, du\right) \mid \mathcal{F}_t\right) \quad (3.19)$$

for $t_0 \leq t \leq s \leq T$, where $E$ denotes expectation under the measure $P$. Combining (3.16), (3.18) and (3.19) we see that

$$\bar{e}_t = K \int_t^{T^*} \mathbb{B}_{t,s} \left[\mu_{t,s} N\left(\frac{\mu_{t,s} - f_T(s)}{\sigma_{t,s}}\right) + \sigma_{t,s} N'\left(\frac{f_T(s) - \mu_{t,s}}{\sigma_{t,s}}\right)\right] \, ds \quad (3.20)$$

for $t_0 \leq t \leq T^*$. We now have an explicit representation of the early exercise premium $\bar{e}_t$ at time $t$ for the extended Vasicek model (3.1). This representation is crucial for the efficient valuation of American put prices for the two-factor extended Fong and Vasicek model as we will see in the next section. It is also of independent mathematical interest because it provides for a one factor Gaussian interest rate model of the type specified in (3.1) an exact formula for the early exercise premium. The integral representation given by (3.20) can be compared to results obtained by, for example, Carr, Jarrow & Myneni (1992) for the Black-Scholes model.

The critical exercise boundary function $f_T \bar{r}$ for the short rate $\bar{r}$ can now be obtained by solving (implicitly) the equation

$$\hat{P}_t = K - \bar{B}(t,T,\bar{r}_t) = \hat{p}(t,\bar{r}_t) + \hat{e}(t,\bar{r}_t) \quad (3.21)$$

for $t_0 \leq t \leq T^*$, Thus we have to solve equation (3.21) for $\hat{r}_t$, using different values of $t$. These values then become the critical exercise boundary values $f_T \bar{r}$ for the short rate $\bar{r}$. Note that equation (3.21), if evaluated at time $t$ using (3.20), requires computation of the values $f_T(s)$ at all future times $s$, $t \leq s \leq T^*$. This observation means that the critical exercise boundary is usually more conveniently obtained or estimated using a backward iteration technique starting at time $T^*$. An alternative is to find some approximation for $f_T \bar{r}$, say $f^{(1)}_T \bar{r}$, use this in (3.20) and then compute a new approximation for $f_T \bar{r}$, say $f^{(2)}_T \bar{r}$ from (3.21).

The terms $\bar{B}_{t,s}, \mu_{t,s} \text{ and } \sigma_{t,s}$ appearing in (3.20) can all be efficiently evaluated using non-stochastic numerical methods. Consequently, for a given choice of $f_T \bar{r}$, the early exercise premium $\bar{e}_t$ can also be efficiently computed. However, several computations of the integral in (3.20), working backwards in time or using the successive boundary approximations $f^{(i)}_T \bar{r}$, $i = 1, 2, \ldots$, are required to estimate the critical exercise function $f_T \bar{r}$. Once this is obtained calculation of the price of an American put is very fast as it requires only one additional estimate of the integral.

Repeated evaluations of the integral in (3.20) require many computations of the terms $\bar{B}_{t,s}, \mu_{t,s}$ and $\sigma_{t,s}$ using different values of $t$ and $s$, $t_0 \leq t \leq s \leq T^*$. These terms can be estimated at certain discrete points $t_i$, $i \in \{1, \ldots, N\}$, over a two-dimensional
grid. We can then use two-dimensional interpolation methods to estimate values between the points. However this evaluation procedure over a two-dimensional grid is computationally expensive.

In the next section we will show how the terms \( \hat{B}_{t,s} \), \( \mu_{t,s} \) and \( \sigma_{t,s} \) can be effectively calculated in a manner which avoids the requirement of dealing with two-dimensional interpolation problems.

### 5.4 Computation of Terms for the Extended Vasicek Model

To compute the early exercise premium \( \hat{e}_t \) at time \( t \), \( t_0 \leq t \leq T^* \), given in equation (3.20), requires computation of the terms \( \hat{B}_{t,s} \), \( \mu_{t,s} \) and \( \sigma_{t,s} \). In this section we will find formulas of these terms that will enable \( \hat{e}_t \) to be computed more easily and efficiently.

According to (4.1.6) with \( s \) replacing \( T \), the discount bond price \( \hat{B}_{t,s} \) for a discount bond maturing at times \( s \geq t \) has the form

\[
\hat{B}_{t,s} = \hat{B}(t,s,\hat{r}_t) = \exp\left\{-\hat{r}_t D(t,s) + \hat{v}_t \hat{F}(t,s) + \hat{G}(t,s)\right\},
\]

for \( t_0 \leq t \leq T \), where \( D(\cdot,\cdot) \), \( \hat{F}(\cdot,\cdot) \) and \( \hat{G}(\cdot,\cdot) \) are real valued functions that satisfy the ordinary differential equations

\[
\frac{dD(t,s)}{dt} = \alpha_t D(t,s) - 1, \\
\frac{d\hat{F}(t,s)}{dt} = \gamma \hat{F}(t,s) + \left(\lambda - \frac{1}{2} D(t,s)\right) D(t,s), \\
\frac{d\hat{G}(t,s)}{dt} = \alpha_t \hat{r}_t D(t,s) - \gamma \hat{v}_t \hat{F}(t,s)
\]

for \( t_0 \leq t \leq T \) with boundary conditions

\[
D(s,s) = \hat{F}(s,s) = \hat{G}(s,s) = 0
\]

and \( \hat{v}_t \) is as given in (3.2).

With this formulation we use the symbols \( \hat{F} \) and \( \hat{G} \) because these functions may be different from the corresponding functions \( F \) and \( G \) for the extended Fong and Vasicek model. As noted previously the function \( D \) will be the same for both models.

Consequently if we use the value \( s = T \), then the function \( \hat{H} \) as used in (3.15) is given by

\[
\hat{H}(t,T) = \hat{v}_t \hat{F}(t,T) + \hat{G}(t,T).
\]

The ordinary differential equations in (4.2) with boundary conditions (4.3) can be solved in integral form as follows:

\[
D(t,s) = \phi_{t,s} \int_t^s \phi_{u,s}^{-1} du,
\]
5.4. COMPUTATION OF TERMS - EXTENDED VASICEK MODEL

\[ \hat{F}(t, s) = -e^{\gamma t} \int_t^s e^{-\gamma u} \left[ \lambda - \frac{1}{2} D(u, s) \right] D(u, s) du, \]

\[ \hat{G}(t, s) = -\int_t^s \left[ \alpha_u \bar{r}_u D(u, s) - \gamma \bar{v} \hat{F}(u, s) \right] du, \]

for \( t_0 \leq t \leq s \leq T^* \), where \( \phi \) is as given in equation (3.12). From this expression for \( \phi \) together with the above integral solutions for the functions \( D \), the following representations for the functions \( \phi \) and \( D \) can be derived in a straightforward manner.

\[ \phi_{t, s} = \frac{\phi_{t, T^*}}{\phi_{s, T^*}} \]

\[ \phi_{t, s} = \frac{\phi_{t_0, s}}{\phi_{t_0, t}}, \]

\[ D(t, s) = D(t, T^*) - \frac{\phi_{t, T^*}}{\phi_{s, T^*}} D(s, T^*). \] (4.4)

This means that \( \phi \) and \( D \) considered as functions of two variables can be computed from the one variable functions \( \phi, T^* \) and \( D(\cdot, T^*) \).

Similarly the functions \( \hat{F} \) and \( \hat{G} \) can also be expressed as combinations of one variable functions. These calculations are somewhat lengthy so we will not include all of the details here. However to illustrate the first part of the procedure for the function \( \hat{F} \) we write \( \hat{F} \) as the sum \( \hat{F} = \hat{F}_1 + \hat{F}_2 \), where

\[ \hat{F}_1(t, s) = -e^{\gamma t} \int_t^s \lambda e^{-\gamma u} D(u, s) du \] (4.5)

for \( t_0 \leq t \leq s \leq T^* \). Define the one variable functions \( I_1 \) and \( I_2 : [t_0, T] \to \mathbb{R} \) by

\[ I_{1, t} = \int_t^{T^*} e^{-\gamma u} D(u, T^*) du \]

\[ I_{2, t} = \int_t^{T^*} e^{-\gamma u} \phi_u T^* du \]

for \( t_0 \leq t \leq T^* \). Using these functions, (4.4) and (4.5) we can express the function \( \hat{F}_1 \) in the form

\[ \hat{F}_1(t, s) = -e^{\gamma t} \int_t^s \lambda e^{-\gamma u} \left( D(u, T^*) - \frac{\phi_u T^*}{\phi_{s, T^*}} D(s, T^*) \right) du \]

\[ = \lambda e^{\gamma t} \frac{D(s, T^*)}{\phi_{s, T^*}} (I_{2, t} - I_{2, s}) - \lambda e^{\gamma t} (I_{1, t} - I_{1, s}) \]

for \( t_0 \leq t \leq s \leq T^* \). This means that the function \( \hat{F}_1 \) can be computed from one variable functions. This procedure can be continued for the functions \( \hat{F}_2 \) and \( \hat{G} \). Consequently the bond price function \( \hat{B}(\cdot, \cdot, \hat{r}_t) \) given by (4.1) can also be computed from combinations of one variable functions.
Let us now define functions $\psi$ and $\chi$ by the relations:

$$
\psi_{t,s} = \int_t^s \phi_{t,u}^{-1} \left( \alpha_u \tilde{r}_u + \hat{v}_u \lambda + \sqrt{\hat{v}_u \hat{\theta}^*_u} \right) \, du,
$$

$$
\chi_{t,s} = \int_t^s \frac{\hat{v}_u}{\phi_{t,u}^2} \, du,
$$

(4.6)

for $t_0 \leq t \leq s \leq T^*$, where the functions $\hat{v}$ and $\hat{\theta}^*$ are given by equations (3.2) and (3.10), respectively. From (3.10) and (4.1) we have

$$
\hat{\theta}^*_u = -\sqrt{\hat{v}_u} D(u,T^*),
$$

so that

$$
\psi_{t,s} = \int_t^s \phi_{t,u}^{-1} \left( \alpha_u \tilde{r}_u + \hat{v}_u (\lambda - D(u,T^*)) \right) \, du.
$$

(4.7)

This result, the definition of $\chi$ given in (4.6) and the second equation in (4.4) means that we can express the functions $\psi$ and $\chi$ in the form

$$
\psi_{t,s} = \phi_{t_0,t} \left( \psi_{t_0,s} - \psi_{t_0,t} \right)
$$

$$
\chi_{t,s} = \phi_{t_0,t}^2 \left( \chi_{t_0,s} - \chi_{t_0,t} \right).
$$

(4.8)

The mean and variance parameters $\mu_{t,s}$ and $\sigma_{t,s}^2$ defined by (3.13) and (3.14) can now be represented using the equations

$$
\mu_{t,s} = \phi_{t_0,s} \left( \frac{\tilde{r}_t}{\phi_{t_0,t}} + \psi_{t_0,s} - \psi_{t_0,t} \right)
$$

$$
\sigma_{t,s}^2 = \phi_{t_0,s}^2 \left( \chi_{t_0,s} - \chi_{t_0,t} \right).
$$

(4.9)

These results show that the terms $\mu_{t,s}$ and $\sigma_{t,s}^2$ can be evaluated at any point $(t,s)$, $t_0 \leq t \leq s \leq T^*$, using the one variable functions $\phi_{t_0,t}$, $\psi_{t_0,s}$, and $\chi_{t_0,t}$. We have already seen that the bond price $B_{t,s}$, $t_0 \leq t \leq s \leq T^*$ can also be conveniently computed from one variable functions. Consequently the early exercise premium $\hat{e}_{t,s}$ as given by (3.20) can be efficiently evaluated using these one variable functions.

To compute the critical exercise boundary function and from this American put values, these one variable functions would typically be pre-computed at discrete points and stored in one-dimensional arrays. Appropriate interpolation procedures would then be used to calculate values for $\hat{B}_{t,s}$, $\mu_{t,s}$ and $\sigma_{t,s}^2$, $t_0 \leq t \leq s \leq T^*$, at intermediate points, so that the integral in (3.20) can be estimated. These results are therefore of considerable practical value as they reduce both the memory and computational demands of the problem.

We remark finally that these calculations for $\hat{B}_{t,s}$, $t_0 \leq t \leq s \leq T^*$, also hold in the case of $B_{t,s}$, the discount bond price for the extended Fong and Vasicek model. That is $B_{t,s}$ can be conveniently computed from combinations of one variable functions.
5.5 American Pricing for the Extended Fong and Vasicek Model

As an example of a two-factor term structure model, we will now consider an extended version of the Fong & Vasicek (1991a,b) model as defined by equation (4.1.4). As noted previously we regard this model as an extension of the original version proposed by Fong and Vasicek because the back driving intensity $\alpha$ and short rate level $\bar{r}$ are both specified as time-dependent functions.

The early exercise premium $e_t$ for the extended two-factor Fong and Vasicek model given by (2.13) is more difficult to evaluate than that for the corresponding one-factor extended Vasicek model because the critical exercise function $f$ now represents a two-dimensional surface.

However, by using the values $\hat{e}_t$, obtained from the one-factor valuation formula (3.20) we can construct control variate estimates for both $e_t$ and the two-dimensional exercise boundary. The integral representation (3.20) can also be used with $B_{t,s}$ replacing $\tilde{B}_{t,s}$ in (4.7) as this choice provides the basis for a better control variate estimation of $e_t$, $t_0 \leq t \leq T^*$. As in the case for the extended Fong and Vasicek model these tasks need to be integrated as an estimate of the critical exercise function is required to compute the integral in (2.13) however these critical exercise function estimates themselves require approximations of the integral appearing in (2.1.3). A procedure similar to that described in the previous section, therefore needs to be employed for this two-factor model.

To be more explicit, if we write $e_t = \hat{e}_t + (e_t - \hat{e}_t)$, then we can use stochastic numerical methods of the type described in the previous chapter to estimate $e_t - \hat{e}_t$. The variance associated with estimates of $e_t - \hat{e}_t$ will in general be much less than those obtained from a direct simulation of $e_t$. To ensure that the computational loads for this problem do not explode we need to have a good starting approximation for the critical boundary function $f$. As a natural choice we can use $\tilde{f}$, the critical boundary function for the extended Vasicek model as a suitable starting point.

One method for obtaining better approximations for the critical exercise function $f$ is as follows: We divide the time interval $[t_0, T^*]$ into say three segments $[t_0, t_1]$, $[t_1, t_2]$ and $[t_2, T^*]$ with $t_0 < t_1 < t_2 < T^*$. At $t_2$ we estimate $e_{t_2}$ at four points on a square positioned over the curve that corresponds to the critical boundary for the one-factor model. By comparing the values $p_{t_2} + e_{t_2}$ with the corresponding intrinsic value, $K - B_{t_2}$, we can determine how close we are to the critical exercise boundary for the extended Fong and Vasicek model. We use four points here because the two state variables $B_{t_2}$, the price of the long bond at time $t_2$, and $B^*_t$, the price of the short bond at time $t_2$, provide two degrees of freedom in the model. Different values for $B_{t_2}$ and $B^*_t$ can be obtained by using different values for the short rate $r_{t_2}$ and its instantaneous variance $v_{t_2}$ as prescribed by the bond pricing formula (4.1.6). A similar procedure is then used at time points $t_1$ and $t_0$. Using these 12 estimates of $e_t$ at time...
$t_0$, $t_1$ and $t_2$ we can now construct a new estimate for the boundary function say $f^{(2)}$, where $f^{(1)}$ is specified directly from $f$.

This method can be repeated if necessary to obtain finer approximations to the boundary function $f$. However in general only one or two iterations of this procedure are needed because the American price is not so sensitive to small movements of the critical exercise boundary.

Although somewhat heuristic in its formulation simulation experiments have shown that this algorithm can be efficiently and effectively used to determine the two-dimensional critical exercise boundary for the extended Fong and Vasicek model. The method can be partially validated by evaluating American prices at intermediate points located on the estimated critical boundary. These prices can then be compared to the corresponding intrinsic values $K - B_t$ which should be the same.

Another very simple but effective control variate technique is as follows: Let $\hat{p}_{t_0}$ and $\hat{e}_{t_0}$ denote the price of a European put and the early exercise premium, respectively, at time $t_0$, for the extended Vasicek model and let $p_{t_0}$ be the corresponding price of a European put for the extended Fong and Vasicek model. This price can be estimated very accurately using the methods described in Chapter 4 and does not require any knowledge of the critical exercise boundary. Then for a wide class of parameter values a good estimate for $e_{t_0}$, the early exercise premium for the corresponding American put at time $t_0$, is

$$e_{t_0} = \hat{e}_{t_0} \left( \frac{p_{t_0}}{\hat{p}_{t_0}} \right).$$

The computation of American prices using this formula is very fast as it requires only one estimate of $\hat{e}_{t_0}$ using the exact representation (3.20) and one simulation estimate of $p_{t_0}$. The price $\hat{p}_{t_0}$ can be determined analytically.

### 5.6 Computational Results for American Pricing

In this section we will describe some simulation results that were obtained for the pricing of American puts for the extended Fong and Vasicek model given by (4.1.4). The valuation procedure used consisted of three main steps. Firstly, the calculation of American put values for the corresponding extended Vasicek or related models. Secondly, the evaluation of European puts for the extended Fong and Vasicek model. Thirdly, the approximation of prices for American puts for this model using results from the two preceding steps together with a heuristic control variate method as outlined in the previous section.

The main aim of this section is to highlight the computational power of the methods used rather than to explain in detail how these results were obtained and validated. Also, computation of hedge ratios for American puts was not undertaken in these simulation experiments. Note however that these hedge ratios can be computed either using finite difference approximations, for rough estimates, or the powerful exact representa-
5.6. COMPUTATIONAL RESULTS FOR AMERICAN PRICING

To reduce the systematic error, see Kloeden & Platen (1992), associated with the computation of European puts for the extended Fong and Vasicek model, we used the predictor-corrector method (4.3.1). The variance reduction techniques of control variates, as outlined in Section 4.3, and the antithetic variates and stratified sampling as described in Section 3.4, were also applied to reduce the corresponding statistical error.

For the predictor-corrector method (4.3.1) we used 16 discretization points with 20 batches each with 128 paths divided into 32 groups of 4 constructed from the antithetic variance reduction procedure described in Section 3.4. We also employed the technique of stratified sampling using two-point approximations for Gaussian increments together with the partitioning sets (3.4.2).

The computation of prices for American puts for the extended Fong and Vasicek model was based on the exact representation (3.20) using the results of Section 5.4. The critical exercise boundary was estimated from this representation using a backward iterative technique. Approximation of the integral appearing in (3.20) required extensive use of non-stochastic numerical and integration routines. For these simulation experiments we estimated this integral using Romberg extrapolation with a 128 x 256 point grid. However, this integral needs to be evaluated repeatedly to obtain the critical exercise boundary.

Using a 486, 33 MHz personal computer with 16 discretization points and 2560 (= 20 x 128) sample paths a relative statistical error of 0.3% at a 99% confidence level was obtained. Note that this error term applies only to the European component of the American price. For all of the results presented in this section, American option prices, using the three-step procedure mentioned above, were each computed in about 15 seconds.

For these simulation experiments we used the following default parameter values:

\[ T = 5.0 \text{ (years)}, \quad T^* = 0.6, \quad \lambda = 0.0, \quad \eta = 0.0, \quad \gamma = 0.2, \quad \hat{\sigma} = 0.04, \quad \xi = 0.04, \quad \rho = 0.0, \]

\[ K = 70.0 \text{ with initial values } T_0 = 0.022 \text{ and } v_0 = 0.04 \text{ at time } t_0 = 0. \]

Linear functions for the time-dependent values \( \alpha_t \) and \( \bar{r}_t \), \( t_0 \leq t \leq T^* \) were chosen as given in Section 4.5 with both the long and short bonds scaled to give a face value of 100 monetary units.

As has been previously explained, the critical exercise boundary for the extended Fong and Vasicek model is two-dimensional and takes the form \( B_t = f(t, B_t^*) \), \( t_0 \leq t \leq T^* \). Using equation (4.4.12) we can also replace critical values of the long bond price \( B_t \), with critical values of the short rate \( r_t \), at time \( t \), \( t_0 \leq t \leq T^* \).

Figures 5.6.1 and 5.6.2 show critical values of the long bond price \( B_t \) and critical values of the short rate \( r_t \) respectively, for different values of time to maturity and the short bond price \( B_t^* \) at time \( t \), \( t_0 \leq t \leq T^* \). We remark that to the best of our knowledge this is the first characterization of the critical exercise boundary for a two-dimensional American problem.
CHAPTER 5. VALUATION OF AMERICAN BOND OPTIONS

Figure 5.6.1: Critical values of the long bond price $B_t$.

Figure 5.6.2: Critical values of the short rate $r_t$. 
Figure 5.6.3 displays prices of American and European puts for different values of the long bond price $B_0$ at time $0$. We also plot the intrinsic value of the option to show the convergence of the American price to its intrinsic value as the critical exercise boundary is approached. Note that for points close to this boundary the main component in the price of an American put is its early exercise premium.

The effects of stochastic volatility are illustrated in Figures 5.6.4 and 5.6.5. The first of these diagrams shows price differences, $(u_0' - \tilde{u}_0')$, between American puts for the extended Fong and Vasicek and corresponding extended Vasicek models for different values of the long bond price $B_0$ at time $0$ and time to maturity $T^*$. The smile effect in prices is not symmetric partly because for low values of the long bond price the American prices for both models quickly approach the intrinsic value of the option. A different view of these stochastic volatility effects is shown in Figure 5.6.4, where we plot price differences between American puts for the above two models for different values of the parameter $\xi$ (volatility of volatility) and the long bond price $B_0$. This figure shows that the smile effect becomes less pronounced as the parameter $\xi$ approaches zero as is to be expected from (4.1.4).
CHAPTER 5. VALUATION OF AMERICAN BOND OPTIONS

Figure 5.6.4: Price differences due to stochastic volatility for different values of the long bond price $B_0$ and time to maturity $T^*$. 

Figure 5.6.5: Price differences due to stochastic volatility for different values of $\xi$ and the long bond price $B_0$. 
Bibliography


BIBLIOGRAPHY

Control Optimization 16, 252–269.


