Pricing and Hedging American Options Using Monte Carlo Simulation

Fredrik Fagerlund

September 2, 2007
Abstract

This thesis is devoted to pricing and hedging of American style options by the use of Monte Carlo simulation. We describe and implement numerous methods, developed for pricing American options by simulation. We show how Monte Carlo simulation can be used to achieve a plausible, and accurate, price approximation and illustrate this by numerical results. Both single asset, and multiple asset, contract structures have been applied to these methods. This study points out the strengths, and weaknesses, of Monte Carlo simulation when using it for pricing an American style option. Finally, we use Monte Carlo simulation to estimate the hedge ratios of American options and the result is illustrated in tables.
Acknowledgements

I would like to thank my supervisor, Associate Professor Kaj Nyström, for his guidance during this thesis work. I also would like to thank Karl Larsson for discussions and ideas. Finally, I would like to thank my family, Camilla and our daughter Isabelle, for all their love and support.
## Contents

1 Introduction ............................................. 5

2 Theory .................................................. 7
   2.1 The market, Portfolio and Arbitrage ................. 7
   2.2 Introduction to American Option Pricing .......... 10
   2.3 Models of Asset Dynamics .......................... 13

3 Pricing Methods .................................... 15
   3.1 Dynamic Programming Formulation .................. 15
   3.2 Binomial Tree Method in 1-D ....................... 16
   3.3 Binomial Tree Method in Higher Dimension ........ 18
   3.4 Random Tree Method ................................ 19
   3.5 Stochastic Mesh Method ............................ 22
   3.6 Regression Based Method ............................ 24
   3.7 Duality ............................................ 25

4 Estimating Sensitivities ............................ 27
   4.1 Finite Difference Approximations ................. 27
   4.2 The Hedge Parameter Delta ....................... 29
   4.3 The Hedge Parameter Vega ....................... 30

5 Numerical Results ................................ 32
   5.1 Single Asset American Options ................... 32
   5.2 Multi Asset American Options .................... 33
   5.3 Hedge Ratios ...................................... 35
   5.4 Comments on the numerical results .............. 35
   5.5 Suggestions for further studies .................. 37

6 Conclusions ........................................ 37

References ............................................. 39
1 Introduction

Options are financial derivatives, contracts, that give the holder the right to buy (or sell) an underlying asset, typically a stock, at a specific price and after (or during) a specific time horizon. There are two main types of options, European and American style options. While European style options can only be exercised at the maturity date, American options can be exercised at any time up to its expiration.

In 1973 a paper published by Fischer Black and Myron Scholes revolutionized the financial market. In the published paper they had outlined an analytic model that would determine the fair market value for European style call options paying no dividends. This model was later on to be called the Black-Scholes model for option pricing.

Ever since the Black-Scholes model was presented for the first time in 1973 the development of financial derivatives has grown rapidly. Today, the theory of option pricing is often considered among the most mathematically demanding of all applied areas of finance. In today’s financial market there are hundreds of different types of options which all differ in their payoff structures, path dependence or termination conditions.

American style options, are constructed such that the holder can choose the time of exercise. This privilege of early exercise come at a cost, known as the early exercise premium. It is the feature of early exercise that makes American option pricing more complex than it’s European style counterpart. The valuation of early exercise features remain a difficulty and is still an intensive area of research. The research covering American option pricing originate from the work done by Cox, Ross and Rubenstein(1979) who successfully developed a model for pricing one asset American options by the use of a binomial lattice. Convergence of the Cox, Ross and Rubenstein(1979) binomial method was proved fifteen years later by Amin and Khanna(1994). Boyle, Evenine and Gibbs(1989) extended the work done by Cox, Ross and Rubenstein(1979) to a multiple dimensional binomial lattice, making pricing of multiple asset options possible. Other approaches for pricing American options include finite difference approximations and finite elements methods.

The binomial tree method, the finite element method and the finite difference approximation can be used successfully to approximate the price of an American option with one or two underlying assets. Nevertheless, in many real-life finance applications, the number of underlying securities exceeds two - or even three. The high dimension, and the feature of early exercise, makes PDE methods, finite differential approximations and binomial tree methods inadequate. This is why the focus of research has turned to the attempt of pricing high

---

1A European style call option is an option where the holder has the right to buy the underlying asset at the maturity date. A European style put option is an option where the holder has the right of selling the underlying asset at maturity.
dimensional American options by the use of simulation. The research with main focus on American option pricing by the use of simulation covers research done by, Longstaff and Schwartz (2001), Brodie and Glasserman (1997) and Haugh and Kogan (2001), to only name a few. The high dimensionality is an attractive approach for MC, since the convergence rate of Monte Carlo simulation is independent of the number of underlying state variables. However, any method for pricing a high dimensional American option by simulation requires substantial computational effort, both storage capacity and processor speed.

For investors and option traders the price of the option is not the only thing of interest. In order to be able to measure the risk in their portfolio, knowledge about the sensitivity of the option and other derivatives in the market is valuable. These sensitivities are known as the option greeks. The professional market uses the greeks to measure exactly how much they need to hedge their portfolio and to measure how much risk their portfolio is exposed to.

Two main questions arise regarding this subject: How accurate is Monte Carlo simulation for estimating the price of an American option?
If simulation have been used to estimate the option price - Is it also possible to estimate the option greeks by simulation?

The objective of this thesis is to try to give answers to these questions. The thesis starts with explaining some basic theory connected to financial mathematics and, in particular, American option pricing. In chapter three, a number of proposed methods for American option pricing is treated. This section covers methods based on simulation and methods based on numerical computation. Each method description is based on the original research article(s) and the advantages and drawbacks are regarded. Chapter four considers estimation of hedge parameters. That section also starts with a brief background on how partial derivatives can be estimated by the use of simulation. Next, two commonly used hedge parameters are described. Finally, in chapter five, some numerical values based on implementation of the methods covered in chapter three and four are presented.
2 Theory

In this chapter some definitions and theory will be given and used as a background for the American option nature. It will cover definitions, theorems and assumptions, important for the understanding of both the problem of pricing and hedging an American option and the proposed solution. The following definitions used are found in Øksendal [14] and Detemple [8].

2.1 The market, Portfolio and Arbitrage

The economy, i.e the market, is represented as a complete probability space $(\Omega, \mathcal{F}, P)$, where $\Omega$ is the total set of events with generic elements $\omega$. $\mathcal{F}$ is the $\sigma$-algebra generated by the random variable (Brownian motion) $\{B(s)\}_{0 \leq s \leq t}$. One often thinks of $\mathcal{F}_t$ as the "the history of $B_s$ up to time t". $P$ is a probability measure defined on $(\Omega, \mathcal{F})$.

Definition 1 A (mathematical) market is an $\mathcal{F}_t$-adapted $(n+1)$-dimensional Itô process

$$X(t) = (X_0(t), X_1(t), \ldots, X_n(t)) \quad 0 \leq t \leq T$$

which is assumed to have the form

$$dX_0(t) = \rho(t, \omega)X_0(t)dt \quad X_0(0) = 1$$

and

$$\frac{dX_i(t)}{X_i(t)} = \mu_i(t, \omega)dt + \sum_{j=1}^{m} \sigma_{ij}(t, \omega)dB_j(t)$$

$$= \mu_i(t, \omega)dt + \sigma_i(t, \omega)dB(t)$$

$$X_i(0) = x_i$$

where $\sigma_i$ is the $i$:th row of the $n \times m$ matrix $[\sigma_{ij}]$; $1 \leq i \leq n \in N$.

From this definition it is evident that there are two main types of investment opportunities, one risky and one risk free. $X(t)$ can be seen as the total market consisting of assets or securities where $X_i(t)$ is asset $i$ at time $t$. The assets $X_1(t), \ldots, X_n(t)$ are risky assets because of the existence of a diffusion term and are governed by a geometric Brownian motion (a stochastic differential equation), while $X_0(t)$ is a risk free investment e.g a bond. Note also that the market $\{X(t)\}_{t \in [0,T]}$ is normalized if $X_0(t) \equiv 1$. This can be accomplished by defining
\[ X_i(t) = X_0(t)^{-1}X_i(t) := \xi(t)X_i(t) \quad 1 \leq i \leq n \]  

The market

\[ X(t) = (1, X_1(t), \ldots, X_n(t)) \]  

is then called the normalization of \( X(t) \).

**Definition 2** A portfolio in the market \( \{X(t)\}_{t \in [0,T]} \) is an \((n+1)\)-dimensional \((t, \omega)\) measurable and \(\mathcal{F}_t\)-adapted stochastic process

\[ \Theta(t, \omega) = (\Theta_0(t, \omega), \Theta_1(t, \omega), \ldots, \Theta_n(t, \omega)) \quad 0 \leq t \leq T \]

The elements \( \Theta_0(t, \omega), \Theta_1(t, \omega), \ldots, \Theta_n(t, \omega) \) represents the number of units (of an asset) an investor hold at time \( t \). The total collection of assets is called a portfolio.

**Definition 3** The value at time \( t \) of a portfolio \( \Theta(t) \) is defined by

\[ V(t, \omega) = V^\Theta(t, \omega) = \Theta(t) \cdot X(t) = \sum_{i=0}^{n} \Theta_i(t)X_i(t) \]  

where \( \cdot \) denotes the inner product in \( \mathbb{R}^{n+1} \)

This definition simply defines that the value of the portfolio is the total value of all investments held at time \( t \).

**Definition 4** The portfolio \( \Theta(t) \) is called self-financing if

\[ \int_0^T \left\{ \left| \Theta_0(s)\rho(s)X_0(s) + \sum_{i=1}^{n} \Theta_i(s)\mu_i(s) \right| + \sum_{j=1}^{m} \left[ \sum_{i=1}^{n} \Theta_i(s)\sigma_{ij}(s) \right]^2 \right\} ds < \infty \]  

and

\[ V(t) = V(0) + \int_0^t \Theta(s) \cdot dX(s) \quad t \in [0, T] \]

The concept of self financing portfolios is of great importance for defining what is meant by an arbitrage. To keep it simple, a self financing portfolio is a portfolio where the purchase of a new asset is financed by selling an existing asset in the portfolio.
A natural condition in real life finance is that there has to be a lower bound to how much debt the creditors can tolerate, or in other words, limitations on a portfolio. This argument leads to the following definition

**Definition 5** A self financing portfolio is called admissible if the corresponding value process \( V^\Theta(t) \) is lower bounded i.e there exists a \( L = L(\Theta) < \infty \) such that

\[
V^\Theta(t) \geq -L \quad (t, \omega) \in [0, T] \times \Omega
\]  

**Definition 6** An admissible portfolio is called an arbitrage in the market \( \{X(t)\}_{t \in [0, T]} \) if the value process \( V^\Theta(0) = 0 \) and \( V^\Theta(T) > 0 \) a.s and \( P(V^\Theta(T) > 0) > 0 \).

The portfolio is an arbitrage if there is an increase in value of the portfolio \( \Theta(t) \) form time \( t = 0 \) to \( t = T \). An arbitrage exists if the state of the market is not in equilibrium and, hence, there exists an opportunity to generate money without the risk of losing money.

**Theorem 1** Suppose the there exists a process \( u(t, \omega) \in \mathcal{V}^m(0, T) \) such that

\[
\sigma(t, \omega)u(t, \omega) = \mu(t, \omega) - \rho(t, \omega)X(t, \omega)
\]  

and such that

\[
E\left[\exp \left(\frac{1}{2} \int_0^T u^2(t, \omega)dt\right)\right] < \infty
\]  

Then the market \( \{X(t)\}_{t \in [0, T]} \) defined in definition 1 has no arbitrage.

For proof see Øksendal [14] page 268-269.

The next theorem originate from the theory of stochastic differential equations. It concerns change of probability measure and it is important for the construction of portfolios later on.

**Theorem 2** (The Girsanov Theorem 2)
Let \( Y(t) \in \mathbb{R}^n \) be an Itô process of the form

\[
dY(t) = \beta(t, \omega)dt + \Theta(t, \omega)dB(t); \quad t \leq T
\]  

where \( B(t) \in \mathbb{R}^m \), \( \beta(t, \omega) \in \mathbb{R}^n \) and \( \Theta(t, \omega) \in \mathbb{R}^{n \times m} \).

Suppose there exists processes \( u(t, \omega) \) and \( \alpha(t, \omega) \) such that

\[
\Theta(t, \omega)u(t, \omega) = \beta(t, \omega) - \alpha(t, \omega)
\]
\begin{align*}
M_t &= \exp \left( -\int_0^t u(s, \omega)dB_s - \frac{1}{2} \int_0^t u^2(s, \omega)ds \right); \quad t \leq T \quad (14) \\
dQ(\omega) &= M_t dP(\omega) \quad \text{on } \mathcal{F}_T \quad (15)
\end{align*}

Assume that $M_t$ is a martingale. Then $Q$ is a probability measure on $\mathcal{F}_T$. The process

$$
\tilde{B}(t) := \int_0^t u(s, \omega)ds + B(t) \quad (16)
$$

is a Brownian motion w.r.t $Q$ and in terms of $\tilde{B}(t)$ the process $Y(t)$ has the stochastic integral representation

$$
dY(t) = \alpha(t, \omega)dt + \Theta(t, \omega)d\tilde{B}(t) \quad (17)
$$

For proof see Øksendal [14] page 165.

\section*{2.2 Introduction to American Option Pricing}

The American style contingent claim differs, as mentioned before, from the European style contingent claim because of the option holders privilege of early exercise. An important element in American option theory is optimal stopping time. A stopping time, in general, is defined as

\textbf{Definition 7} Let $\{\mathcal{F}_t\}$ be an increasing family of $\sigma$-algebras. A function $\tau : \Omega \to [0, \infty]$ is called a stopping time with respect to $\{\mathcal{F}_t\}$ if

$$
\{\omega; \tau(\omega) \leq t\} \in \mathcal{F}_t \quad \forall t \geq 0 \quad (18)
$$

So, the information given by the filtration $\mathcal{F}_t$ is enough for deciding whether $\tau \leq t$ has occurred or not.

The following definition defines what is meant by an American style option.

\textbf{Definition 8} An American $T$-claim is an $\mathcal{F}_t$-adapted $(t, \omega)$ measurable stochastic process $F(t, \omega); t \in [0, T], \omega \in \Omega$. An American option on such a claim gives the holder the right (but not the obligation) to choose any stopping time $\tau(\omega) \leq T$ as exercise time for the option resulting in a payoff $F(\tau(\omega), \omega)$.

\textbf{Definition 9} An American $T$-claim, $F(t, \omega)$, is called attainable in the market $\{X(t)\}_{0 \leq t \leq T}$ if there exists an admissible portfolio $\Theta(t)$ and a real number $z$ such that

$$
F(\omega, t) = V^\Theta_z(T) := z + \int_0^t \Theta(s)dX(s) \quad a.s \quad (19)
$$
and such that
\[ V_{\theta}^z(\tau) := z + \int_0^\tau \xi(s) \sum_{i=1}^n \Theta_i(s)\sigma_i(s)d\tilde{B}(s); \quad 0 \leq \tau \leq T \] (20)
is a \( Q \) martingale. If such a portfolio exists, we call it a replicating or hedging portfolio for \( F \).

In the mathematical financial setting, it is the properties, or better, the assumed properties of the market that makes fair pricing of derivative securities possible. The next definition and theorem, we first define what is meant by a complete market, and secondly we state an important condition on the market in order for it to be complete.

**Definition 10** The market \( \{X(t)\} \) is called complete if every \( T \)-claim \( F(t,\omega) \) is attainable.

**Theorem 3** The market \( \{X(t)\} \) is complete if and only if \( \sigma(t,\omega) \) has a left inverse \( A(t,\omega) \) for almost any \( (t,\omega) \), i.e. there exists an \( \mathcal{F}_t \)-adapted matrix valued process \( A(t,\omega) \in \mathbb{R}^{m \times n} \) such that
\[ A(t,\omega)\sigma(t,\omega) = I_m \quad \text{for a.a}(t,\omega) \] (21)
For proof see Øksendal [14] page 275-276.

Up to this point we have the given restrictions on both the market and the portfolio to make it possible to state the fair price of an American option. First, a definition of the common price and then the pricing formula for the American option can be given.

**Notation 1** The price the buyer is willing to pay for an American \( T \)-claim, \( F \), is denoted by \( p_A(F) \). The price the seller of the same American \( T \)-claim, \( F \), is willing to accept is denoted \( q_A(F) \).

**Definition 11** Suppose that the price, the seller and the buyer, of an American \( T \)-claim is willing to accept is denoted as in notation 1. If
\[ p_A(F) = q_A(F) \] (22)
Then we will call this price the common price of the American option.

**Theorem 4** Suppose that the market \( \{X(t)\} \) is complete and that the conditions of no-arbitrage in the market is fulfilled. Suppose that the sellers price
and the buyers price of the American T-claim, $F$, is denoted as in notation 1. Define the measure $Q$ on $\mathcal{F}_t$ by

$$
dQ(\omega) = \exp\left(-\int_0^T u(t,\omega)dB(t) - \frac{1}{2} \int_0^T u^2(t,\omega)dt\right) dP(\omega)
$$

(23)

Then there exists a unique price of the American option and it is given by

$$
p_A(F) = \sup_{\tau \leq T} E_Q[\xi(\tau)F(\tau(\omega),\omega)] = q_A(F)
$$

(24)


Thus, the fair price of the American option can be determined by exercising optimally.

Given that the market is defined as a complete probability space and that the environment (i.e. the underlying asset) is Markovian, then the information required for pricing an American option is contained in the two sets called Immediate exercise region and Continuation region.

**Definition 12** The immediate exercise region, denoted $\mathcal{E}$ is a finite set of values at which the optimal policy is to exercise the option immediately:

$$
\mathcal{E} = \{(t,X) \in \mathbb{R}_+^n \times [0,T] : C(t,X) = F(t,X)\}
$$

(25)

where $C(t,X)$ denotes the price of the option and $F(t,X)$ denotes the payoff.

**Definition 13** The continuation region denoted $\mathcal{C}$ is the finite set of values at which immediate exercise is suboptimal (the complement set of the immediate exercise region).

$$
\mathcal{C} = \{(t,X) \in \mathbb{R}_+^n \times [0,T] : C(t,X) > F(t,X)\}
$$

(26)

where $C(t,X)$ denotes the price of the option and $F(t,X)$ denotes the payoff.

The boundary connecting the immediate exercise region and the continuation region is another important element in the theory of American options. The following proposition states some properties regarding the boundary for the one-dimensional case.

**Proposition 1** The boundary $b^*$ of the immediate exercise region in (25) is continuous on $[0,T)$, non increasing with respect to time and has the limiting values

$$
\lim_{t \to T} b^*_t = \max \left\{K, \left(\frac{\beta}{\delta}\right)K\right\} \text{ and }
\lim_{T-t \to \infty} b^*_t = b^*_\infty \equiv \frac{K(\beta + f)}{\beta + f - \sigma^2}
$$

(27)
where $\beta \equiv \delta - r + \frac{1}{2}\sigma^2$ and $f \equiv (\beta^2 + 2r\sigma^2)^{1/2}$. At maturity $b^*_T = K$. $r$ is the interest rate, $\delta$ the dividend yield and $\sigma$ the volatility.

The following proposition states some properties of the special case of the American style call option.

**Proposition 2** Let $C(t, X)$ be the value of the American style call option. Then

(i) $C(t, X)$ is continuous on $\mathbb{R}_+ \times [0, T]$.

(ii) $C(t, X)$ is non-decreasing and convex on $\mathbb{R}_+$ for all $t \in [0, T]$.

(iii) $C(\cdot, X)$ is non-increasing on $[0, T]$ for all $X \in \mathbb{R}_+$.

(iv) $0 \leq \frac{\partial C}{\partial X} \leq 1$ on $\mathbb{R}_+ \times [0, T]$.

(v) $\frac{\partial C(t, X)}{\partial X} = 1$ for $(t, X) \in E^0$

where $E^0$ denotes the interior of the exercise region

### 2.3 Models of Asset Dynamics

Options are governed by the so called Black-Scholes model. The Black-Scholes model is a model of the evolving price of financial instruments, in particular stocks. The key to option pricing in the Black-Scholes market rely on a few basic assumptions, namely

- The price of the underlying instrument $X_i(t)$ follows a geometric Brownian motion with constant drift $r_i$ and volatility $\sigma_i$:

$$dX_i(t) = r_iX_i(t)dt + \sigma_iX_i(t)dB_i(t); \quad 1 \leq i \leq n$$

(28)

where $B_i(t)$ is $n$-dimensional Brownian motion.

- There are no arbitrage opportunities.

- Trading in the stock is continuous.

- There are no transaction costs or taxes.

- A constant risk-free interest rate exists and is the same for all maturity dates.

The standard model for the underlying asset movement can easily be simulated by
\[ X(t_{k+1}) = X(t_k) \exp \left( \left( r - \frac{1}{2} \sigma_i^2 \right) (t_{k+1} - t_k) + \sigma_i \sqrt{t_{k+1} - t_k} Z_{k+1} \right) \] (29)

with \( Z_1, Z_2, \ldots, Z_n \) are independent standard normals.

The distribution of the expression in (29) is log-normal. An extension is the multivariate log-normal distribution, i.e when the contract structure is dependent upon several correlated underlying assets. The model for asset movements can then be simulated using

\[ X_i(t_{k+1}) = X_i(t_k) \exp \left( \left( r_i - \frac{1}{2} \sigma_i^2 \right) (t_{k+1} - t_k) + \sigma_i \sqrt{t_{k+1} - t_k} Y_i \right) \]

where each \( Y_i \) is a one dimensional standard normal and where \( Y_i \) and \( Y_j \) have correlation term \( \rho_{ij} \).

Generation of samples from a multivariate log-normal distribution can be accomplished by the following

\[ X_i(t_{k+1}) = X_i(t_k) \exp \left( \left( r_i - \frac{1}{2} \sigma_i^2 \right) (t_{k+1} - t_k) + \sigma_i \sqrt{t_{k+1} - t_k} \sum_{j=1}^{d} A_{ij} Z_{k+1,j} \right) \]

with \( Z_k = (Z_{1k}, Z_{2k}, \ldots, Z_{dk}) \in N(0, I), \ i = 1, \ldots, d, \ k = 0, \ldots, n - 1 \) and where \( Z_1, Z_2, \ldots, Z_n \) independent normals.

\( A_{ij} \) is the Cholesky factor of the covariance matrix \( \Sigma \) in the multivariate normal distribution \( N(\mu, \Sigma) \).

In the original Black-Scholes model of the evolving price of the asset, the volatility parameter \( \sigma \) is constant. But there other models for the asset movement, such as stochastic volatility models, where the volatility is modeled as a random variable - a stochastic process. The aim for stochastic volatility models is to try and capture the empirical observation that the volatility appears to act, not as a constant, but rather as a stochastic process itself. If the volatility is modeled as a stochastic process the underlying asset would be governed by

\[ dX_t = rX_t dt + X_t \sigma(Y_t) dB_t \] (30)

where \( \sigma(Y_t) \) is an additional stochastic process, in most cases a mean reverting stochastic process, e.g the Heston model\(^2\)

\(^2\)See Heston(1993) in the reference section
3 Pricing Methods

The pricing of American options is a demanding task due to the feature of possible early exercise. Several methods have been proposed as price approximations. Among the major ones are integral equations, PDE approaches, finite difference approaches, variational inequalities, analytic approximations and Monte Carlo simulation. In this chapter the focus will be on the latter of the methods mentioned. The chapter starts with explaining the common procedure of pricing an American option by simulation - the dynamic programming formulation. Even though this method is not invented for the specific purpose of pricing an American option, and of course is applicable for any other problem where optimal stopping problems occur, it is of such great importance in the nature of American option simulation that it deserves to be treated individually. The second and third methods, the binomial tree method in one or several dimensions, is a deterministic method and should be viewed only as a benchmark model for the rest of the methods covered in this chapter. The four remaining methods are based on Monte Carlo simulation. Each method is treated with the background theory and assumptions about the market covered in section 2.

The notation is based on the notation in section 2, i.e

\[ X_t = \text{State variable of the underlying asset, where } X_0 = x \]
\[ K = \text{Strike price} \]
\[ r = \text{Risk free interest rate} \]
\[ \sigma = \text{Constant volatility of the underlying asset} \]
\[ T = \text{Time to maturity(yrs)} \]
\[ \delta = \text{Continuous dividend yield} \]

3.1 Dynamic Programming Formulation

Most methods for pricing an American option relies on the dynamic programming method. The method uses a backward induction principle to estimate future values and can be described by the following:

Let \( h_i(x) \) denote the discounted payoff for exercise at time \( t_i \) and let \( V_i(x) \) denote the the value of the option, given \( X_i = x \), assuming that the option has not previously been exercised. We are interested in \( V_0(X_0) \). The dynamic programming procedure is determined by

\[
\begin{align*}
V_m(x) &= h_m(x) \\
V_{i-1} &= \max \{ h_{i-1}(x), E[V_i(X_i)|X_{i-1} = x] \} \\
&\text{for } i = m, \ldots, 1
\end{align*}
\]
The conditional expectation in (31) is called the continuation value. When pricing an American option and, hence, solving the optimal stopping problem one must determine if the payoff at that particular time step is larger or smaller than holding the option up to the next time step. The continuation value is simply the predicted value of the option at the next time step. The dynamic programming procedure can then be seen as "Taking the largest value at every time step between the payoff and the continuation value".

The dynamic programming recursion in (31) focuses on option values but there is also a stopping rule (defined in definition 7) which, in this setting, is described by the first time the Markov chain (underlying asset) hits the boundary, or in other words enters the exercise region. The \( \tilde{\tau} \) is defined by

\[
\tilde{\tau} = \min \{ i \in 1, \ldots, m : h_i(X_i) \geq V_i(X_i) \} \tag{32}
\]

The American option feature is given by the following expression, where the value of the option can be determined by

\[
V_i(x) = \max(h_i(x), E[V_{i+1}(X_{i+1})|X_i = x]) = \max(h_i(x), C_i(x)) \tag{33}
\]

So, the value of the option \( V_0(X_0) \) will determine the price of the option and, hence

\[
V_0^\tau(X_0) = E[h_\tau(X_\tau)] \tag{34}
\]

### 3.2 Binomial Tree Method in 1-D

This method proposed by Cox, Ross and Rubinstein [7] assumes that the price of the underlying asset follows a binomial process. The movement in the underlying asset can then be modeled in the recombining binomial lattice by the parameters \( u \) and \( d \), which corresponds to "up" and "down" movement, respectively.

The up and down factors are calculated using the underlying volatility \( \sigma \) and the time duration of a step \( dt \), i.e,

\[
u = e^{\sigma \sqrt{dt}} \quad \text{and} \quad d = e^{-\sigma \sqrt{dt}} = \frac{1}{u} \tag{35}\]

The probabilities for up and down movement is determined by the relation:

\[
p = \frac{e^{(r-\delta)dt} - d}{u - d} \tag{36}\]
where $\delta$ is the dividend yield, $r$ is the risk free interest rate, $u$ and $d$ is given by the expression above.

So, an "up" movement in the underlying has a probability $p$ and a "down" movement has probability $(1 - p)$. The following procedure is then simply the usual dynamic programming formulation where the value of the option now can be determined by taking the maximum of the payoff at time $i$ and the continuation value at time $i$, i.e

$$V_i(x) = \max(V_{\text{bin}}(x), h_i(x))$$ (37)

where $V_{\text{bin}}(x) = e^{-rdt}(pu + (1 - p)d)$ is the discounted continuation value of the option and where $V_u$ and $V_d$ corresponds to "option up" and "option down", respectively.

Table 1 shows numerical results of an American option priced by the binomial tree method. The table shows the fast convergence rate of the binomial tree method. The CPU of each value using 10,000 time steps is less then 10 seconds.
Table 1: Standard American put option priced by the binomial tree method.

<table>
<thead>
<tr>
<th>K</th>
<th>σ</th>
<th>T (yr)</th>
<th>Binomial (N=10,000)</th>
<th>Binomial (N=5,000)</th>
<th>Binomial (N=1,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>0.2</td>
<td>1</td>
<td>2.4724</td>
<td>2.4725</td>
<td>2.4731</td>
</tr>
<tr>
<td>90</td>
<td>0.3</td>
<td>1</td>
<td>5.5517</td>
<td>5.5519</td>
<td>5.5537</td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>1</td>
<td>8.8920</td>
<td>8.8924</td>
<td>8.8926</td>
</tr>
<tr>
<td>100</td>
<td>0.3</td>
<td>1</td>
<td>6.0903</td>
<td>6.0902</td>
<td>6.0896</td>
</tr>
<tr>
<td>100</td>
<td>0.4</td>
<td>1</td>
<td>9.8699</td>
<td>9.8698</td>
<td>9.8687</td>
</tr>
<tr>
<td>110</td>
<td>0.2</td>
<td>1</td>
<td>13.6674</td>
<td>13.6672</td>
<td>13.6656</td>
</tr>
<tr>
<td>110</td>
<td>0.3</td>
<td>1</td>
<td>11.9728</td>
<td>11.9728</td>
<td>11.9738</td>
</tr>
<tr>
<td>110</td>
<td>0.4</td>
<td>1</td>
<td>15.6177</td>
<td>15.6180</td>
<td>15.6167</td>
</tr>
</tbody>
</table>

Columns (1)-(3) represent the parameter values, K (Strike price), σ (Volatility) and T (Time to maturity). Columns (4)-(6) represent numerical results of option prices by the binomial tree method where N is the number of time steps. Initial price $X_0 = 100$, dividend yield $\delta = 0\%$ and interest rate $r = 5\%$.

3.3 Binomial Tree Method in Higher Dimension

An extension of the 1-D binomial tree is the multivariate binomial tree method proposed by Boyle, Evenine and Gibbs [2]. In analogy with Cox, Ross and Rubenstein the method uses a recombining binomial lattice but in this case the aim is to price a contingent claim written on several underlying assets. To demonstrate the method, BEG considers the two dimensional case (contract written on 2 underlying assets):

In contrast to the 1-D binomial tree, where the nature of "asset jumps" is "up" or "down", the 2-D case has four pairs of possible values with four corresponding probabilities (see Table 2).

Table 2: The 2-D binomial lattice and its returns

<table>
<thead>
<tr>
<th>Nature of jumps</th>
<th>Probability</th>
<th>Asset prices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up,up</td>
<td>$p_1 = p_{uu}$</td>
<td>$X_1u_1, X_2u_2$</td>
</tr>
<tr>
<td>Up,down</td>
<td>$p_2 = p_{ud}$</td>
<td>$X_1u_1, X_2d_2$</td>
</tr>
<tr>
<td>Down,up</td>
<td>$p_3 = p_{dd}$</td>
<td>$X_1d_1, X_2u_2$</td>
</tr>
<tr>
<td>Down,down</td>
<td>$p_4 = p_{du}$</td>
<td>$X_1d_1, X_2d_2$</td>
</tr>
</tbody>
</table>

Table 2 shows the possible values of the 2-D binomial tree model after one time step. Middle and right columns show the corresponding probabilities and asset prices, respectively.

The risk neutral probabilities $p_i, i = 1, \ldots, 4$ are in the 2-D case given by
\[
\begin{align*}
    p_1 &= \frac{1}{4} \left( 1 + \rho + \sqrt{dt} \left( \frac{\mu_1}{\sigma_1} + \frac{\mu_2}{\sigma_2} \right) \right) \\
    p_2 &= \frac{1}{4} \left( 1 - \rho + \sqrt{dt} \left( \frac{\mu_1}{\sigma_1} - \frac{\mu_2}{\sigma_2} \right) \right) \\
    p_3 &= \frac{1}{4} \left( 1 - \rho + \sqrt{dt} \left( -\frac{\mu_1}{\sigma_1} + \frac{\mu_2}{\sigma_2} \right) \right) \\
    p_4 &= \frac{1}{4} \left( 1 + \rho - \sqrt{dt} \left( \frac{\mu_1}{\sigma_1} + \frac{\mu_2}{\sigma_2} \right) \right)
\end{align*}
\] (38)

where \( \mu_i \) is the drift term of the continuous log normal distribution and is given by \( \mu_i = r - \delta - \frac{1}{2} \sigma_i^2 \) for \( i = 1, 2 \) and the \( \rho \) denotes the correlation between the two underlying assets.

All of the probabilities given in (38) is nonnegative as long as the time step \( dt \) is chosen sufficiently small.

An evident limitation in the high dimensional binomial tree is storage requirements of terminal payoff values. Storing all nodes requires order \( m^n \) storage, where \( m \) is the number of time steps and \( n \) is the number of underlying assets.

### 3.4 Random Tree Method

One major disadvantage of the binomial tree method is the fact that the lattice is recombining. For path dependent contracts, like Asian-American options (sometimes called Hawai’i options), the use of a recombining tree in order to price such a contract is inadequate.\(^3\) The storage requirements of nodes in the high dimensional binomial tree also make the method unsuitable for high dimensional options. One way of dealing with path dependence and high dimensionality is the Random tree method, see figure 2, proposed by Brodie and Glasserman [3]. The random tree is constructed such that the lattice nodes are sampled randomly. This should be compared with the binomial tree, where lattice nodes are deterministic. The random structure of the tree makes this method well suited for multiple asset contracts with a small number of exercise opportunities. This, however, is also the main drawback of this method since the computational requirements grow exponentially as the number of exercise dates \( m \) increases.

One of the advantages on the other hand is that the random tree method produces two price estimators, one biased high and one biased low. A combination of these estimators can then be used to obtain valid confidence intervals for the probable price of the option.

\(^3\) Chalasani et al (1999) have developed a refined binomial lattice for the valuation of American-Asian options based on the Cox, Ross and Rubinstein binomial lattice.
The first step is generating the generic nodes by simulation of the underlying Markov chain, i.e.

\[ dX_t = rX_t dt + \sigma X_t dB_t \]  

(39)

Because of the large number of generic nodes needed to be sampled in order to build the tree structure, the simulation requires substantial computational capacity. A way to avoid this is by using a deep-first processing of the tree. In the deep-first algorithm, we follow a single branch at a time rather than generating them all simultaneously. This algorithm will minimize the storage requirements of the nodes. For a detailed pseudo code of the deep-first algorithm see Brodie and Glasserman [3] Appendix C.

The distinction between the two estimators (high and low) is that the high estimator uses the dynamic programming formulation described in section 3.1 and focus on option values. The value of the option at each node can be computed by

\[ \hat{V}_{i}^{j_1 \ldots j_i} = \max \left\{ h_i(X_i^{j_1 \ldots j_i}), \frac{1}{b} \sum_{j=1}^{b} \hat{V}_{i+1}^{j_1 \ldots j_i} \right\} \]  

(40)

where \( h_i(X_i^{j_1 \ldots j_i}) \) denotes the payoff at time \( i \), node \( j_1 \ldots j_i \) and \( \hat{C}_i = \frac{1}{b} \sum_{j=1}^{b} \hat{V}_{i+1}^{j_1 \ldots j_i} \) is the continuation value. The values \( \hat{V}_{i}^{j_1 \ldots j_i} \) are values of the high estimator at node \( X_i^{j_1 \ldots j_i} \).
Like mentioned in section 3.1 the target is $\hat{V}_0$. Let $\bar{V}_0(n, b)$ denote the sample mean of the $n$ times replicated random tree with branching parameter $b$. It is now possible to create a asymptotically valid confidence interval for the value $\bar{V}_0(n, b)$ by the central limit theorem

$$\bar{V}_0(n, b) \pm \frac{s_V(n, b)}{\sqrt{n}}$$

where $s_v(n, b)$ denotes the sample standard deviation. The interval (41) is a valid $1 - \delta$ confidence interval for $E[\hat{V}_0]$.

A way to remove the high bias in the high estimator is to separate the exercise decision value from the continuation value. Brodie and Glasserman splits the values of $\hat{V}$ into two disjoint subsets. To make this reasoning concrete consider the problem of estimating

$$\max \{h, E[Y]\}$$

where $h$ is a constant, e.g. the payoff at a specific node and time step and where $E[Y]$ would correspond to the continuation value.

In the next step let $\bar{Y}$ be the sample mean of $Y_i$, $i = 1, \ldots, b$. Divide this set of sampled values into two disjoint sets $Y_1$ and $Y_2$ and calculate their mean values, $\bar{Y}_1$ and $\bar{Y}_2$. Brodie and Glasserman chooses to use all but one of the $Y_i$ to calculate $\bar{Y}_1$ and the remaining one for $\bar{Y}_2$. The value of the option may now be computed via

$$\hat{v} = \begin{cases} h & \text{if } \bar{Y}_1 \leq h \\ \bar{Y}_2 & \text{otherwise} \end{cases}$$

By taking the expected value of this estimate it can be shown that the estimator is in fact biased low, i.e

$$E[\hat{v}] = P(\bar{Y}_1 \leq h)h + (1 - P(\bar{Y}_1 \leq h))E[Y] \leq \max(h, E[Y])$$

As in the programming formulation, all terminal nodes are set so that the estimator equals the payoff

$$\hat{v}^{j_1j_2\ldots j_m}_m = h_m \left( X^{j_1j_2\ldots j_m}_m \right)$$

By recursion, the low estimator is then defined as follows
\[ \hat{v}_{ik}^{j_1 j_2 \cdots j_i} = \begin{cases} h(X_i^{j_1 j_2 \cdots j_i}) & \text{if } \frac{1}{b} \sum_{j=1, j \neq k}^{b} \hat{v}_{i+1}^{j_1 j_2 \cdots j_i} \leq h(X_i^{j_1 j_2 \cdots j_i}) \\ \hat{v}_{ik}^{j_1 j_2 \cdots j_i} & \text{otherwise} \end{cases} \] (46)

As for the high biased estimator it is now possible to simulate \( n \) replications of the random tree and establish a valid confidence interval of the low estimator
\[ \overline{v}_0(n,b) = \hat{v}_0 \pm \frac{s_v(n,b)}{\sqrt{n}} \sqrt{n} \] (47)

By taking the upper limit of the high estimator and the lower limit of the low estimator it is possible to create an interval which will secure that the true value of the option price is contained in this interval.

### 3.5 Stochastic Mesh Method

The exponential growth of computational effort in the random tree framework gave rise to this method. Brodie and Glasserman [4] use a mesh rather than a tree in order to avoid the obstacle of increasing computation time. Like in the random tree method, the stochastic mesh method is based on generating randomly sampled nodes. The main distinction between these two methods is that in the stochastic mesh method the mesh uses values from all nodes at time step \( i+1 \) for valuing the option at time step \( i \). This is not the case in the random tree method where only successors of the current node is utilized. Since the number of nodes at each time step is fixed the difficulty of exponential growth in computational requirement is avoided. This makes the method applicable for a large number of contract structures and captures features ranging from the standard American option to more complicated structures like basket options and contracts with a large number of exercise dates.

The algorithm for option price estimation starts by generating mesh nodes, \( X_{ij} \), of the Markov chain where \( ij \) denotes the \( j \)-th node at time step \( i \). The option value is computed using the dynamic programming formulation described in section 3.1 and is given by
\[ \hat{V}_{ij} = \max \left\{ h_i(X_{ij}), \frac{1}{b} \sum_{k=1}^{b} W_{jk} \hat{V}_{i+1,k} \right\} \] (48)

As mentioned, the advantage of the stochastic mesh method is that the continuation value at each time step \( i \) uses information from all nodes in the following time step \( i+1 \). Since all nodes at time step \( i \) interconnect with those at time step \( i+1 \) the continuation value can be computed by
\[ C_i(X_{ij}) = \frac{1}{b} \sum_{k=1}^{b} W_{jk}^i \hat{V}_{i+1,k} \]  

(49)

where \( W_{jk}^i \) denotes some set of weights. At the root node the value is

\[ \hat{V}_0 = \frac{1}{b} \sum_{k=1}^{b} \hat{V}_{1k} \]  

(50)

or the maximum of \( \hat{V}_0 \) and \( h_0(X_0) \) is exercise at time 0 should be allowed.

Figure 3: Stochastic mesh illustrated in 1-D.

Each weight \( W_{jk}^i \) is a deterministic function of \( X_i \) and \( X_{i+1} \) and is set to

\[ W_{jk}^i = \frac{f_{i+1}(X_{ij}, X_{i+1,k})}{g(X_{i+1,k})} \]  

(51)

where \( f_{i+1}(X_{ij}, X_{i+1,k}) \) is the transition density from one generic node \( j \) at time step \( i \) to another generic node \( k \) at time step \( i + 1 \) and is has the general form

\[ f(x, y) = \prod_{i=1}^{n} \frac{1}{\sigma \sqrt{\Delta t}} \phi \left( \frac{y_i - x_i - (r - \delta - \frac{1}{2} \sigma^2) \Delta t}{\sigma \sqrt{\Delta t}} \right) \]  

(52)

with \( \phi \) the standard normal density and \( n \) denoting the dimension, i.e \( x = (x_1, x_2, \ldots, x_n) \) and \( y = (y_1, y_2, \ldots, y_n) \).
The function $g(X_{i+1,k})$ is the mean transition density of the successor nodes, meaning that $g(X_{i+1,k})$ is the average transition density of the generic nodes at the following time step

$$g(X_{i+1,k}) = \frac{1}{b} \sum_{i=1}^{b} f_{i+1}(X_{it}, X_{i+1,k})$$

(53)

The weights are constructed in such a way that by fixating $k$ and sum over $j$ the average into a node is 1, i.e

$$\frac{1}{b} \sum_{j=1}^{b} W_{jk} = 1$$

(54)

As in the random tree method, Brodie and Glasserman creates conditions which makes it possible to create a low biased estimator based on the option values evaluated in the high biased estimator. The low biased estimator (called path estimator) is defined by simulating a trajectory of the underlying asset process until the exercise region determined by the mesh is reached. So, given the values computed in the high estimator an exercise policy is determined. By simulating a trajectory independent of the mesh nodes the continuation value can be estimated by

$$\hat{C}_i = e^{-rdt} \frac{1}{b} \sum_{i=1}^{b} W_{low} \hat{V}_{i+1,k}$$

(55)

where $W_{low}$ is the set of weights given by the transition density connecting the simulated trajectory against the mesh nodes (compare with the expression in (52)).

### 3.6 Regression Based Method

The main difficulty in pricing an American option is estimating the continuation value at every time step $t_i$, $i = 1, \ldots, n$. This method, proposed by Longstaff and Schwartz [12] and Carriere [5] uses regression to estimate the continuation value. Each continuation value $C_i(x)$ is approximated by a linear combination of known functions of the current state and then by the use of regression (least squares) the best coefficients that fits the values $V_{i+1}(X_{i+1})$ can be approximated. The continuation value is approximated by

$$E[V_{i+1}(X_{i+1})|X_i = x] = \sum_{r=1}^{M} \beta_r \psi_r(x)$$

(56)
where $\psi$ is some basis function and $\beta$ is some constant in $\mathbb{R}$.

The choice of basis functions has a large influence on the price estimates. The table below displays how the choice of basis functions may affect the price estimates of an American max option when priced using regression.

### Table 3: American Max-option priced using regression.

<table>
<thead>
<tr>
<th>Basis function</th>
<th>Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1, X_1, X_2, X_3i$</td>
<td>15.74</td>
</tr>
<tr>
<td>$1, X_1, X_2^3, X_3i, X_1X_2$</td>
<td>15.24</td>
</tr>
<tr>
<td>$1, X_1, X_2^3, X_3i, X_1X_2, \max{X_1, X_2}$</td>
<td>15.23</td>
</tr>
<tr>
<td>$1, X_1, X_2^3, X_3i, X_1X_2, X_1X_2^2$</td>
<td>15.07</td>
</tr>
<tr>
<td>$1, X_1, X_2^3, X_3i, X_1X_2, X_2^2X_2, X_1X_2^2, h(X_1, X_2)$</td>
<td>14.06</td>
</tr>
<tr>
<td>$1, X_1, X_2^3, X_1X_2, \tilde{h}(X_1, X_2)$</td>
<td>14.08</td>
</tr>
</tbody>
</table>

Table 3 shows the affect basis functions has on the price estimate. Parameter values are:

- Initial price $X^0_1 = X^0_2 = 100$
- Interest rate $r = 0.05$
- Volatility $\sigma = 0.2$
- Strike price $K = 100$
- Dividend yield $\delta = 0.1$
- Time to maturity $T = 3$

The option can be exercised at nine equally spaced dates. The true price is 13.90 and is valued in the two dimensional binomial tree lattice.

Using the notation $\hat{C}_i$ for the continuation value we may write (56) as

$$
\hat{C}_i = \beta^T \psi(x)
$$

with $\beta^T = (\beta_1, \beta_2, \ldots, \beta_M)^T$ and $\psi(x) = (\psi_1(x), \ldots, \psi_M(x))^T$.

If $b$ denotes the number of paths generated by the underlying Markov chain then we may calculate the value of the option at every time step. So, by setting $\hat{V}_{ij} = \max \left\{ h_i(X_{ij}), \hat{C}_i(X_{ij}) \right\}$, $j = 1, \ldots, b$.

The next step is to set $\hat{V}_0 = (\hat{V}_{11} + \cdots + \hat{V}_{1b})/b$. This will estimate the option price.

A low biased estimator is possible to produce since the continuation values in the regression states an exercise policy. At every time step use the programming formulation and focus on stopping times. So, by simulating paths, the first time the Markov chain hits the boundary will determine the low biased price estimate.

### 3.7 Duality

All of the methods up to this point is based upon maximizing over stopping times. Another formulation is the dual method, proposed by Haugh and Kogan [10] and Rogers [15], in which a minimization problem arises. In the dual method we minimize over a class of supermartingales or martingales. Recall
from the dynamic sampling property that
\[ V_i(x) \geq E[V_{i+1}(X_{i+1})|X_i = x] \] (58)
for all \( i = 0, 1, \ldots, m - 1 \).
This is the defining property of a supermatingale. We also have that
\[ V_i(X_i) \geq h_i(X_i) \text{ for all } i \] (59)
So the value function \( V_i(X_i) \) is in fact the minimal supermartingale dominating \( h_i(X_i) \). For any stopping time \( \tau \) the sampling property of martingales gives
\[ E[h_\tau(X_\tau)] = E[h_\tau(X_\tau) - M_\tau] \leq E[\max_k \{h_k(X_k) - M_k\}] \] (60)
for any \( k = 1, \ldots, m \)
An arbitrage argument says that the price of the option is the supremum over \( \tau \), i.e the largest value of \( E[h_\tau(X_\tau)] \). The relation in (60) can be expressed as
\[ E[h_\tau(X_\tau)] \leq \inf_M E[\max_k \{h_k(X_k) - M_k\}] \]
and since the inequality holds for every \( \tau \), it also holds for the supremum over \( \tau \), hence
\[ V_0(X_0) = \sup_\tau E[h_\tau(X_\tau)] \leq \inf_M E[\max_k \{h_k(X_k) - M_k\}] \] (61)
where the infimum taken over martingales with initial value 0, i.e \( M_0 = 0 \).
In order to find the optimal martingale \( M_i, i = 1, \ldots, m \) define,
\[ \Delta_i = V_i(X_i) - E[V_i(X_i)|X_{i-1}] \]
and set
\[ M_i = \Delta_1 + \cdots + \Delta_i \]
If we now can find an optimal martingale, then an upper bound for the option price can be estimated using
\[ V_0(X_0) = \max_{k=1, \ldots, m} (h_k(X_k) - M_k) \] (62)
Finding the optimal martingale can easily fail due to the fact that there is no guarantee that the martingale difference \( \Delta_i \) fulfills the martingale property.
$E[\Delta_{i}|X_{i-1}] = 0$. If the martingale differences $\Delta_{i}$ fail to satisfy the martingale property then so will their sum $M_k$. The reason for this is that the continuation value $C_{i-1}(X_{i-1})$ does not necessarily equal $E[V_i(X_i)|X_{i-1}]$.

A way to extract martingale differences from approximate value functions is by using a nested simulation at every step $X_{i-1}$. This means that at every step of the Markov chain, generate $n$ successor nodes $X_{i}^{(1)}, X_{i}^{(2)}, \ldots, X_{i}^{(n)}$ and estimate the continuation value by

$$
\hat{C}_{i-1}(X_{i-1}) = \frac{1}{n} \sum_{j=1}^{n} V_i \left( X_{i}^{(j)} \right) \tag{63}
$$

By using this procedure we end up with the simulated values

$$
\Delta_{i} = V_i(X_i) - E[V_i(X_i)|X_{i-1}] = V_i(X_i) - \frac{1}{n} \sum_{j=1}^{n} V_i \left( X_{i}^{(j)} \right) \tag{64}
$$

which are martingale differences and hence - the conditional expectation at node $X_i$ given $X_{i-1}$ is zero as required.

4 Estimating Sensitivities

As mentioned before, the hedge ratio parameters are used to construct portfolios that hedge against risk. In order to estimate the partial derivatives of a contract using Monte Carlo simulation, consider a contract $Y$ depending on a parameter $\theta$ i.e

$$
Y = Y(\theta) \quad \text{and let} \quad \alpha(\theta) = E[Y(\theta)]
$$

4.1 Finite Difference Approximations

Consider the output of a contract given by the setting above. The idea is to simulate $n$ replications $Y_1(\theta), Y_2(\theta), \ldots, Y_n(\theta)$ of the Markov chain and $n$ additional samples $Y_1(\theta + h), Y_2(\theta + h), \ldots, Y_n(\theta + h)$ where $h > 0$.

Define:

$$
\bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i(\theta)
$$

Create the forward difference estimator:
\[ \hat{\Delta}_F = \frac{Y(\theta + h) - Y(\theta)}{h} \]  
(65)

and the central difference estimator:

\[ \hat{\Delta}_C = \frac{Y(\theta + h) - Y(\theta - h)}{2h} \]  
(66)

where \( \hat{\Delta}_F \) and \( \hat{\Delta}_C \) are estimators of \( \alpha'(\theta) \).

The problem now is that the estimators \( \hat{\Delta}_F \) and \( \hat{\Delta}_C \) are not unbiased, since if \( \alpha \) is twice differentiable at \( \theta \) then for \( \hat{\Delta}_F \) we have:

\[ \alpha(\theta + h) = \alpha(\theta) + \alpha'(\theta)h + \frac{1}{2} \alpha''(\theta)h^2 + O(h^2) \]  
(67)

In this case the bias in (65) of the forward estimator is:

\[ \text{Bias}(\hat{\Delta}_F) = E[\hat{\Delta}_F - \alpha'(\theta)] = \frac{1}{2} \alpha''(\theta)h + O(h) \]  
(68)

In the case of the central difference estimator we need to simulate not just one additional point \( \theta + h \) but two additional points \( \theta + h \) and \( \theta - h \). This computational effort is time consuming but however, it yields an improvement in the convergence of the rate of the bias. If \( \alpha \) is twice differentiable at \( \theta \) then

\[ \alpha(\theta + h) = \alpha(\theta) + \alpha'(\theta)h + \frac{1}{2} \alpha''(\theta)h^2 + O(h^2) \]
\[ \alpha(\theta - h) = \alpha(\theta) - \alpha'(\theta)h + \frac{1}{2} \alpha''(\theta)h^2 + O(h^2) \]  
(69)

Subtracting these equations yields:

\[ \text{Bias}(\hat{\Delta}_C) = \frac{\alpha(\theta + h) - \alpha(\theta - h)}{2h} - \alpha'(\theta) = O(h^2) \]  
(70)

We see that the bias in (70) is of smaller order than the bias in (68). There is no general rule in deciding which method to use. The choice of estimator should be based upon whether the computation is supposed to be carried out instantaneously or, if time consumption is not to be taken into consideration.
To estimate second derivatives using the finite difference method we create the central difference:

\[
\frac{\overline{y}_{n}(\theta + h) - 2\overline{y}_{n}(\theta) + \overline{y}_{n}(\theta - h)}{h^2} \tag{71}
\]

Assuming that \( \alpha \) is four times differentiable and taking the expectation value of (71) we get:

\[
\frac{\alpha(\theta + h) - 2\alpha(\theta) + \alpha(\theta - h)}{h^2} = \alpha''(\theta) + O(h^2) \tag{72}
\]

When looking at the bias in the equations above one may think that we would achieve higher accuracy by taking the smallest \( h \) possible. Due to the fact that there is a dependence between \( h \) and \( \text{Var} (\overline{y}(\theta + h) - \overline{y}(\theta - h)) \) meaning that the effect \( h \) has on bias must be weighted against it’s affect on variance. There are three cases which arise when investigating the dependence between \( h \) and \( \text{Var} (\overline{y}(\theta + h) - \overline{y}(\theta)) \)

\[
\text{Var} (\overline{y}(\theta + h) - \overline{y}(\theta)) = \begin{cases} 
O(1) & \text{Case(i)} \\
O(h) & \text{Case(ii)} \\
O(h^2) & \text{Case(i)} 
\end{cases} \tag{73}
\]

Case(i) applies if we simulate \( \overline{y}(\theta + h) \) and \( \overline{y}(\theta) \) independently. Case(ii) arise when simulating \( \overline{y}(\theta + h) \) and \( \overline{y}(\theta) \) using common random numbers. Case(iii) is the same as for Case(ii) but we also impose continuity to the output \( \overline{y}(\cdot) \).

Due to the dependence between \( h \) and \( \text{Var} (\overline{y}(\theta + h) - \overline{y}(\theta)) \) estimating second order derivatives is much more difficult than estimating first order derivatives. The main difficulty when estimating partials, first order in general and second order in particular, is finding the optimal \( h \). (See Glasserman [9] p.378-385 for further information regarding this issue)

\subsection*{4.2 The Hedge Parameter Delta}

The delta of an option is defined as the rate of change of the option price w.r.t. the price of the underlying asset. The delta of an option dependent on a single asset \( X \) is mathematically expressed as:

\[
\Delta = \frac{\partial C}{\partial X} \tag{74}
\]

Recall from proposition 2 that call deltas are positive - ranging from 0 to 1. Put deltas are negative - ranging from -1 to 0. In the exercise region, call option delta will approach 1.00 meaning that the option is in-the-money and
in the continuation region the call option delta will approach 0.00. The option is then out-of-the-money.

### Table 4: Delta hedge ratios of a standard American put option.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td></td>
<td></td>
<td>T</td>
<td>Binomial</td>
<td>Binomial</td>
</tr>
<tr>
<td></td>
<td>(yrs)</td>
<td></td>
<td>(European)</td>
<td>(American)</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>0.2</td>
<td>1.0000</td>
<td>0.1903</td>
<td>-0.2078</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>0.2</td>
<td>0.0833</td>
<td>-0.0270</td>
<td>-0.0272</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>0.4</td>
<td>1.0000</td>
<td>0.2781</td>
<td>-0.2916</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>0.0833</td>
<td>-0.1571</td>
<td>-0.1579</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>1.0000</td>
<td>0.3632</td>
<td>-0.4111</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.4</td>
<td>1.0000</td>
<td>0.3726</td>
<td>-0.3944</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>0.2</td>
<td>0.0833</td>
<td>0.5504</td>
<td>-0.6552</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>0.4</td>
<td>1.0000</td>
<td>0.4598</td>
<td>-0.4689</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>0.4</td>
<td>0.0833</td>
<td>0.4626</td>
<td>-0.4666</td>
<td></td>
</tr>
</tbody>
</table>

Columns (1)-(3) represent the parameter values, \( K \) (Strike price), \( \sigma \) (Volatility) and \( T \) (Time to maturity). Column (4) represent numerical results of the European delta hedge ratio valued numerically in the binomial lattice. Column (5) represent numerical results of the American delta hedge ratio valued numerically in the binomial lattice. Initial price \( X_0 = 100 \), dividend yield \( \delta = 0\% \), interest rate \( r = 5\% \). The number of time steps in the binomial tree is \( N = 5,000 \).

In Table 4, a comparison between deltas of a standard American put option and its corresponding European counterpart is displayed. Since the delta of a put option will lie between -1 and 0 it gives a hint on the nearness to the immediate exercise region, i.e free boundary.

### 4.3 The Hedge Parameter Vega

The Vega of an option indicates how much, theoretically at least, the price of the option will change as the volatility of the underlying asset changes. Vega is quoted to show the theoretical price change for every 1 percentage point change in volatility. For example, if the theoretical price is 2.5 and the Vega is showing 0.25, then if the volatility moves from 0.2 to 0.21 the theoretical price will increase to 2.75. Vega, which is not a greek letter, is often denoted the greek letter \( \kappa \). The vega of the option is defined as the partial derivative of the price w.r.t the volatility, hence

\[
\kappa = \frac{\partial C}{\partial \sigma} \quad (75)
\]

Table 5 displays a comparison of numerical values of the vega hedge ratio of standard American and European put options, valued in the binomial lattice.
Table 5: Vega hedge ratios of a standard American put option.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>Column (4)</th>
<th>Column (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Binomial</td>
<td>Binomial</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(European)</td>
<td>(American)</td>
</tr>
<tr>
<td>90</td>
<td>0.2</td>
<td>1</td>
<td>27.2167</td>
<td>28.2378</td>
</tr>
<tr>
<td>90</td>
<td>0.2</td>
<td>0.0833</td>
<td>1.8039</td>
<td>1.8147</td>
</tr>
<tr>
<td>90</td>
<td>0.4</td>
<td>1</td>
<td>33.5074</td>
<td>33.9673</td>
</tr>
<tr>
<td>90</td>
<td>0.4</td>
<td>0.0833</td>
<td>6.8972</td>
<td>6.9192</td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>1</td>
<td>37.5226</td>
<td>37.4873</td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>0.0833</td>
<td>11.4550</td>
<td>11.4266</td>
</tr>
<tr>
<td>100</td>
<td>0.4</td>
<td>1</td>
<td>37.8400</td>
<td>37.9300</td>
</tr>
<tr>
<td>100</td>
<td>0.4</td>
<td>0.0833</td>
<td>11.4630</td>
<td>11.4541</td>
</tr>
<tr>
<td>110</td>
<td>0.2</td>
<td>1</td>
<td>39.7608</td>
<td>33.6586</td>
</tr>
<tr>
<td>110</td>
<td>0.2</td>
<td>0.0833</td>
<td>3.4426</td>
<td>0.2937</td>
</tr>
<tr>
<td>110</td>
<td>0.4</td>
<td>1</td>
<td>39.7256</td>
<td>38.9816</td>
</tr>
<tr>
<td>110</td>
<td>0.4</td>
<td>0.0833</td>
<td>8.7456</td>
<td>8.4436</td>
</tr>
</tbody>
</table>

Columns (1)-(3) represent the parameter values, $K$ (Strike price), $\sigma$ (Volatility) and $T$ (Time to maturity). Column (4) represent numerical results of the European vega hedge ratio valued numerically in the binomial lattice. Column (5) represent numerical results of the American vega hedge ratio valued numerically using finite differences in the binomial lattice. Initial price $X_0 = 100$, dividend yield $\delta = 0\%$, interest rate $r = 5\%$. The number of time steps in the binomial tree is $N = 5,000$.

When looking in the table it seems like the vegas of the European and the American option are correlated. One value may be surprising, namely 0.2937, i.e when $K = 110$, $\sigma = 0.2$ and $T = 0.0833$. The explanation why the value of the American vega suddenly drops to less than 10% of the european vega can be found in the chosen parameter values. Due to the specific conditions, the time horizon is rather short, the initial price is lower than the strike price and the volatility is reasonably low, the underlying asset will start off close to the free boundary. This will trigger the model for early exercise and hence, the price will be less sensitive of a volatility change. As a comparison consider a european option with a short maturity date, $T = 0.0027$ (one day), say. The vega of a european style call option can be determined analytically by

$$\kappa = \frac{\partial C}{\partial \sigma} = X \phi(d_1) \sqrt{T-t}$$ \hspace{1cm} (76)

where $\phi(x)$ is the normal probability density function and $d_1$ is given by

$$d_1 = \frac{\ln(X/K) + (r + \sigma^2/2)(T-t)}{\sigma \sqrt{T-t}}$$ \hspace{1cm} (77)

Now, since the time horizon is short the factor $\phi(d_1)$ in equation(76) will be less then 1 and since $\sqrt{T-t}$ is small the value of the vega will approach zero. The vega of the American option can of course not be determined analytically but the punch line of this reasoning is that the time horizon to maturity, the
difference between the strike price and initial price and the volatility all contribute to a specific state of nature for the American option. That is why the imaginable conclusion that there is a high correlation between European and American vega hedge ratios should be distrusted.

5 Numerical Results

This section provides some numerical results to demonstrate the pricing methods covered in the earlier sections. Various contract structures have been implemented to give an illustration to the effectiveness different methods have as price approximations. All of the contracts are written on an underlying stock and they are governed by a geometric Brownian motion, the Black-Scholes model and it’s assumptions about the market. The first part considers the one dimensional case (puts and calls). The second part considers the multi-asset case, i.e contracts written on several underlying stocks and in the last part some results regarding hedge ratios will be presented.

5.1 Single Asset American Options

Table 6 below displays numerical results of different methods by the use of simulation. Those parameter values which are varied are strike price, \( K \), and volatility, \( \sigma \). The reason for this decision is that they are assumed to have the largest affect on price fluctuations in the models tested.

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K )</td>
<td>( \sigma )</td>
<td>Binomial (( N=8,000 ))</td>
<td>Stoch.Mesh (High Est)</td>
<td>Stoch.Mesh (Low Est)</td>
<td>Random tree (High est)</td>
<td>Regression</td>
</tr>
<tr>
<td>90</td>
<td>0.2</td>
<td>4.6476</td>
<td>5.6721</td>
<td>4.6319</td>
<td>5.0917</td>
<td>4.7154</td>
</tr>
<tr>
<td>90</td>
<td>0.3</td>
<td>9.8518</td>
<td>12.6865</td>
<td>9.5210</td>
<td>10.6058</td>
<td>10.0269</td>
</tr>
<tr>
<td>90</td>
<td>0.4</td>
<td>15.3298</td>
<td>16.6056</td>
<td>14.7856</td>
<td>15.4403</td>
<td>15.6200</td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>8.3209</td>
<td>9.8505</td>
<td>7.8333</td>
<td>9.4207</td>
<td>8.3672</td>
</tr>
<tr>
<td>100</td>
<td>0.3</td>
<td>14.3040</td>
<td>18.8725</td>
<td>14.1477</td>
<td>15.1904</td>
<td>14.4387</td>
</tr>
<tr>
<td>100</td>
<td>0.4</td>
<td>20.3290</td>
<td>26.8847</td>
<td>20.0718</td>
<td>23.3848</td>
<td>20.6118</td>
</tr>
<tr>
<td>110</td>
<td>0.2</td>
<td>13.4220</td>
<td>16.9556</td>
<td>12.7856</td>
<td>14.2748</td>
<td>13.4722</td>
</tr>
<tr>
<td>110</td>
<td>0.3</td>
<td>19.6403</td>
<td>24.5641</td>
<td>18.8800</td>
<td>20.8245</td>
<td>19.8020</td>
</tr>
<tr>
<td>110</td>
<td>0.4</td>
<td>25.9550</td>
<td>33.5773</td>
<td>24.9931</td>
<td>26.7533</td>
<td>26.2277</td>
</tr>
</tbody>
</table>

Columns (1)-(2) represent the parameter values, \( K \) (Strike price), \( \sigma \) (Volatility). Columns (3)-(7) represent numerical results of option prices by the binomial tree method, (\( N \) number of time steps), Stochastic mesh (High and Low estimators) and Random tree (High estimator), respectively. Initial price \( X_0 = 100 \), dividend yield \( \delta = 0\% \), interest rate \( r = 5\% \) and time to maturity (yrs) \( T = 3 \). The option can be exercised at four equally spaced dates.

The stochastic mesh method is simulated using 100 nodes in each time step and the low estimator is computed using 50 subpaths. The simulation is then
batched 5 times to stabilize the variance. For the random tree method, simulation is based on $b = 20$ as the branching parameter and the tree is batched 10 times. Using a branching parameter of only 20 may seem inadequate, however, since the tree is batched five times, a branching parameter of 20 is acceptable. The computational burden of the random tree method is the reason for this decision. The number of nodes generated and the number batches and are chosen with the purpose that computation time of both the random tree and the stochastic mesh is more or less the same. This is also why the low estimator of the random tree method is not presented, otherwise computation time comparison would be impossible.

In the regression based method 10,000 paths are generated and batched 100 times. The basis functions used is $1, X_i, X_i^2, X_i^3, X_i^4$.

### 5.2 Multi Asset American Options

Multi-asset options like max options and basket options are contracts well suited to demonstrate the strength of using Monte Carlo simulation. In the first case, a max option written on two underlying assets are simulated. The main focus will be held on the regression based method with 4000 simulated paths and 100 batches. The dual method is presented because of it’s characteristics of producing an upper bound for the price estimate.

<table>
<thead>
<tr>
<th>$K$</th>
<th>$\sigma$</th>
<th>Binomial (N=270)</th>
<th>Regression (Low bias)</th>
<th>Regression (n=10)</th>
<th>Dual-$\hat{V}$ (n=10)</th>
<th>Dual-$\hat{V}$ (n=100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>0.2</td>
<td>20.0433</td>
<td>20.1468</td>
<td>19.9182</td>
<td>22.0994</td>
<td>20.3776</td>
</tr>
<tr>
<td>90</td>
<td>0.3</td>
<td>29.5723</td>
<td>30.0117</td>
<td>29.477</td>
<td>33.2631</td>
<td>30.2106</td>
</tr>
<tr>
<td>90</td>
<td>0.4</td>
<td>39.4820</td>
<td>40.2262</td>
<td>39.4166</td>
<td>44.8936</td>
<td>40.6976</td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>13.8992</td>
<td>14.9423</td>
<td>13.8669</td>
<td>15.4051</td>
<td>14.1636</td>
</tr>
<tr>
<td>100</td>
<td>0.3</td>
<td>23.8442</td>
<td>24.2175</td>
<td>23.7808</td>
<td>26.8912</td>
<td>24.4489</td>
</tr>
<tr>
<td>100</td>
<td>0.4</td>
<td>34.0964</td>
<td>34.5350</td>
<td>34.0653</td>
<td>39.4485</td>
<td>35.2222</td>
</tr>
<tr>
<td>110</td>
<td>0.2</td>
<td>9.3890</td>
<td>9.5656</td>
<td>9.3581</td>
<td>10.5052</td>
<td>9.5607</td>
</tr>
<tr>
<td>110</td>
<td>0.3</td>
<td>19.1601</td>
<td>19.7434</td>
<td>19.1158</td>
<td>21.7751</td>
<td>19.6098</td>
</tr>
<tr>
<td>110</td>
<td>0.4</td>
<td>29.4830</td>
<td>30.4185</td>
<td>29.4306</td>
<td>33.5819</td>
<td>30.4695</td>
</tr>
</tbody>
</table>

Columns (1)-(2) represent the parameter values, $K$ (Strike price), $\sigma$ (Volatility). Columns (3)-(6) represent numerical results of option prices by the binomial tree method with 270 time steps, regression, Dual-$\hat{V}$ with 10 and 100 generated subpaths, respectively. Initial price $X_0^1 = X_0^2 = 100$, dividend yield $\delta = 10\%$, interest rate $r = 5\%$ and $T = 3$ (Time to maturity). Correlation $\rho = 0$. The option can be exercised at nine equally spaced dates.

The regression method is based on the following basis functions

$$1, X_i, X_i^2, X_i^3, X_1X_2, X_1X_2^2, X_1^2X_2, \text{max}(\text{max}(X_1, X_2) - K, 0), i = 1, 2$$

The binomial tree is used as a benchmark. Computation time for the regression based method is significantly higher than the binomial tree method, however
one should keep in mind the small number of generated nodes in each time step, \( N = 270 \).

Table 8 displays numerical results of a basket option written on two underlying assets simulated using the regression based method and the 2-D binomial lattice, used as a benchmark. The European style counterpart is also presented to demonstrate the affect of the early exercise premium. The discrepancy between the the American option price and the European option price will then be the price for the privilege of early exercise. The European style is valued in the binomial lattice using the same number of time steps, \( N = 270 \), as for the American style.

Table 8: American two-asset basket-option priced by various methods.

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K )</td>
<td>( \sigma )</td>
<td>Binomial</td>
<td>Regression</td>
<td>Dual-( \hat{V} )</td>
<td>European</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(N=270)</td>
<td>(n=100)</td>
<td>(N=270)</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>0.2</td>
<td>10.0000</td>
<td>9.9475</td>
<td>10.1308</td>
<td>5.9119</td>
</tr>
<tr>
<td>90</td>
<td>0.3</td>
<td>13.2250</td>
<td>13.3297</td>
<td>13.5550</td>
<td>9.7451</td>
</tr>
<tr>
<td>90</td>
<td>0.4</td>
<td>17.0622</td>
<td>17.4539</td>
<td>17.6275</td>
<td>13.6964</td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>4.9279</td>
<td>4.9807</td>
<td>5.0744</td>
<td>3.3114</td>
</tr>
<tr>
<td>100</td>
<td>0.3</td>
<td>8.8997</td>
<td>9.1038</td>
<td>9.1864</td>
<td>6.9441</td>
</tr>
<tr>
<td>100</td>
<td>0.4</td>
<td>13.1492</td>
<td>13.5730</td>
<td>13.7138</td>
<td>10.9082</td>
</tr>
<tr>
<td>110</td>
<td>0.2</td>
<td>2.3956</td>
<td>2.5296</td>
<td>2.5361</td>
<td>1.7722</td>
</tr>
<tr>
<td>110</td>
<td>0.3</td>
<td>6.0250</td>
<td>6.2985</td>
<td>6.3143</td>
<td>4.9068</td>
</tr>
<tr>
<td>110</td>
<td>0.4</td>
<td>10.2202</td>
<td>10.6987</td>
<td>10.8979</td>
<td>8.6971</td>
</tr>
</tbody>
</table>

Columns (1)-(2) represent the parameter values, \( K \) (Strike price), \( \sigma \) (Volatility). Columns (3)-(7) represent numerical results of option prices by the binomial tree method with 270 time steps, regression, Dual-\( \hat{V} \) with 100 generated subpaths and the corresponding European price, respectively. Initial price \( X_{10}^1 = X_{10}^2 = 100 \), dividend yield \( \delta = 10\% \), interest rate \( \tau = 5\% \) and \( T \) (Time to maturity). Correlation \( \rho = 0 \). The option can be exercised at nine equally spaced dates.

When the number of underlying assets exceeds two, Monte Carlo simulation becomes a necessity. In table 9 a basket option of five underlying assets has been simulated using the regression based method. Because of the obstacle of not having a general benchmark to compare values against, the American five-asset contract is compared with the European style counterpart simulated by Monte Carlo.

In theory, the value of the American style contingent claim should always be higher than the European style contingent claim due to the early exercise premium. High dimension imposes, not only mathematical, but also computational problems typically connected with machine precision and storage requirements. This becomes evident when choosing basis functions in the regression based method. Since the node values at every time step is rather large their exponential can not exceed five. It is also not certain that by taking as many basis functions as possible that this will end up in a better approximation of the continuation value - the value may be over estimated. In the table 9, simulated values are based on the following basis functions.
Table 9: Comparison of American and European five-asset basket option.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>σ</td>
<td>European Regression Dual-(\hat{V}) (n=100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>0.2</td>
<td>3.2828</td>
<td>8.3494</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>0.3</td>
<td>5.7807</td>
<td>9.6099</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>0.4</td>
<td>8.5692</td>
<td>11.9292</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>1.1781</td>
<td>2.3784</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.3</td>
<td>3.2225</td>
<td>4.8151</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.4</td>
<td>5.8059</td>
<td>7.7352</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>0.2</td>
<td>0.3552</td>
<td>0.6793</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>0.3</td>
<td>1.7066</td>
<td>2.5136</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>0.4</td>
<td>3.9139</td>
<td>5.1969</td>
<td></td>
</tr>
</tbody>
</table>

Columns (1)-(2) represent the parameter values, \(K\) (Strike price), \(\sigma\) (Volatility). Column (3) represent numerical results of a European five-asset basket option priced by Monte Carlo. Column (4) and (5) represent numerical results of the American basket option by the use of regression and Dual-\(\hat{V}\) with 100 generated subpaths, respectively. Initial price \(X_1^0 = X_2^0 = X_3^0 = X_4^0 = X_5^0 = 100\), dividend yield \(\delta = 10\%\), interest rate \(r = 5\%\) and time to maturity(ys) \(T = 3\). Correlation \(\rho = 0\). The option can be exercised at nine equally spaced dates.

\[
1, X_1, X_1^2, X_1^3, X_1 X_2, X_1 X_2^2, X_1^2 X_2, \max \left( \frac{1}{5} \sum_{i=1}^{5} X_i - K, 0 \right)
\]

5.3 Hedge Ratios

Estimating sensitivities, like the option delta and vega, in high dimension using Monte Carlo is time consuming. This is because the simulation needs to be extended to \(n\) additional independent simulations, where \(n\) denotes the dimension of underlying assets. Since the simulation also needs to be executed simultaneously to avoid high biased effects (discussed in section 4.1), the computation time is multiplied by a factor \(2^n\).

Table 10 displays a comparison between Monte Carlo simulated option deltas and numerical evaluated option deltas of an American max option written on two underlying assets. The regression based method seems to under estimate the values of those in the binomial tree.

5.4 Comments on the numerical results

The numerical values in table 6 are all simulated with the aim of having approximately the same computation time. This is the only way of evaluating the methods from an "accuracy and time" point of view. Still, there are other
Table 10: Delta hedge ratios of an American max option (2 assets).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Binomial</td>
<td>Regression</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(N=270)</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>0.1</td>
<td>3</td>
<td>0.4570</td>
<td>0.4645</td>
</tr>
<tr>
<td>90</td>
<td>0.2</td>
<td>3</td>
<td>0.4131</td>
<td>0.4108</td>
</tr>
<tr>
<td>100</td>
<td>0.1</td>
<td>3</td>
<td>0.2813</td>
<td>0.2758</td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>3</td>
<td>0.3348</td>
<td>0.3210</td>
</tr>
<tr>
<td>110</td>
<td>0.1</td>
<td>3</td>
<td>0.1103</td>
<td>0.0998</td>
</tr>
<tr>
<td>110</td>
<td>0.2</td>
<td>3</td>
<td>0.2455</td>
<td>0.2320</td>
</tr>
</tbody>
</table>

Columns (1)-(3) represent the parameter values, K (Strike price), \( \sigma \) (Volatility) and \( T \) (Time to maturity). Column (4) represent numerical results of the delta hedge ratio valued numerically in the 2-D binomial lattice using \( N = 270 \) time steps. Column (5) represent numerical results of the American delta hedge ratio estimated using Monte Carlo simulation. Initial price \( X_1^0 = X_2^0 = 100 \), dividend yield \( \delta = 10\% \), interest rate \( r = 5\% \). The option can be exercised at nine equally spaced dates.

ways of looking at it. Sometimes a computation time of, say 20 minutes, is acceptable and would certainly make the choice of method different than if a computation time of 10 seconds is maximum. To generalize - the longer you can endure the plague of waiting for computations to be carried out, the more accurate will your estimate become. This is especially true for methods based on tree structures.

It is noticeable that the high estimator of the stochastic mesh gives rather high values, compared with the high estimator of the random tree. This maybe has to do with the small number of exercise opportunities (four, to be precise). By comparing the total number of nodes generated by each method, the random tree has generated 160,000 nodes, while the stochastic mesh only has generated 400 nodes. A conclusion is that the stochastic mesh is not a good method to use when the number of exercise opportunities is less then, say, five. On the other hand, the low biased estimator of the stochastic mesh evidently gives a lower bound close to the true price, evaluated in the binomial tree.

The regression based method gives a good approximation, regardless of parameter values. This is not the case in the stochastic mesh and the random tree, where accuracy is apparently dependent upon the chosen parameter values.

In table 7 numerical results of a max option written on two underlying assets is presented. We see that the regression based method still does a good job at approximating the true price. The accuracy of the dual method is highly dependent upon the number of subpaths generated. A suggestion is that, if the dual method is used in order to given a valid upper bound for the option price, then the number of generated subpaths should as many as possible. This conclusion is only based on empirical results, and further studies regarding this subject is requested. The table also shows that the low estimator of the regression based method gives a lower bound close to the true price.
In table 8 numerical results of a two-asset basket option is presented. Having exchanged the max option for a basket option does not seem to affect the accuracy of the regression based method. It still gives a good approximation. Numerical results for the European style counterpart is also presented, and the cost of the early exercise privilege clearly appears. In table 9 the two-basket option is extended to a five-asset basket option, and the simulated values of the regression method is compared with corresponding European five-asset basket option.

Finally, table 10 display numerical values of the delta hedge ratios for an American max option, written on two underlying assets. The simulated delta values are compared with delta values numerically evaluated in the binomial tree. The regression based method is here used to demonstrate how versatile this method is. The delta values are slightly lower than the binomial deltas in most cases, but nevertheless acceptable.

5.5 Suggestions for further studies

The framework done by the authors of the methods covered in this thesis can be extended in many directions. For example, investigation of the effect the choice of basis functions has on the estimated price of an American option. Implementing variance reduction techniques could be an interesting addition, making computation less time consuming in the stochastic mesh and the random tree. Impose stochastic volatility in the dynamics of the underlying asset. Models for mean reverting volatility, e.g the Heston model, would give a more realistic model for the dynamics of the underlying asset. In higher dimension there is much work left to be done. High dimensional path dependent payoff structures and the option hedge ratios in high dimension, is such areas where there is little research done.

6 Conclusions

Except for a few special cases, Monte Carlo simulation is to present day, the only way of pricing high dimensional American options. However, for path independent payoff structures written on one or two underlying assets there is no need to use Monte Carlo simulation. In this area the binomial tree stands out as a universal method by three reasons - it is fast, accurate and easy to implement. Methods based on simulation requires substantial computational effort, and the choice of method should be based upon the characteristics of the contract structure. The random tree method has its limitations in the exponential growth of generated nodes which only makes this method acceptable if time consumption is not to be considered. This is, however, not the case in real-life finance where computations need to be carried out instantaneously.
The stochastic mesh method does not have the same problem of exponential growth as in the random tree framework, but has instead requirements of storage capacity for weight functions. As can be seen in table 6 the high estimator of the stochastic mesh seems to overestimate the option price significantly while the high estimator of the random tree gives a much closer higher bound in most studied cases. The reason for this could be that, in order for the stochastic mesh to give a plausible price approximation, the number of generated nodes in each time step needs to be significantly higher than the number of nodes used here. Yet, the low estimator of the stochastic mesh gives a sustainable low biased price approximation. Common of the random tree and the stochastic mesh seems to be that accuracy is highly dependent on the discrepancy of the strike and initial price and also the volatility. This is evident for the high estimator of the stochastic mesh where the price estimate seems sensitive for increase in volatility.

If the random tree or the stochastic mesh should be used to price an American option some kind of variance reduction technique is a necessity in order to improve the speed and efficiency of the simulation algorithm. Parallel computing would also lead to large improvements and a decrease in computation time.

The regression based method in an interesting approach. It is rather easy to implement and, as can be seen in table 6, gives a good approximation of the option price. It has the advantage of not being to much demanding when it comes to computational capacity and storage requirements and it is applicable for a large number of contract structures. The one thing that makes the method somewhat demanding is the problem of choosing an appropriate set of basis functions, in order to produce a plausible price approximation.
References


