Nikolaos Mourtzanos

Athens University of Economics and Business

Email: mourtzanosn@yahoo.gr

Comparison of Numerical Methods for Option Pricing

Abstract

This study goes through a range of methods for option pricing. We begin with the celebrated Black-Scholes formula, and then we begin examining methods that do not provide closed-form solutions, namely the finite-difference method, binomial tree and simulations. We examine the accuracy of Least Squares Monte Carlo method, and we also examine how simulation can be used for options with stochastic volatilities.

We used GAUSS v3.2.32 to develop the routines of the algorithms we had to examine. The routines were compiled on a single desktop with a 2.6 GHz Intel Pentium Processor and 1GB RAM. Analytic results of al the methods are cited, and extra weight is given to simulations.

December 2006

Chapter 1
Introduction

Due to their nature, it is very difficult to value American options, especially in more than one underlying asset, and more than one stochastic factor, other than the price of the underlying asset. Valuation of American options is critical not only for traders but for corporate practitioners too. These types of derivatives are very popular in markets around the world, most important being CBOE, and include equity, commodity, foreign exchange, energy, credit, mortgage and emerging markets. That is why effective option pricing is crucial.

Unfortunately, closed-form solutions, like the celebrated Black-Scholes-Merton (B-S) formula, do not exist for the majority of options. The above formula can well price at-the-money European options, on one underlying asset, assuming lognormal distribution for the asset's price and constant, pre-determined volatility for the changes of the asset's price, during the life of the option. The purpose of this study is to cover numerical methods for option pricing, methods that don't provide a closed-form analytical solution for the price of an option derivative. These methods have been developed to mainly value American options, and further, ones that have more than one underlying asset, and regard volatility as a stochastic variable. The first methods to be mentioned are the lattice methods, the most famous being the binomial tree, numerical integration methods and the finite difference method. As it will extensively explained later, those methods cannot handle options with more than two or three underlying assets and additively ones with stochastic volatility. A more effective method for dealing with this kind of options is simulation, which allows for effective option pricing, in the expense of computational time, especially when we cannot use a network of workstations.

This study describes the assumptions used for option pricing, and then provides a comprehensive citation of the methods mentioned above, giving special attention to simulation as being more effective and more adaptive than the other methods.

Chapter 2 is an introduction to some basic definitions regarding options. Chapter 3, provides a thorough literature review, concerning option pricing theory. Chapter 4, introduces basic assumptions made for the pricing of options namely the risk-neutral assumption, the arbitrage free theory and the assumption for the lognormal distribution of the changes in the price of the underlying asset's price. Chapter 5 provides a description of the B-S method, finite difference method, and a binomial approach of the lattices. Chapter 6 cites Monte Carlo simulation as it was first introduced for option pricing, by Boyle (1976), and the Least Squares Monte Carlo simulation algorithm as it was introduced by Longstaff and Schwartz (2001). So far, this study will be covering options with one underlying asset and constant volatility, both American and European. In Chapter 7, we will also cover options with stochastic volatility being a second underlying asset, using the Least Squares approach to value them.

Chapter 2
Basic definitions

An option is a tradable security that provides its owner with the right to exercise a claim at the expense of the counterparty that issued the option (the writer of the option). The latter charges a fee for the risk of incurring possible loss.

There are numerous types of options like energy options, weather options, some simple options like vanilla options on one stock, index options, and real options. The main characteristics of an option are its strike price, the variance of the changes in the price of its underlying asset, its time to expiration, and its payoff function.

Options are categorized as being call or put options, and European or American. A call option, i.e. on a stock, gives the holder of the option the possibility to buy the stock for the contractually agreed strike price, while the put option gives him the right to sell it. The European option can be exercised only at the specified date that the option expires, while the American option can be exercised at any time till the expiration date. If we buy a 3-months European call option on i.e. 100 IBM stocks, we will exercise it only after 3 months and only if the price of the IBM shares that time is greater than the strike price we had agreed when we bought the option. If we on the other hand possessed an American put option with the same time to expiration and the same strike price, we would exercise our option to sell IBM shares at any time till the expiration of the option, and only in the case where the strike price of our option had been greater than the price of IBM stocks, at any given time less than, or equal to 3 months.

Apart from the strike price, and the time to expiration, another major parameter of an option contract is the volatility in the changes in the price of its underlying asset. The simplest case is to assume a constant volatility rate, like Black-Scholes and Merton have done to derive their celebrated closed-form solution for the pricing of European options. Clearly this assumption doesn't hold so we will provide other methods to deal with non constant, stochastic volatility, later in this study. It suffices to say that the first derivative between volatility rates and option price is positive. That happens because a greater volatility makes the expected value of the option after its assumption more uncertain, that is, the probabilities that the price of the underlying asset does pretty good or pretty bad are greater when its volatility rate is increased. So the holder of a call option can have great earnings if the price of the underlying asset goes unexpectedly high, and the maximum loses he can suffer is the option value, if the course of the price of the asset underlying her option moves very low. The opposite happens for the holder of a put option. So it would be intuitive to say that the buyer of the option has to pay a greater price, for the possible gains he can have, while keeping the possible losses stable.

Finally options are characterized from their payoff function, elements of which were cited with the example of the option on the IBM shares. The payoff function for a European call option, at expiry T, strike price K, and the price of its underlying asset (i.e. a stock) at expiry, namely S_T, is:

$$C_E(S_T,T) = \max\{S_T-K,0\}$$

and the corresponding payoff function for a put option is:

$$P_E(S_T,K) = \max\{K-S_{T,}0\}$$

What the above equations are telling us is that the holder of a European call option on a stock, with a strike price of \$50, will only exercise her option if the price of the stock at the time of expiry, say 1 year, is greater than the strike price of the option, i.e. if that price reaches \$55, then the option will be exercised, giving its holder a payoff of 55-50=\$5 and a gain of \$5 minus the value of the option. If the price at the time of expiry is \$45, then the option will not be exercised, its payoff will be \$0, and the holder will suffer losses equal to the value of the option. In the same way, if we were to have a put option, with the same characteristics, we would only exercise it if the price of the stock at expiry was less than the strike price of the option, say \$45. Then the payoff would be 50-45=\$5 and the net gain would be our payoff minus the option value. In any other way, the option would not be exercised, giving a zero payoff, and losses equal with its value. More generally when the underlying asset's price at the expiration of the option equals the strike price, then the option is said to be at-the-money. When the asset's price at expiration is such that we have zero payoff the option is out-of-the-money, and when a positive payoff occurs, the option is in-the-money.

As far as American options are concerned, their payoff function cannot be written in two lines, with two little equations, and it will be cited in great detail in the following chapters.

Chapter 3
Literature review

Before proceeding with the description of the methods we will cover, it would be useful to review what the evolution has been, in terms of literature, of the option pricing theory. We will see when certain methods were introduced, what improvements have been made to them, and in the sequel we will provide the reader with a detailed description of the methods that this study covers.

We will begin with the celebrated Black-Scholes-Merton framework, which provides a closed-form solution, published in 1973, in the well famous paper titled: *"Black, F., M. Scholes. 1973. The pricing of options and corporate liabilities. Journal of Political Economy. 81 637-654".* That was a paper providing us with an analytical solution for the pricing of European options, based on some pretty strong assumptions to which we will refer in the next chapters. An important assumption was to regard volatility as a constant term through the life of the option. By empirical observation of the data provided by the markets, this assumption was impossible to hold, even though the B-S framework worked pretty well for at-the-money options for relatively small expirations like less than 2 or even one year.

Another category of models which this study will not cover is the jump-diffusion models, which were first introduced by Merton (1976). In this model, the volatility is also regarded as constant, and we assume a Poisson distribution for the returns of the stock. The model is:

$$\frac{dS_t}{S_t} = (\mu - \delta)dt + \sigma dW_t + d\left(\sum_{n=1}^{N_t} (e^{Z_n} - 1) \right)$$

where N_t is the Poisson process, with intensity λ and $Z_n \sim (\mu_s, \sigma_s^2)$ is the distribution of the jumps in returns, which are independent of the Brownian motion W_t. λ actually determines the arrival rate of jumps in the stock's returns, while μ_s and σ_s constitute the mean and volatility of the returns. While this model produces some good results when compared with the empirical asset price data, it still suffers from the assumption of constant volatility. Other models that enable jumps in returns are the ones published by Bates (1996), Scott (1997), Madal et al. (1998), Carr et al. (2002), Kou in the same year and Carr and Wu (2003).

In order to address the problem of the volatility not being constant, option pricing theory evolved by producing models that regarded volatility as a stochastic process, just like the process followed by a stock underlying an option to be valued. The first models to mention is the ones of Hull and White (1987), Stein and Stein (1991), and Heston's model published in 1993. The most famous model among them is the Heston model for pricing options with stochastic volatility, and that's the model we will mainly cover, among others. What made this model so famous is that it assumes a degree of correlation between the returns of the stock underlying the option and the volatility itself, while for example Hull and White assume zero correlation. Heston also provided a closed-form analytical solution for the pricing of European options, which worked pretty well not only for at-the-money options, with short expirations like B-S framework did, but it also provided us with an efficient way to deal with out-of-the-money and in-the-money options, with larger expirations.

We will proceed with the part of the option pricing theory concerning Numerical Methods. The main reason that made the development of these methods necessary is the fact that B-S model and all the models that assumed constant volatility had by nature certain disadvantages. As far as the B-S formula was concerned it only gave us reliable results for at-the-money European options based on the assumptions of perfect, frictionless markets, and the assumption of constant volatility, among other assumptions we will cite at the next chapters. So we could not get a price for American options, not to mention more complicated options like the ones with multiple underlying assets, with stochastic volatility and options that had big expirations like 15 years. But even for the options that this formula could manage, we needed numerical computations, in order to calculate the integrals of the probability that option ends up in the money area. Furthermore we couldn't handle options with stochastic volatilities as already mentioned, which leaded to the creation of Stochastic Volatility models, which in turn could not also provide a solid solution for American style options with more than one underlying assets.

The major categories of Numerical Methods are: a) formulas and approximations which we will not cover, b) lattice methods, which are covered, c) finite difference methods and d) Monte Carlo Simulation which are both covered.

In the first category, the most important method to be mentioned is the application of transform methods, asymptotic expansion techniques, and also Fourier and Laplace transformations. These techniques were used by many authors for the pricing of options with stochastic volatility, as mentioned above, namely Heston (1993), Stein and Stein (1991), and Duffie et al. (2000), Hull and White (1987), Hagan and Woodward (1999), Fouque et al. (2000) and Leif Andersen et al. (2001). They were also used for the pricing of more complicated options like Asian options (Reiner (1990), Geman and Yor (1993)).

The first publication on the *lattice methods* is due to Parkinson (1977). But the most famous form of lattices is the one introduced by Cox, Ross and Rubinstein in 1979. The latter is a binomial tree which is easy to implement and gives pretty accurate results but is used mainly for pricing "easy" options, with constant volatility and no more than 3 underlying assets, a drawback named "the curse of dimensionality" We will explain later why that is so. Lattices can be used for relatively complex derivatives, as shown by Heston and Zoo (2000) and Alford and Webber (2001). The major drawback is that these advanced algorithms require a lot of computational time, more than is needed for other methods to give accurate results. This was shown among others to some specialized papers like the one of Broadie and Detemple (1996). Other publications on lattices include the ones of Coval, J. E. and t. Shumway (2001), Rendleman, R., and B. Bartter (1979), Figlewski, S., and B. Gao (1999), Hull, J. C., and A. White (1988).

The use of *finite difference method* for option pricing was first introduced by Brennan and Schwartz (1977, 1978). Finite difference methods value a derivative by solving the stochastic differential equation (SDE) that the latter satisfies. Again as we will explain in great detail, the major drawback of this approach is its inability to value options with more than three underlying assets, as solutions for this types of SDEs are not yet available. They can incorporate though options with stochastic volatilities and jump diffusions. More publications on the finite difference approach include those of Hull, J. C., and A. White (1990), and Wilmott, P., (1998).

As shown in Glasserman (2004), a complete reference on *Monte Carlo simulation* approach would require over 350 references. So we will include the most important among them leaving the description of this method for later chapters. It suffices here to say that in Monte Carlo simulations we produce a large number of paths that the price of the underlying asset could follow in subsequent timesteps using the solution of the SDE that characterizes it. Then we average the discounted at the risk-free rate option values observed at expiration, in every single path, obtaining the option price. We will cover this method in a very great extend as it can handle every kind of option, from the simple plain vanilla options to exotic options, and every option one can come with.

Monte Carlo simulation was first introduced to value options by Boyle (1976). As long as this simulation was traditionally presented as a forward-looking technique, it had been impossible to value American, and in general path-dependent options, where we had to check at a continuous time basis whether it would be optimal to exercise the option or not. In recent years, a lot of authors have developed new approaches that enabled Monte Carlo simulation to implement backward-looking algorithms. These publications include those of Tilley (1993), Barraquand and Martineau (1995), Carriere (1996, 2001), Rayman and Zwecher (1997), Broadie and Glasserman (1997), Tsitsiklis and Van Roy (1999, 2001), Garcia (2002), Rogers in the same year, Ibanez and Zapatero (2004), Haugh and Kogan (2004), Anderson and Broadie (2004) and others. It is of great importance to mention in this point a great evolution that took place in Monte Carlo simulation techniques, which was the introduction of the Least Squares Method (LSM), by Longstaff and Schwartz in 2001, which due to its effectiveness and simplicity, was given a great deal of attention. Some authors that tested or tried to ameliorate the LSM, are Clement, Laberton and Protter (2002), Stentoft (2004), Rasmussen (2002), Pizzi and Pellizzari (2002) and Moreno and Navas (2003). This study will cover in extend the work done by Longstaff and Schwartz, and Stentoft.

Chapter 4

Basic assumptions of the options pricing theory

4.1 The price of the underlying asset

The first assumption made for the pricing of options, is the process followed by the underlying asset's price, and from now on we assume that the underlying asset of our options is a stock. We begin with the premise that the stock price S follows a geometric Brownian motion process

$$\frac{dS_t}{S_t} = (\mu - \delta)dt + \sigma dW_t \qquad (1)$$

where μ is the total expected (above the risk-free rate) return on the stock, δ is the dividend rate, σ is the volatility of the changes of the stock's price, that is the volatility of its returns, and W_t is a Wiener process. For the time being we assume μ, δ and σ as being constants. Later we will assume non-constant volatility, but later. If V is a function of S and time t, using Ito's lemma we derive the process followed by V

$$dV = \left(\frac{\partial V}{\partial S}(\mu - \delta)S + \frac{\partial V}{\partial t} + \frac{1}{2}\left(\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2 \right) \right) dt + \frac{\partial V}{\partial S}\sigma S dW_t \qquad (2)$$

This is a fundamental equation that any derivative whose price depends on a dividend paying stock must follow. In option pricing it is convenient to transform the stock's price to its logarithmic form, so we set G = lnS, and by computing the derivatives of equation (2), we have the process of the changes in the asset's price, written

$$dG = \left(\mu - \delta - \frac{\sigma^2}{2} \right) dt + \sigma dW_t \qquad (3)$$

lnS follows a generalized Wiener process, with constant drift rate $\mu - \frac{\sigma^2}{2} - \delta$ and constant volatility rate σ. From (3), we see that changes in S are lognormally distributed with a standard deviation $\sigma\sqrt{T}$. These conclusions will prove to be crucial for the derivation of option pricing algorithms covered by this study.

4.2 Arbitrage-free pricing

A very important concept in the option pricing theory is the assumption of arbitrage-free pricing. This assumption states that the total expected return of all the traded assets must equal the risk-free rate. We examine this statement in more detail.

To begin, we state that the presence of arbitrage opportunities in the market, allows for certain players to make riskless profit by buying and selling securities. This assumption leads to the derivation of the fair price of the option. Any price below or above this price, will provide the opportunity for riskless profit and will be vanished by the powers of supply and demand. To see why this happens, we first assume the existence of a risk-free security, a security that has a zero probability of default and that returns a continuously compounded rate of return r, the risk-free rate of return. We also assume that investors can borrow and lend at this rate, suffering no costs. Imagine two riskless portfolios A and B, each one consisting of one traded security and the risk-free asset, with yields a and b, where a>b, and current prices A_0 and B_0, where $A_0 = B_0$. Portfolio C can be constructed with a long position in A and a short one in B, and equal amount of money invested in both securities. The present value of the portfolio is

$$C_0 = A_0 - B_0 = 0 \qquad (4)$$

and its value after time T is

$$C_T = A_T - B_T = A_0e^{aT} - B_0e^{bT} = A_0e^{bT}(e^{(a-b)T} - 1) > 0 \qquad (5)$$

as $A_0 = B_0$ and $A_0 = B_0 > 1$. Portfolio C cost nothing to construct and it also has a positive payoff. Same think could happen if we had the opposite positions in the two securities and if b>a. So the player possessing this portfolio would make riskless gains, above the risk-free rate, if she for the first case bought, all the time, security A and sold B. The power of demand for security A would cause its price to rise and its yield to decline. The opposite would happen for security B by the power of supply. Eventually the two yields become equal, so we conclude that the return of each riskless portfolio must be the risk-free rate of return r in order to achieve the necessary equilibrium in the market. That is r = a = b.

4.3 Risk-neutral valuation

The natural sequel after the description of the arbitrage-free theory is the next strong assumption made in option pricing theory, the risk-neutral valuation. As far as all asset returns have to be equal to the risk-free interest rate, one can say that the expected return μ of the investors doesn't play any role to option pricing. That is exactly the essence of risk-neutral valuation theory, first introduced by Cox and Ross in 1976. More sophisticated forms of this theory can be found in Harrison and Kreps (1979) and Harrison and Pliska (1981). The description of the theory goes as follows:

Suppose that the stock price follows the process cited by equation (1). The risk-neutral valuation theory implies that the expected total return on the stock equals the risk-free rate, that is

$$\mathrm{E}_t\left[\frac{dS_t}{S_t}+\delta dt\right]=rdt \qquad (6)$$

where E_t is the conditional expectation at time t with respect to the Brownian motion W_t. Using the last assumption, we examine whether using the results of (6) in equation (5), the last one holds, which evidently does. The following argument is structural for the composition and implementation of the option pricing methods analyzed later in this study. So, if one would use the Feynman – Kac formula (similarly as in Karatzas and Shreve in 1988), he would end up that the derivatives price is simply the expected payoff discounted at the risk-free rate of return.

The next conclusion is that we can substitute investors total expected rate of return on the stock, namely μ in (1) with r, and thus use the externally determined and in this study concerned as constant, risk-free interest rate for option valuation.

Chapter 5
Methods for simple option pricing

5.1 The Black-Scholes-Merton framework

Although the purpose of this study is to cover numerical methods on options pricing, that is methods that don't provide a closed-form analytical solution for the price of an option derivative, it is however meaningful to describe the Black-Scholes-Merton (B-S) framework, for reasons of consistency, as it will help us to better understand the next methods and of course because we will be mentioning B-S framework a lot of times in the next chapters.

First we cite the assumptions underlying the B-S world:

- The underlying asset's price (here a stock's price) follows the process in (1)
- The short selling of securities with full use of proceeds is permitted
- There are no transaction costs or taxes and all securities are perfectly divisible
- There are no dividends during the life of the derivative. ($\delta=0$ in (1)).
- We are in an arbitrage free world
- We are in a risk-neutral world
- Security trading is continuous

We have assumed that the changes in the value of a derivative, dependent on a stock and on time t, follow the process

$$dV = \left(\frac{\partial V}{\partial S} \mu S + \frac{\partial V}{\partial t} + \frac{1}{2} \left(\frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 \right) \right) dt + \frac{\partial V}{\partial S} \sigma S dW_t$$

and that the process followed by the changes of the price of a share is

$$\frac{dS}{S_t} = (\mu - \delta)dt + \sigma dW_t \xrightarrow{\delta=0} \frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$

$$\text{or } dS = \mu S\, dt + \sigma S dW_t$$

where we have to notice that dW_t is the uncertainty factor, same for the two equations. We can construct a portfolio with a short position on one derivative and a long one on a number of shares. We then construct an equation for the changes in the value of this portfolio. So let Π be the value of the portfolio, V be the derivative and S the share. The weights of the portfolio are

-1 derivative

$\frac{\partial V}{\partial S}$ shares

The value of the portfolio is

$$\Pi = -V + \frac{\partial V}{\partial S} S$$

and the changes in its value would be the sum of the changes of the value of its components, that is

$$d\Pi = -dV + \frac{\partial V}{\partial S} dS$$

and by substitution we get

$$\mathrm{d}\Pi = \underbrace{-\left(\frac{\partial \mathrm{V}}{\partial \mathrm{S}}\mu S + \frac{\partial V}{\partial T} + \frac{1}{2}\left(\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2\right)\right)dt - \frac{\partial \mathrm{V}}{\partial \mathrm{S}}\sigma S dW_t}_{\mathrm{dV}} + \frac{\partial \mathrm{V}}{\partial \mathrm{S}}\underbrace{(\mu S dt + \sigma S dW_t)}_{dS}$$

$$= \left(-\frac{\partial V}{\partial t} - \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2\right)dt$$

The uncertainty factor dW_t has been simplified, and that's why the portfolio is riskless, returning the risk-free rate of return. So

$$d\Pi = r\Pi dt$$

and substituting for dΠ and Π we get the partial differential equation (PDE) of Back and Scholes:

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} - rV = 0$$

which solved under the restrictions of the payoff functions and two simple transformations of them gives the famous Black-Scholes formula for the pricing of European options. The restrictions and their transformations are:

for call options : $V(S,T) = \max(S_T - K, 0)$

$V(0,T) = 0$

and for put options : $V(S,T) = \max(K - S_T, 0)$

$V(0,T) = K$

where T is the expiration time. The B-S pricing formulas are:

$$c = S_0 N(d_1) - Ke^{-rT} N(d_2)$$

for call options and

$$p = Ke^{-rT} N(-d_2) - S_0 N(-d_1)$$

for put options. S_0 is the share price at the time the option contract is signed and function N(.) is the cumulative probability distribution function for a standardized normal distribution, of the probability that the option expires at the in-the-money region. d_1, d_2 and N(x) are

$$d_1 = \frac{\ln(S_0 / K) + (r + \sigma^2 / 2)\mathrm{T}}{\sigma\sqrt{\mathrm{T}}}$$

$$d_2 = \frac{\ln(S_0 / K) + (r - \sigma^2 / 2)\mathrm{T}}{\sigma\sqrt{\mathrm{T}}} = d_1 - \sigma\sqrt{\mathrm{T}}$$

$$N(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} e^{-\frac{1}{2}s^2} ds$$

So the only think one has to do is to calculate d_1 and d_2 , and then by using a software package, or simply Microsoft's Excel, calculate the integrals of N(.). B-S model can be modified to adapt expected dividends from the share underlying the option. We just have to subtract the present value of the expected dividends from the current price of the share, and then implement the formulas.

The simplicity of this formula and the fact that it was the first, to our knowledge, closed-form solution in the option pricing theory, made it very popular. But this method has some important drawbacks, conveyed by its strong assumptions:

- We assume perfect structured markets, with no transaction costs and no taxes, which one doesn't have to be Warren Buffett to deny it.
- We assume a constant volatility through the life of the option, which of course is not the true, shown by basic analysis of the market data. In this study we will try to break this assumption, and assume a stochastic volatility for the underlying asset.
- We also assume a stable risk-free rate, which doesn't hold, but this topic will not be covered in this study.
- We finally assume perfectly divisible securities and continuous trading, which is not the case.

We also observe that the price of the option does not in any way depend on the risk preferences of the investors, which are part of their total expected return μ. This observation agrees with the assumption of risk-neutral valuation. Usage has also been done of the arbitrage free pricing assumption in order to reach the PDE of Black and Scholes. Finally the assumption of continuous lognormal distribution of the changes in stocks price, enabled us to select a distribution, with known and easy to compute probability distribution function (pdf).

Due to the above strong assumptions, the B-S formula only works well in a few categories of European options. It cannot price options with American characteristics, or options with multiple underlying assets. As Heston (1993) shows, B-S model, effectively prices at-the-money options, with relative small maturities, up to 3 years maximum.

Results for option pricing with the B-S formula will be cited along with other numerical methods' results at the end of the chapter.

5.2 Finite difference approach

Finite difference methods value a derivative by providing numerical results for the fundamental PDE that the derivative satisfies, using a discrete-time, discrete-state approximation. We will soon make clear what this means. The fact that these methods rely on solving differential equations, makes them more adaptive than other methods, in the sense that we can price more complicated derivatives, like ones with stochastic volatility, jump-diffusion processes and more than one underlying asset. Their major drawback is that we cannot price options with more than 3 underlying assets, as solutions for these types of differential equations are not yet available. This drawback was also pointed out by Barraquand and Martineau (1995). There are two approaches in finite difference approach: a) implicit finite difference method and b) explicit finite difference method.

5.2.1 Implicit finite difference method

Suppose that we divide life T of an option in N equal intervals of length $\Delta t = T/N$, having N+1 times varying 0, Δt, $2\Delta t$,…,T. Suppose we do the same with the asset's (i.e. a stock) price S. The only difference is that in this case we first consider a price level S_{max} , such that the value of a put option at this level is 0, and that of a call option is the strike price. So we have M+1 equal stock spaces, each one with a length of $\Delta S = S_{max}/M$. Imagine a grid, defined by the time and the stock points, having (N+1)(M+1) dimensions, with time at the horizontal axis and stock price at the vertical. The (i,j) point of the grid, corresponds to time $i\Delta t$ and stock price $j\Delta S$, and the variable $f_{i,j}$ is the option's value at the (i,j) point. The analysis made so far is also applicable to the explicit finite difference method. We also assume arbitrage-free pricing, risk-neutral world, and that the stock price follows the same process as in the B-S framework. Finally, the stock is paying a dividend rate δ.

A forward difference approximation for the interior point (i,j) is

$$\frac{\partial f}{\partial S} = \frac{f_{i,j+1} - f_{i,j}}{\Delta S} \qquad (1)$$

and a backward approximation for the same point would be

$$\frac{\partial f}{\partial S} = \frac{f_{i,j} - f_{i,j-1}}{\Delta S} \qquad (2)$$

and for the point (i,j+1), the backward approximation is

$$\frac{\partial f}{\partial S} = \frac{f_{i,j+1} - f_{i,j}}{\Delta S} \qquad (3)$$

We can also use a symmetrical formula, written

$$\frac{\partial f}{\partial S} = \frac{f_{i,j+1} - f_{i,j-1}}{2\Delta S} \qquad (4)$$

As far as $\partial f / \partial t$ is concerned, we use a forward difference approximation for the point (i,j) so that

$$\frac{\partial f}{\partial t} = \frac{f_{i+1,j} - f_{i,j}}{\Delta S} \qquad (5)$$

If we wanted to make a finite approximation for the changes in the option value, between two points, at the same timestep, we would need an approximation for $\partial^2 f / \partial S^2$ at the point (i,j), that is

$$\frac{\partial^2 f}{\partial S^2} = \left(\frac{f_{i,j+1} - f_{i,j}}{\Delta S} - \frac{f_{i,j} - f_{i,j-1}}{\Delta S} \right) \Big/ \Delta S$$

or

$$\frac{\partial^2 f}{\partial S^2} = \frac{f_{i,j+1} + f_{i,j-1} - 2 f_{i,j}}{\Delta S^2} \qquad (6)$$

Substituting (4), (5), (6) in the B-S differential equation (assuming a dividend rate is paid) we get

$$\frac{f_{i+1,j} - f_{i,j}}{\Delta t} + (r-\delta) j\Delta S \frac{f_{i,j+1} - f_{i,j-1}}{2\Delta S} + \frac{1}{2}\sigma^2 j^2 \Delta S^2 \frac{f_{i,j+1} + f_{i,j-1} - 2f_{i,j}}{\Delta S^2} = rf_{i,j} \qquad (7)$$

with j=1,2,…,M-1 and i=0,1,2,…,N-1. Rearranging terms, we get

$$a_j f_{i,j-1} + b_j f_{i,j} + c_j f_{i,j+1} = f_{i+1,j} \qquad (8)$$

with

$$a_j = \frac{1}{2}(r-\delta) j\Delta t - \frac{1}{2}\sigma^2 j^2 \Delta t$$

$$b_j = 1 + \sigma^2 j^2 \Delta t + r\Delta t$$

$$c_j = -\frac{1}{2}(r-\delta) j\Delta t - \frac{1}{2}\sigma^2 j^2 \Delta t$$

(8) constitutes a system of N+1 linear equations, with N+1 unknowns, which can be solved with the Gaussian method of elimination, that is by constructing a triagonal matrix with the coefficients of the unknowns, multiplying it with its inverse, and get the vector of the solutions. But before doing the above calculations, we must first cite some arguments that will help us solve this system. First of all a_j , b_j and c_j are (1xM) vectors that can be calculated straight away. Secondly we know that the last column of the grid, that is $f_{N,j}$, equals the payoff of the option and by transforming the payoff function we can get the same boundaries with those used for the calculation of the B-S PDE. Namely for call options

$$f_{N,j} = \max(j\Delta S - K, 0)$$

$$f_{i,0} = 0 \text{ and}$$

$$f_{i,M} = K$$

and for put options

$$f_{N,j} = \max(K - j\Delta S, 0)$$

$$f_{i,0} = K \text{ and}$$

$$f_{i,M} = 0$$

where j = 0,..,M and i = 0,…,N, and when j = M we get S_{max}. Now we have a N-1 system with the same number of unknowns, for each timestep. We start by calculating the unknowns of T-Δt, that is the elements of the penultimate column of the grid, with the use of the known elements of the last column. We use (8), which for each element of the penultimate column gives

$$a_j f_{N-1,j-1} + b_j f_{N-1,j} + c_j f_{N-1,j+1} = f_{N,j}$$

After calculating all elements of the penultimate column, we can proceed to the calculation of the elements corresponding to time T-2Δt, a timestep prior from T-Δt. As one can realize, we use the elements of the column right to the one we want to cover, that is the now known elements of the penultimate column. We apply this backward looking technique until T=0 is reached, and the option price is the element that corresponds to S_0. Three important observations must be made. First of all, if the option is American, we compare the solutions of the system corresponding to each timestep, with the value of early exercise, the value of the corresponding elements of the last column. We use the results of the comparison as the known elements with which we proceed to the solution of the next system. Secondly, we choose S_{max} and M, such that jΔS passes through S_o. Finally we have used the dividend rate δ in the formulas, because finite difference methods do not require zero dividends, as we had assumed for the derivation of the B-S formula. These observations also apply for explicit finite difference method.

5.2.2 Explicit finite difference method

An easier way to deal with the finite difference approach for option pricing is to regard the values of $\partial f/\partial S$ and $\partial^2 f/\partial S^2$ for the point (i,j) as being equal with those for the point (i+1,j). Equations (4) and (6) and (7) then become

$$\frac{\partial f}{\partial S} = \frac{f_{i+1,j+1} - f_{i+1,j-1}}{2\Delta S}$$

$$\frac{\partial^2 f}{\partial S^2} = \frac{f_{i+1,j+1} + f_{i+1,j-1} - 2f_{i+1,j}}{\Delta S^2}$$

and

$$\frac{f_{i+1,j} - f_{i,j}}{\Delta t} + (r-q)\, j\Delta S\, \frac{f_{i+1,j+1} + f_{i+1,j-1}}{2\Delta S}$$
$$+\frac{1}{2}\sigma^2 j^2 \Delta S^2\, \frac{f_{i+1,j+1} + f_{i+1,j-1} - 2f_{i+1,j}}{\Delta S^2} = rf_{i,j}$$

Equation (8) now becomes

$$f_{i,j} = a_j^* f_{i+1,j-1} + b_j^* f_{i+1,j} + c_j^* f_{i+1,j+1}$$

and a_j, b_j, and c_j are written

$$a_j^* = \frac{1}{1+r\Delta t}\left(-\frac{1}{2}(r-\delta)\, j\Delta t + \frac{1}{2}\sigma^2 j^2 \Delta t\right)$$

$$b_j^* = \frac{1}{1+r\Delta t}\left(1-\sigma^2 j^2 \Delta t\right)$$

$$c_j^* = \frac{1}{1+r\Delta t}\left(\frac{1}{2}(r-\delta)\, j\Delta t + \frac{1}{2}\sigma^2 j^2 \Delta t\right)$$

This is a much easier approximation as the only think one has to do is, under the boundaries developed earlier, to calculate each element of the penultimate column of the grid, by the known elements of the final grid corresponding to a distance of one ΔS, above and beneath the unknown element, and the element of the same jΔS. Then by using the two first observations made for the implicit approximation, one calculates the option price of consideration. Though easier in its use, explicit finite difference approximation has a stability issue as cited by Jouni Kerman (2002). For a given step in the stock price, of size ΔS, Δt cannot be larger than $\frac{1}{2}\Delta S^2$. Especially, changing ΔS by a factor of ζ, one has to multiply the number of timesteps by ζ^2. Instead of this, we assumed in our compiling the same ratio of timesteps and stock steps as Longstaff and Schwartz (2004, p.126) did, and the method gave accurate results.

Results from the implementation of implicit and explicit finite difference methods will be cited in the sequel, along with results obtained from Black-Scholes pricing framework and binomial tree option pricing.

5.3 Lattices – The binomial tree approach

Lattice approaches were first introduced in Parkinson (1977) and Cox, Ross and Rubinstein (1979). They actually use the same discrete-time, discrete-state approximation with finite difference approach. The easiest lattice to handle is the binomial tree introduced by Cox et al. (1979), which can value simple European and American options. By simple options, we mean options with maximum three underlying assets, and assuming the underlying asset's price being the only stochastic parameter for the valuation of the option. Computational time and the necessary computer memory grow exponentially with the number of the underlying assets, an effect that is named as the curse of dimensionality. We will explain later in the chapter why this happens.

5.3.1 The one timestep approach

Suppose we have an option, with a share as the underlying asset, whose price follows the same stochastic process, that is

$$dS = \mu S\, dt + \sigma S dW_t$$

and we take S_0 as the current stock price. Instead of directly solving the above stochastic differential equation, lattices assume the existence of a random variable X, with X = xi, i = 1,2,…,m and probabilities pi for each xi. Taking a discrete-time approximation for the stock price, this method gives a distribution of asset prices for different values of m. If m=2, then we have a binomial distribution and for example S_0 can take values equal to S_0u and S_0d. The percentage increase in the tock price when there is an up movement is *u*-1, and the percentage decrease when there is a down movement is 1-*d*. u and d can either be given constants, where d=1-u, or they can be matched with the volatility parameter, in which case $u = e^{\sigma\sqrt{\Delta t}}$ and *d=1/u*. From each price, from each node, we get two different prices, when moving from one timestep to another. If m=3, we get a trinomial distribution, with three different prices rising from each "node" and so on. This study will cover the case of m=2, the case of the binomial tree. To continue with our description of the method, u and d are the rate of returns that the stock can provide after each node. If for instance S_0 = \$20 and u=0.1, then the price of the stock can move to either \$22 or \$18. We assume u and d to be constant so at the next timestep it can either move to \$24.2, \$19.8 or \$16.2. The price of \$19.8 is a downward movement from \$22 and also an upward movement of the same percentage from \$18.

Suppose that we want the same riskless portfolio we structured in the B-S framework, based on the assumptions of zero dividend rate, arbitrage free pricing and the risk neutral pricing hypothesis. We take a long position in a number of Δ shares underlying the option and a short position in the option. After an upward movement from the first timestep, we assume that the intrinsic value of the option is f_u and after a downward one, the value of the option is assumed to be f_d. In order for the portfolio to be riskless, its value after an upward movement must equal its value after a downward movement, and the correspondent values are written

$$S_0 u\Delta - f_u = S_0 d\Delta - f_d$$

We calculate Δ, in order for the portfolio to be riskless, that is

$$\Delta = \frac{f_u - f_d}{S_0 u - S_0 d} \qquad (1)$$

Now we use the formula of the cost for setting up the portfolio, in order to derive the value of the option, at present. The cost is

$$S_0\Delta - f$$

and it must be equal to the present value of the portfolio after for example an upward movement of the share's price, that is

$$S_0\Delta - f = (S_0 u\Delta - f_u)e^{-rT}$$

Then we solve for *f,* we substitute for Δ from (1), and with simple transformations, we end up writing the value of the option as

$$f = e^{-rT}\left[pf_u + (1-p)f_d\right] \qquad (2)$$

where

$$p = \frac{e^{rT} - d}{u - d} \qquad (3)$$

is the probability of an upward movement for the stock price, and 1-p is the probability for a downward one. Equation (2) gives the value of the option when we use only one timestep for the stock price movements, which is not the usual case. In order for the binomial tree approach to converge to the options' values given by B-S, the number of timesteps must tend to infinity, or reach a number like ten or fifty thousands, depending on the compilation time we want to dispose. Next we will show a time approximation of two steps, and then cite the approach for the construction of a multiperiod tree.

5.3.2 Two timesteps approach

Instead of using one timestep equal to the maturity of the option, we now use a time step equal to Δt, and equations (2) and (3) become

$$f = e^{-r\Delta t}\left[pf_u + (1-p)f_d\right] \qquad (4)$$

and

$$p = \frac{e^{r\Delta t} - d}{u - d} \qquad (5)$$

At the upper and down nodes, the corresponding option values are

$$f_u = e^{-r\Delta t}\left[pf_{uu} + (1-p)f_{ud}\right] \qquad (6)$$

and

$$f_d = e^{-r\Delta t}\left[pf_{ud} + (1-p)f_{dd}\right] \qquad (7)$$

Substituting (6) and (7) into (4), we get the option value

$$f = e^{-2r\Delta t}\left[p^2 f_{uu} + 2p(1-p)f_{ud} + (1-p)^2 f_{dd}\right]$$

where p^2, 2p(1-p), and $(1\text{-}p)^2$ are the probabilities of two upward moves in the stock price, one upward and one downward move and two downward moves accordingly. Figure 1 shows a binomial tree with one timestep and two timesteps to expiration.

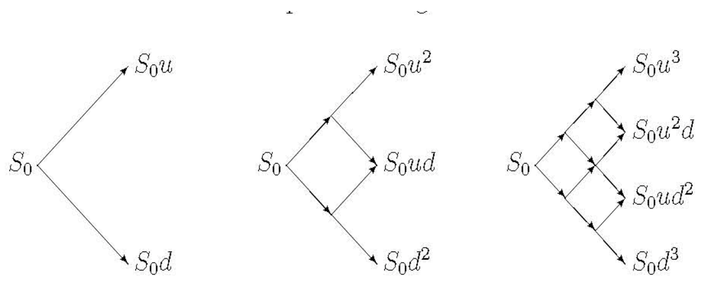

Figure 1. Stock price process for 3 subsequent timesteps

An important observation is that due to the risk neutral pricing assumption, the option value, as affected by p, is not affected by the risk preferences of the investors, namely μ. The stock prices have already absorbed the risk preferences when we simulate their movements by this tree, so there is no need for the probabilities p and (1-p) and so on, to take into account those risk preferences.

If we want to use more than 2 timesteps, we only have to make two nodes leaving each node, as we move forward regarding timesteps, until we reach the final timestep, and the corresponding final nodes. Then we calculate the intrinsic value of the option at each of the final nodes, and we begin moving backwards, one timestep at a time. For each timestep, we use (4) to calculate the option value at each node, and we proceed the same way until we reach the initial node. If the option is American, after the calculation of the option's value from (4), we compare the results with the corresponding intrinsic value of the option, and we use the outcome of the comparison in order to calculate the option prices one node backwards. This comparison is done for the same reason that we did it in the finite difference method, and that is because we want in each timestep, and each node, to find if the early exercise of the option is optimal or not. Results of the binomial tree approach are presented at table 1, along with results of the B-S framework and finite difference methods.

To conclude let us say a little words for the curse of dimensionality, and its effects on the usage on lattices. For the simplest case of the binomial tree we see that each node is "separated" into two nodes, when we are talking of course for an option with one underlying asset. Anyone could imagine, simply by looking at figure 1, what the tree would like if we had two underlying assets, that is four nodes leaving each node. In general, the nodes leaving each node on a lattice are m^L, L being the number of stochastic factors. Needless to describe what the situation would be if we had a tree of m = 4, three underlying assets and 10.000 timesteps. We see that the nodes are increasing exponentially with the number of the underlying assets, and that costs a lot in terms of computational time and of the means required to estimate the value of the option. That's the essence of the curse of dimensionality and that's why lattices are avoided in the valuation of multi-factor options. Also, as Barraquand and Martineau (1995) have shown, simply by using the binomial tree of Cox, Ross and Rubinstein, or the finite difference approximation, we cannot valuate options with more than three underlying assets. That's where simulation methods come in and help us deal with this problem, as we will thoroughly discuss in the next chapter.

Now we shall cite table 1 with the results obtained from the methods we covered so far. We value a European put option giving two different current stock prices, S_0 = \$40 and S_0=\$36, two times to expiration, T = 1 and T=1/2, and two volatility rates, σ = 0.4 and σ=0.2. The strike price is K = \$40, and the risk-free rate of return is r = 0.06. We assume 2000 timesteps for the binomial trees. For the implicit finite difference method, we have assumed 2000 steps for the stock price and 50 timesteps that the American option can be exercised. For the explicit approach we assumed 200 steps for the stock price and 8000 timesteps, due to the instable nature of this method. We have also used the control variate technique in the results obtained by the explicit approach, using the price of a similar European option obtained by the use of explicit finite difference approach as control variate, and Black-Scholes framework as an analytic solution. A detailed description of the control variate technique will be given in the next chapter.

Table 1

Black-Scholes, Finite difference approaches and binomial tree pricing

	Black-Scholes	Binomial European	Binomial American	Finite Diffrence Implicit Approach		Finite Diffrence Explicit Approach	
				European	American	European	American
	Price	Price	Price	Price	Price	Price	Price
s0=40 T=1 σ=0.4	5.05962	5.05885	5.31792	5.04698	5.29048	5.10884	5.32085
s0=36 T=1 σ=0.4	6.71140	6.71118	7.10897	6.70245	7.08102	6.75919	7.11119
s0=40 T=1 σ=0.4	5.05962	5.05885	5.31792	5.04698	5.29048	5.10884	5.32085
s0=40 T=1/2 σ=0.4	3.86569	3.86513	3.97775	3.85803	3.96196	3.90376	3.97910
s0=40 T=1 σ=0.4	5.05962	5.05885	5.31792	5.04698	5.29048	5.10884	5.32085
s0=40 T=1 σ=0.2	2.06640	2.06600	2.31943	2.06137	2.30280	2.08613	2.32199

Chapter 6
Monte Carlo simulation in option pricing

Simulation has been a promising alternative in the last decade to value American options and more generally options with multiple state variables and under general stochastic processes. Simulation techniques were first introduced in option pricing by Boyle (1976). This paper simulated a European option, with asset price as the only stochastic factor. There has been a great deal of research since then, and today we can accurately value, most of multifactor, path-dependent options, with general stochastic processes, say stock price together with volatility. The breakthrough was in our opinion the introduction of the least-squares method (LSM) by Longstaff and Schwartz in 2001. This method introduced a simple backward-looking simulation algorithm to value options with American characteristics. In this study, we will cover the simulation techniques, as they were seen in the papers of Boyle and Longstaff and Schwartz, for plain vanilla options, both European and American. We will also examine the contribution of Stentoft (2004a), in the application of LSM.

6.1 The Monte Carlo Simulation method

To begin with, we assume an arbitrary function g(y) and its probability density function (pdf) f(y) with $\int_A f(y)dy = 1$, A being the range of integration. For the calculation of an estimate $\hat{g}$, of $\bar{g}$, we peak in random, a number *n* of sample values *(yi)*, from the pdf, which is the same as calculating the definite integral $\int_A g(y)f(y)dy = \bar{g}$. That's actually what Monte Carlo simulation does; it simulates *n* number of paths for a function g(y), using values from its pdf. The final calculation is to average the values *g(yi)* obtained, written

$$\hat{g} = \frac{1}{n}\sum_{i=1}^{n} g(yi). \qquad (1)$$

In option pricing g(y) is the process followed by the stock price. So we generate sample paths for the stock price, using an arbitrary number of time intervals, or the definite number of time intervals that the option can be contractually exercised. We end up with a number of possible outcomes for the stock price at the expiration. Then we calculate the intrinsic value of the option and we average the values obtained, using an equation like (1). The option price will be the outcome of (1), discounted at the risk-free rate. So we end up using a more dense formula, written

$$\hat{g} = \frac{1}{n}\sum_{i=1}^{n} e^{-rT} g(yi). \quad (2)$$

That formula works for European options, but what about American ones? We will see what happens with them in a later stage. We will first examine the basic aspects of the simulation techniques. The most important aspect of this technique is its advantages, which we cite.

6.1.1 Advantages and disadvantages

a) Simulation techniques can adopt any desired assumption, i.e. on the process followed by a stochastic factor, like the stock price, or the arbitrage free and the risk neutral assumption, which are actually implemented in (2).

b) We can price options with multiple underlying assets and more than one stochastic factor, i.e. the volatility.

c) Since a Monte Carlo estimator is an average of *n* individual draws of a random variable, then according to the central limit theorem, this estimator has normal distribution for a large number of draws. So we can use the standard deviation as a statistical measure of the uncertainty of the estimator. This is done with the square root of the variance of the estimator, written

$$\hat{\sigma}_n = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n} \hat{g}(yi)^2\right) - \left(\frac{1}{n}\sum_{i=1}^{n} \hat{g}(yi)\right)^2}$$

from which we can derivate the standard error as another measure of uncertainty, written

$$\hat{\varepsilon}_{\mathrm{K}} = \frac{\hat{\sigma}_{\mathrm{K}}}{\sqrt{n}}$$

which tells us that we can calculate the accuracy of the results for each simulation done, and that, accuracy only depends on the number of paths simulated. As far as a number of 100.000 paths can be easily applied by modern computers, we can be rather certain that we have very accurate results.

An alternative measure of the accuracy of our simulated estimators is the Root Mean Square Error (RMSRE), which equals

$$RMSRE = \sqrt{\frac{1}{L}\sum_{j=1}^{L}\left[\frac{g(y) - \hat{g}_j(y)}{g(y)}\right]^2}$$

where L is the number of estimators whose performance we want to investigate, $\hat{g}_j(y)$ is the j-th estimator and g(y) is an option value we use as a benchmark, acquired from an other method, i.e. a binomial tree with 10.000 timesteps.

d) Another advantage of simulation techniques is the possibility of using two global techniques, namely i) the control variate method, and ii) the antithetic variate, as an alternative to the increase of the number of paths $\left(\sqrt{n}\right)$, in decreasing standard error. These methods are described as follows:

   i. *The control variate method.* In this method, apart from using simulation to price the option in consideration, we also run the same simulation, (same number of paths, and same range of timesteps) in order to price a similar option that has an analytic solution. Let A be the option that we want to price, B be the similar option we have in mind, f_A be the amended estimation of the price of the option, f_B be the known analytic solution of B, and f_A^*, f_B^* be the values obtained by simulating the two options. Then f_A is written

$$f_A = f_A^* - f_B^* + f_B$$

The known analytic solution is usually regarded as the one obtained by the B-S formula, and in simulations, we can use this technique in order to improve the accuracy of American options pricing, by simulating the option in consideration as being a European one. In this study the usage of this technique didn't prove to be necessary, mainly because we made use of the antithetic variate technique.

ii) *The antithetic variate technique.* A Wiener process dW_t has two components, the random term ε, and $\sqrt{\Delta t}$. This technique consists of running two similar simulations, and obtaining two different option values. In the first simulation we sample a number of ε_t values from the distribution of ε, and in the second simulation we sample the same values, but in their antithetic form $-\varepsilon_t$. Then we average the two option prices obtained and we end up with a price $\bar{f}$ of the option, written

$$\bar{f} = \frac{f_1 + f_2}{2}$$

f_1 and f_2 being the two prices obtained by the two similar simulations. If $\bar{\omega}$ is the standard deviation of $\bar{f}$, then the standard error of the estimate can be written

$$\hat{\sigma}_n = \frac{\bar{\omega}}{\sqrt{n}}$$

which is much less than the standard error calculated by running two simulations with a sum of *2n* paths. In all the simulations of this study, the use of this technique has taken place, and provided us with very accurate results. Instead of running two simulations, and calculating two prices of the option, we just run one simulation, where we sampled ε_t values from the distribution of ε, and the same number of $-\varepsilon_t$ values.

As far as the disadvantages of Monte Carlo simulation are concerned, the only major disadvantage of the Monte Carlo simulation is the computational time entailed in having a large number of estimators; that is, running a significant number of simulations with different parameters (like the number of paths), used in each one of them. This problem is tackled with the usage of networks of workstations, and with the evolution in the technology of home desktops. Unfortunately we couldn't use such a network for the purposes of our study.

6.2 Implementation of the Monte Carlo Simulation

The first think we do is to solve the SDE describing the known process the stock price follows. An explicit Euler scheme, as cited in Barraquand and Martineau (1995) is given by

$$S_i(t+\Delta t) = S_i(t)\left(\left(r - d_{xi} - \frac{1}{2}k_{ii}\right)(X(t),t)\Delta t + \sum_{j=1}^{n} v_{ij}(X(t),t)\sqrt{\Delta t}\, z_j^t\right)$$

where i is the number of paths we simulate , k_{ii} and v_{ij} are the variance and the volatility of the asset price. z_j^t is the random term of the Wiener process of the SDE, what we cited as ε, and all z_j^t follow independent standard normal distributions for all *j* and *t*. d=T/Δt, is the number of timesteps in [0,T]. One has to draw M x d (M stands for the number of paths) standard normal variates in order to generate M x d-dimensional sample paths $X^1(t),\ldots,X^M(t)$ for al t>0. A more simple form of this equation which we use in all our simulations is

$$S_{[i,j]}(t+\Delta t) = S_0 \exp\left(\left(r - \frac{\sigma^2}{2} - \delta\right) j\Delta t + \sigma\sqrt{\Delta t}\sum_{i=1}^{runs} z[i,1:j]\right) \qquad (3)$$

where i stands for the number of paths simulated, which from now on will declared as runs. j=T/Δt is the number of steps of the simulation. S_0 is the price of the share, the day the option is signed, *exp* stands for exponential, r as always, is the risk-free rate of return, σ^2 is the square of the volatility, the variance of the changes on the prices of the share, δ is the dividend rate, which in our simulations is assumed to be zero, and σ is the volatility in interest. Given the parameters of the option, we generate the matrix, from which to pick the random, independent standard normal values εi, using the same number of εi_s and $-\varepsilon i_s$. Barraquand and Martineau (1995) cite that when the joint process *X(t),* that is the natural logarithm lnS of the price of the share, is assumed lognormal, 10 timesteps are sufficient for security pricing with lognormal underlying assets price processes. The last sentence stands both for European and American options. In our simulations, we have used a number of timesteps equal to 50. As long as runs are concerned, we have used a range of (M) 10.000 to 100.000 runs, with increments of 10.000. Table 2 gives the results obtained, from the valuation of a European put option with: S_0 = \$40, K = \$40, T = 1, σ = 0.4, r = 0.06. The price obtained by a binomial tree with 2.000 timesteps, is used to calculate the bias, and then examine if the latter is statistically significant for two critical regions, that is two different α's, namely 0.05 and 0.01. We use 1-α as a measure of goodness of the estimates. The price of the tree is \$ 5.0588458. Table 3 gives the prices obtained by simulation with the use of 100,000 paths, and the prices of the other methods as well, for a comparison to be feasible and straight.

Table 2

Boyle's simulation

M	price	s.e.	bias	Stat. Sig. 95%	Stat. Sig. 99%
10000	5.16258	0.028	0.104	Yes	Yes
20000	5.08958	0.027	0.031	No	No
30000	5.07106	0.027	0.012	No	No
40000	5.07376	0.027	0.015	No	No
50000	5.06873	0.027	0.010	No	No
60000	5.06882	0.027	0.010	No	No
70000	5.06787	0.027	0.009	No	No
80000	5.06808	0.027	0.009	No	No
90000	5.06855	0.027	0.010	No	No
100000	5.06568	0.027	0.007	No	No

Table 3

Boyle's simulation and other methods for option pricing

	Boyle		Black-Scholes	Binomial	Binomial	Finite Diffrence		Finite Diffrence	
				European	American	Implicit Approach		Explicit Approach	
						European	American	European	American
	Price	Standard Error	Price	Price	Price	Price	Price	Price	Price
s0=40 T=1 σ=0.4	5.06568	0.027	5.05962	5.05885	5.31792	5.04698	5.29048	5.10884	5.32085
s0=36 T=1 σ=0.4	6.71500	0.033	6.71140	6.71118	7.10897	6.70245	7.08102	6.75919	7.11119
s0=40 T=1 σ=0.4	5.06568	0.027	5.05962	5.05885	5.31792	5.04698	5.29048	5.10884	5.32085
s0=40 T=1/2 σ=0.4	3.87017	0.017	3.86569	3.86513	3.97775	3.85803	3.96196	3.90376	3.97910
s0=40 T=1 σ=0.4	5.06568	0.027	5.05962	5.05885	5.31792	5.04698	5.29048	5.10884	5.32085
s0=40 T=1 σ=0.2	2.06959	0.007	2.06640	2.06600	2.31943	2.06137	2.30280	2.08613	2.32199

We note that all the methods described so far, converge to the Black-Scholes solution, as far as European options are concerned. For American options, the results from the binomial tree, and both finite difference approaches are quite close to each other and to the results obtained by simulation, as we will discover shortly.

6.3 The Least Squares Monte Carlo Method

The Least Squares Monte Carlo Method (LSM) was introduced by Longstaff and Schwartz (2001) and was a major breakthrough in the use of simulation for option pricing. Most of the simulation methods till that time could not value options with American characteristics, since they incorporated a forward-looking technique. Each path was followed by a single path after each timestep. That worked for European options, as we discounted the average value of the option at expiration. But we could not value options with discrete exercise dates before expiration, because there was not an easy to apply method in order to check in each timestep whether the early exercise of the option would be optimal or not, and obtain the optimal early exercise strategy. As Stentoft (2004) states, "this exercise strategy would have to be calculated recursively, but when simulation techniques are used at any time along any of the paths there is only one future path, and using these values would lead to biased results". Tilley (1993), Barraquand and Martineau (1994) and Broadie and Glasserman (1997) have proposed some techniques to tackle this problem. The greatest evolution took place with the introduction of the LSM, which due to its simplicity and its ease to price options with complicated payoff functions, gained a lot of ground in the option pricing theory.

The basic idea behind the LSM method lies on the fact that the holder of an American option compares the payoff from immediate exercise with the expected payoff from continuation, at each one of discrete number of times the option can be exercised. The payoff from immediate exercise is known to the holder and it is the option's intrinsic value. What is not known is the expected payoff if she doesn't exercise her option, and keeps it in life. This payoff is determined as the conditional expectation function and LSM is the way to determine it. We regress the ex post realized payoffs from continuation, on a constant and a set of functions of the values of the state variables, using the least-squares method of regression. Ex post realized payoffs are not determined as a simple discount of the next period's payoffs. Using actual realized payoffs means that the maximum is taken over each path, over each and all exercise times. The fitted value from this regression gives us an estimate of the conditional expectation function for each exercise date we apply the regression. This way we obtain a complete specification of the optimal exercise strategy along each path. Then we are able to set up the optimal stopping rule, which tells us when the option must be exercised for each path of the simulation. According to the stopping rule, we discount the option values of each path, and by averaging the discounted value, we finally obtain the value of the option in interest.

We will cite the description of Rodrigues and Armada (2006), for a mathematical exposure of the LSM. We value an American option, whose only stochastic factor is the stock price, which in turn follows the known process, described by the SDE

$$dS = (\mu - \delta)S\ dt + \sigma S dW_t$$

which with the assumptions of zero dividend rate, and risk-neutral pricing, becomes

$$dS = rS\ dt + \sigma S dW_t$$

The value of an American option that can be exercised in a discrete number of periods in the interval [0,T], or [t,T], and whose payoff function is $\Pi(t,S_t)$, can be expressed as

$$F(t,S_t) = \max_{\tau}\left\{\mathrm{E}_t^*\left[e^{-r(\tau-t)}\Pi(\tau,S_\tau)\right]\right\} \qquad (4)$$

where τ is the optimal stopping time $(\tau \in [t,T])$ and $\mathrm{E}_t^*[\cdot]$ is the risk-neutral expectation, conditional on the information available at t. Longstaff and Schwartz stated that the optimal stopping time can be obtained using the following Bellman equation

$$F(t_n, S_{t_n}) = \max\left\{\Pi(t_n, S_{t_n}), e^{-r(t_{n+1}-t_n)}E_{t_n}^*\left[F(t_{n+1}, S_{t_{n+1}})\right]\right\} \qquad (5)$$

deriving the continuation value as

$$\Phi(t_n, S_{t_n}) = e^{-r(t_{n+1}-t_n)}E_{t_n}^*\left[F(t_{n+1}, S_{t_{n+1}})\right] \qquad (6)$$

When expiration is reached, option is no longer in life, so its continuation value is zero, written

$$\Phi(T, S_T) = 0 \qquad (7)$$

Beginning from T, and moving backwards, one timestep at a time, we examine whether early exercise of the option is optimal; this occurs when the conditional on information available in time t_n continuation value is less than or equal to the ex post realized payoff from the option, which as already stated, can be calculated at any path, of any timestep. This rule can be written as follows

$$\text{if } \Phi\left(t_{n,} S_{t_n}(\omega)\right) \leq \Pi\left(t_n, S_{t_n}\right) \quad \text{then} \quad \tau(\omega) = t_n \qquad (8)$$

where *τ(ω)* stands for optimal stopping time. When condition (8) holds, $\tau(\omega)$ is updated, because the option can be exercised only once. The value of the option is calculated by averaging the values of all (K) paths

$$F(0, x) = \frac{1}{K}\sum_{\omega=1}^{K} e^{-r\tau(\omega)} \Pi\left(\tau(\omega), S_{\tau(\omega)}(\omega)\right) \qquad (9)$$

What is missing for the calculation of the option's price is to figure out how the expectation value in each path can be found, and that's what the major contribution of Longstaff and Schwartz has been.

We will first state the formula we use to express the expectation, or continuation value that we are looking for and we will explain the ingenious thought of Longstaff and Schwartz, supporting the LSM method. So, let $\Pi(t, s, \tau, \omega)$ be the cash flow we get from the *ω*-path, if we exercise the option in an optimal manner at the timestep $s(t < s < T)$, with the intuitive assumption that it hasn't been exercised at or before time t. The expectation or continuation value is written

$$\Phi\left(t_n, S_{t_n}\right) = E_{t_n}^{*}\left[\sum_{i=n+1}^{N} e^{-r(t_i - t_n)} \Pi\left(t_n, t_i, \tau, \cdot\right)\right] \qquad (10)$$

with:

$$\Pi(t, s, \tau, \omega) = \begin{cases} \Pi(s, S_s(\omega)) \_\_\_ if \tau(\omega) = s \\ 0 \_\_\_\_\_\_\_\_\_ otherwise \end{cases} \qquad (11)$$

Longstaff and Schwartz state that Φ belongs to a Hilbert space L^2, which gives tremendous potentials for its calculation. Due to this ability, Φ can be represented by a countable orthonormal basis and the conditional expectation can be expressed by a linear combination of the elements of the basis, $\Phi(t,S_t)=\sum_{j=1}^{\infty} e^{-r(t_i-t_n)}\Pi(t_n,t_i,\tau,\cdot)$ The continuation value can be calculated using the first $J \prec \infty$ basis: $\Phi^J(t,S_t)=\sum_{j=1}^{J}\phi(t)L_j(t,S_t)$, with *φ(t)* estimated by a least squares regression. The continuation value estimated by the regression is then used to compute the optimal stopping time

$$\hat{\Phi}^J\left(t_n,S_{t_n}\right)=\sum_{j=1}^{J}\hat{\phi}(t)L_j\left(t_n,S_{t_n}\right) \qquad (12)$$

We will now state some significant characteristics of the LSM algorithm, and some implementation issues.

- First of all, the least squares method of regression is simple, very efficient and consumes logical amounts of computational time.
- Only the paths of the in-the-money area are included in the calculation of the option price. As Longstaff and Schwartz state, this significantly increases the efficiency of the algorithm and decreases the computational time.
- The elements constituting the orthonormal basis, are called the basis functions, which in turn are sets of polynomials. Such polynomials include the Laguerre, Chebyshev, Legendre and Shifted Legendre polynomials, the Hermite, Gegenbauer and Jacobi polynomials. We will discuss further on polynomials when we examine the work of Stentoft (2004) on the LSM.
- In our simulations, we value an at-the-money American put option, with one underlying factor, i.e. the stock's price. We also use five different seeds for the generation of the stock's price matrix. The latter is an Mxd dimensional matrix, M being the number of runs, or paths, and d=maturity/dt being the number of timesteps.
- Extra attention should be given to the use of polynomials like Laguerre ones, which include exponential terms of the state variable. Imagine an exponential of a stock's price which equals 40. This is 2.354^{17}, and a use of such numbers

will lead to underflows of the mathematical processor. To avoid this problem, we use the technique of normalization, in which we divide all cash flows and prices by the strike price. In the end of the algorithm, we multiply the matrix of the cash flows, with the strike price, and only then we discount them, leaving this way the price of the option unaffected.

6.4 Stentoft's work on LSM

Stentoft (2004), has stated that there are two kinds of biases in the use of the LSM algorithm:

- an approximation bias as the conditional expectation function is estimated. This leads to a low bias, which should be vanished when the number of regressors tends to infinity. This bias can be written as

$$\Phi(t, S_t) \approx \hat{\Phi}^J(t, S_t)$$

- The stochastic error as a result of the Monte Carlo Simulation, written

$$F(0, x) \approx \frac{1}{K} \sum_{\omega=1}^{K} e^{-r\tau(\omega)} \Pi\left(\tau(\omega), S_{\tau(\omega)}(\omega)\right)$$

  This type of bias occurs because we use the same paths to calculate the conditional expectation function and the value of the option. This leads to a high bias of the estimation and should be vanished as the number of paths increases.

- For reasons of complicity we will also mention a third type of bias, found in Rodrigues and Armada (2006), which is the discretization bias, as a result of restricting the exercise opportunities of an American option to a finite set of *M* dates.

Stentoft worked on the first two types of biases. Especially he tested n argument made by Longstaff and Schwartz, saying that we should increase the number of regressors, until we get the most accurate results possible, or in the same manner, until the price estimated stops increasing. He conducted simulations with an increasing number of regressors and paths, and using alternative families of polynomials, others than the Laguerre polynomials used by Longstaff and Schwartz. These families include (weighted) Laguerre, General Chebyshev, and Shifted Legendre polynomials (the characteristics of these polynomials are cited in the appendix of the chapter). The reason why he tested polynomials like General

Chebyshev, and shifted Legendre is that these polynomials are orthonormal in the space [0,1] and (0,1) accordingly, instead of the space $[0,\infty)$, that Laguerre polynomials satisfy. Except from the polynomials Stentoft used, we also used monomials, in order to show that simple powers of the state variables can give pretty accurate results, compared with other, more complicated polynomials. He also made simulations with a range of paths from 10.000 to 100.000 paths, with increments of 10.000, using K=1,2,3,4,5 regressors.

We value the same option with Boyle's simulations: an American put option with: S_0 = \$40, K = \$40, T = 1, σ = 0.4, r = 0.06, where as usual S_0 is the current price of the stock, which is considered the only underlying factor, K is the strike price, T is the time to expiration of the option, σ is the volatility, and r is the risk-free interest rate. We also used the same number of paths and regressors as Stentoft did. Tables 4,5,6 and 7, show the results from simulations done using Laguerre, General Chebyshev, Shifted Legendre polynomials, and monomials. In the tables, M stands for the number of runs, and k for the number of regressors. As a benchmark we use the price given by a binomial tree with 10.000 timesteps, which is 5.3182198. We also show the standard error of the estimate, for each simulation, and indicate whether this is statistically significant for two sizes of critical regions, that is two different α's, namely 0.05 and 0.01. We use 1-α as a measure of goodness of the estimates. In the appendix we cite a step-by-step approach to the simulation of the LSM algorithm.

Table 4

Price estimates from the LSM algorithm using various numbers of Laguerre Polynomials

	k=1			Stat. Sig.	Stat. Sig.	k=2			Stat. Sig.	Stat. Sig.	k=3			Stat. Sig.	Stat. Sig.	k=4			Stat. Sig.	Stat. Sig.	k=5			Stat. Sig.	Stat. Sig.
M	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%
10000	5.26933	0.021	-0.049	Yes	No	5.29672	0.020	-0.022	No	No	5.31818	0.020	0.000	No	No	5.29817	0.019	-0.020	No	No	5.29017	0.021	-0.028	No	No
20000	5.27383	0.021	-0.044	Yes	No	5.28940	0.020	-0.029	No	No	5.31185	0.021	-0.006	Yes	No	5.34768	0.019	0.029	No	No	5.30228	0.021	-0.016	No	No
30000	5.27654	0.021	-0.042	Yes	No	5.29462	0.020	-0.024	No	No	5.31142	0.020	-0.007	Yes	No	5.31467	0.020	-0.004	No	No	5.31326	0.003	-0.005	No	No
40000	5.27710	0.021	-0.041	Yes	No	5.29292	0.020	-0.025	No	No	5.30853	0.020	-0.010	Yes	No	5.29874	0.019	-0.019	No	No	5.31536	0.003	-0.003	No	No
50000	5.27978	0.021	-0.038	No	No	5.30353	0.020	-0.015	No	No	5.31225	0.020	-0.006	Yes	No	5.31486	0.020	-0.003	No	No	5.28467	0.021	-0.034	No	No
60000	5.28343	0.021	-0.035	No	No	5.29636	0.020	-0.022	No	No	5.30862	0.020	-0.010	Yes	No	5.32020	0.021	0.002	No	No	5.28343	0.021	-0.035	No	No
70000	5.28894	0.021	-0.029	No	No	5.29632	0.020	-0.022	No	No	5.30778	0.020	-0.010	Yes	No	5.31150	0.020	-0.007	No	No	5.28011	0.021	-0.038	No	No
80000	5.28429	0.021	-0.034	No	No	5.29455	0.020	-0.024	No	No	5.30634	0.020	-0.012	Yes	No	5.32342	0.017	0.005	No	No	5.31759	0.002	-0.001	No	No
90000	5.28656	0.021	-0.032	No	No	5.29224	0.020	-0.026	No	No	5.30510	0.020	-0.013	Yes	No	5.31449	0.017	-0.004	No	No	5.31745	0.002	-0.001	No	No
100000	5.29049	0.021	-0.028	No	No	5.29382	0.020	-0.024	No	No	5.30794	0.020	-0.010	Yes	No	5.32013	0.020	0.002	No	No	5.31477	0.002	-0.003	No	No

Table 5

Price estimates from the LSM algorithm using various numbers of General Chebyshev Polynomials

	k=1			Stat. Sig.	Stat. Sig.	k=2			Stat. Sig.	Stat. Sig.
M	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%
10000	4.59830	0.012	-0.720	Yes	Yes	4.69133	0.017	-0.627	Yes	Yes
20000	4.63611	0.011	-0.682	Yes	Yes	4.70755	0.017	-0.611	Yes	Yes
30000	4.65415	0.011	-0.664	Yes	Yes	4.72257	0.017	-0.596	Yes	Yes
40000	4.66199	0.011	-0.656	Yes	Yes	4.71410	0.017	-0.604	Yes	Yes
50000	4.67363	0.012	-0.645	Yes	Yes	4.75285	0.017	-0.565	Yes	Yes
60000	4.67719	0.012	-0.641	Yes	Yes	4.76009	0.017	-0.558	Yes	Yes
70000	4.68417	0.012	-0.634	Yes	Yes	4.75911	0.017	-0.559	Yes	Yes
80000	4.69216	0.012	-0.626	Yes	Yes	4.74130	0.017	-0.577	Yes	Yes
90000	4.70259	0.012	-0.616	Yes	Yes	4.76757	0.017	-0.551	Yes	Yes
100000	4.70594	0.012	-0.612	Yes	Yes	4.77290	0.017	-0.545	Yes	Yes

	k=3			Stat. Sig.	Stat. Sig.	k=4			Stat. Sig.	Stat. Sig.
M	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%
10000	5.17465	0.019	-0.144	Yes	Yes	5.32876	0.020	0.011	No	No
20000	5.17799	0.019	-0.140	Yes	Yes	5.34822	0.019	0.030	No	No
30000	5.18487	0.019	-0.133	Yes	Yes	5.31117	0.019	-0.007	No	No
40000	5.16764	0.019	-0.151	Yes	Yes	5.29644	0.019	-0.022	No	No
50000	5.16670	0.019	-0.152	Yes	Yes	5.29915	0.019	-0.019	No	No
60000	5.16101	0.019	-0.157	Yes	Yes	5.29107	0.019	-0.027	No	No
70000	5.16691	0.019	-0.151	Yes	Yes	5.29515	0.019	-0.023	No	No
80000	5.17017	0.019	-0.148	Yes	Yes	5.29539	0.019	-0.023	No	No
90000	5.16965	0.019	-0.149	Yes	Yes	5.29437	0.019	-0.024	No	No
100000	5.17205	0.019	-0.146	Yes	Yes	5.29573	0.019	-0.022	No	No

	k=5			Stat. Sig.	Stat. Sig.
M	price	s.e.	bias	95%	99%
10000	5.32505	0.021	0.007	No	No
20000	5.34036	0.020	0.022	No	No
30000	5.31231	0.021	-0.006	No	No
40000	5.30387	0.021	-0.014	No	No
50000	5.29960	0.021	-0.019	No	No
60000	5.30536	0.020	-0.013	No	No
70000	5.30226	0.020	-0.016	No	No
80000	5.30404	0.020	-0.014	No	No
90000	5.30552	0.021	-0.013	No	No
100000	5.30747	0.020	-0.011	No	No

Table 6

Price estimates from the LSM algorithm using various numbers of Shifted Legendre Polynomials

	k=1			Stat. Sig.	Stat. Sig.	k=2			Stat. Sig.	Stat. Sig.	k=3			Stat. Sig.	Stat. Sig.	k=4			Stat. Sig.	Stat. Sig.	k=5			Stat. Sig.	Stat. Sig.
M	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%
10000	4.81191	0.011	-0.506	Yes	Yes	5.24605	0.022	-0.072	Yes	Yes	5.28432	0.020	-0.034	No	No	5.33884	0.020	0.021	No	No	5.33908	0.021	0.021	No	No
20000	4.80573	0.011	-0.512	Yes	Yes	5.25103	0.022	-0.067	Yes	Yes	5.27234	0.020	-0.046	Yes	No	5.32943	0.020	0.011	No	No	5.32652	0.020	0.008	No	No
30000	4.81346	0.011	-0.505	Yes	Yes	5.25073	0.022	-0.067	Yes	Yes	5.27704	0.020	-0.041	Yes	No	5.32304	0.020	0.005	No	No	5.32680	0.021	0.009	No	No
40000	4.80890	0.011	-0.509	Yes	Yes	5.25236	0.022	-0.066	Yes	Yes	5.27678	0.020	-0.041	Yes	No	5.31069	0.020	-0.008	No	No	5.31435	0.020	-0.004	No	No
50000	4.82624	0.011	-0.492	Yes	Yes	5.25658	0.022	-0.062	Yes	Yes	5.30061	0.020	-0.018	No	No	5.31619	0.020	-0.002	No	No	5.31722	0.020	-0.001	No	No
60000	4.82842	0.011	-0.490	Yes	Yes	5.26017	0.022	-0.058	Yes	Yes	5.30268	0.020	-0.016	No	No	5.31409	0.020	-0.004	No	No	5.31317	0.020	-0.005	No	No
70000	4.82686	0.011	-0.491	Yes	Yes	5.25529	0.022	-0.063	Yes	Yes	5.30006	0.020	-0.018	No	No	5.31005	0.020	-0.008	No	No	5.31198	0.020	-0.006	No	No
80000	4.82561	0.011	-0.493	Yes	Yes	5.25320	0.022	-0.065	Yes	Yes	5.29797	0.020	-0.020	No	No	5.30734	0.020	-0.011	No	No	5.30882	0.020	-0.009	No	No
90000	4.82212	0.011	-0.496	Yes	Yes	5.25441	0.022	-0.064	Yes	Yes	5.29832	0.020	-0.020	No	No	5.30625	0.020	-0.012	No	No	5.30740	0.020	-0.011	No	No
100000	4.82402	0.011	-0.494	Yes	Yes	5.25494	0.022	-0.063	Yes	Yes	5.30059	0.020	-0.018	No	No	5.30848	0.020	-0.010	No	No	5.30905	0.020	-0.009	No	No

Table 7

Price estimates from the LSM algorithm using various numbers of Monomials

	k=1			Stat. Sig.	Stat. Sig.	k=2			Stat. Sig.	Stat. Sig.	k=3			Stat. Sig.	Stat. Sig.	k=4			Stat. Sig.	Stat. Sig.	k=5			Stat. Sig.	Stat. Sig.
M	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%	price	s.e.	bias	95%	99%
10000	5.25026	0.022	-0.068	Yes	Yes	5.31080	0.020	-0.007	No	No	5.31919	0.020	0.001	No	No	5.32408	0.021	0.006	No	No	5.33002	0.021	0.012	No	No
20000	5.24970	0.022	-0.069	Yes	Yes	5.29468	0.020	-0.024	No	No	5.31366	0.021	-0.005	No	No	5.31328	0.021	-0.005	No	No	5.31644	0.021	-0.002	No	No
30000	5.24717	0.022	-0.071	Yes	Yes	5.29540	0.020	-0.023	No	No	5.31425	0.021	-0.004	No	No	5.31644	0.021	-0.002	No	No	5.31630	0.021	-0.002	No	No
40000	5.24722	0.022	-0.071	Yes	Yes	5.29646	0.020	-0.022	No	No	5.30801	0.020	-0.010	No	No	5.30937	0.021	-0.009	No	No	5.31206	0.021	-0.006	No	No
50000	5.25101	0.022	-0.067	Yes	Yes	5.29656	0.020	-0.022	No	No	5.31280	0.021	-0.005	No	No	5.31382	0.021	-0.004	No	No	5.31540	0.021	-0.003	No	No
60000	5.25244	0.022	-0.066	Yes	Yes	5.29754	0.020	-0.021	No	No	5.31295	0.020	-0.005	No	No	5.31301	0.021	-0.005	No	No	5.31484	0.021	-0.003	No	No
70000	5.25529	0.022	-0.063	Yes	Yes	5.29553	0.020	-0.023	No	No	5.30919	0.020	-0.009	No	No	5.31168	0.020	-0.007	No	No	5.31223	0.020	-0.006	No	No
80000	5.25320	0.022	-0.065	Yes	Yes	5.29653	0.020	-0.022	No	No	5.30692	0.020	-0.011	No	No	5.30791	0.020	-0.010	No	No	5.30780	0.020	-0.010	No	No
90000	5.25024	0.022	-0.068	Yes	Yes	5.29624	0.020	-0.022	No	No	5.30625	0.020	-0.012	No	No	5.30663	0.021	-0.012	No	No	5.30711	0.021	-0.011	No	No
100000	5.25213	0.022	-0.066	Yes	Yes	5.29726	0.020	-0.021	No	No	5.31476	0.021	-0.003	No	No	5.30959	0.021	-0.009	No	No	5.30465	0.020	-0.014	No	No

6.5 Comparison of the results

In table 4, we note that increasing the number of regressors, does not necessarily mean that we get a more accurate estimation. We see that for a low number of regressors, that is k-1 or k=2, the simulation seems to underprice the option. While we get more accurate results with three regressors, when we increase their number to four and five, we don't get more accurate results and we don't have over pricings either. This corresponds to the approximation bias mentioned above. In our results for the Laguerre polynomials, the proposition of Longstaff and Schwartz we mentioned above, does not seem to hold, meaning that by increasing the number of regressors from three to four and five, actually provided us with the same degree of accuracy as having only three regressors used. This means that a number of three regressors is enough to price the option in interest. Stentoft in the other hand, found that a use of 4 or five regressors, seemed to overprice the option, making the proposition of Longstaff and Schwartz meaningless. This difference in results may be due to the fact that we value an at-the-money option, while Stentoft values a deep in the money option, with s0=$36, leaving the rest of the other characteristics, the same. Another point to make is the existence of pretty few estimations, with statistically significant biases, and those are found only on the region of k=1 regressor, with a relatively poor number of paths, and within the critical region of 5%.

In table 5, we agree with the conclusions made by Stentoft, on his simulations with General Chebyshev polynomials. We have to increase the number of paths and regressors used in order to have accurate results. Same goes for table 6, where our conclusions from using Shifted Legendre polynomials are again similar with those of Stentoft. An increasing number of both paths and regressors must be used in order to avoid any under pricings and over pricings. In table 7, we note that monomials act surprisingly well, comparing with other, more complicated polynomials. When a number of three regressors is reached, the results obtained do not lack of accuracy, from the ones obtained by other polynomials families or by monomials with greater k, namely k=4 or k=5.

APPENDIX 6.6

6.6.1. Characteristics of polynomials

Table 8, gives the weights, the interval of orthogonallity, and the definition of the polynomials families, we used in our regressions, as taken by Stentoft (2004a).

Table 8

Characteristics and formulas for the simulated polynomials.

Polynomial family	**Weight**	**Interval**	**Definition**
Laguerre	e^{-x}	$[0,\infty)$	$L_k(x) = \frac{e^x}{k!}\frac{d^k}{dx^k}\left(x^k e^{-x}\right)$
General Chebyshev	$\left(1-(2x-1)^2\right)^{-0.5}$	$[0,1]$	$T_k(x) = \cos\left(k\cos^{-1}(2x-1)\right)$
Shifted Legendre	1	$(0,1)$	$P_k(x) = \frac{(-1)^k}{2^k k!}\frac{d^k}{dx^k}\left(\left(1-(2x-1)^2\right)^k\right)$

Three observations must be made.

- Instead of e^{-x}, we used $e^{-\frac{x}{2}}$ as weight function for the Laguerre polynomials, like Longstaff and Schwartz did in their simulations.
- Due to their intervals of orthogonallity, General Chebyshev and Shifted Legendre polynomials, cannot be used for the valuation of call options. That is more clear in the weight function of the General Chebyshev polynomials. The in-the-money, normalized values of the stock's price are greater than 1, so substituting them for x, we get a negative value raised on a decimal power.
- The definition given by Stentoft for the Shifted Legendre polynomial has proved to be incorrect, with the use of *Mathematica v5.2*. We propose an other representation of the formula, written

$$P_k(x) = (k!)^{-1}\frac{d^k}{dx^k}\left[\left(x^2-x\right)^k\right]$$ [1]

[1] This alternative formula was found in http://en.wikipedia.org/wiki/Legendre_polynomials#Shifted_Legendre_polynomials

6.6.2 LSM Simulation guide

Here is our proposed simulation of the LSM algorithm, in a step-by-step form.

1) Using a random seed, or multiple seeds, simulate the stock's price paths, making an MxT/Dt – dimensional matrix, and normalize the stock price.
2) Next steps must conclude the in-the-money paths, and only them. Discount the payoffs of the option, given in the last timestep, one timestep back, and regress them on a constant, and a set of values of the underlying factor of the penultimate timestep, (i.e. the stock's price), that you choose (that is, you are choosing the polynomial family and the number of regressors).
3) Compare the fitted value of the regression (which is the continuation value), with the payoff of the penultimate timestep, and make a matrix with the results of the comparison. Discount this matrix one period back, and do the same regression, using the set of values of your choice, for the values that are one timestep before the penultimate timestep. Do the same comparison, and make all the subsequent regressions, moving one timestep back, at a time. It is important to say that we don't use the discounted matrix of comparison as regressand, as this leads to upward biases in the option's price. (Longstaff and Schwartz , 2001). We use the maximum value realized for each path.
4) Construct the optimal stopping rule, and discount the proper unormalised values.
5) Finally we average the discounted values and get the value of the option.

We notice that simply by constructing the matrix of the stock's price paths, and average the discounted (to present) values of the last timestep, we actually value a European option, the way Boyle (1976) instructed to.

Chapter 7
Monte Carlo Simulation of stochastic volatility models

7.1 Heston's stochastic volatility model

The most famous stochastic volatility model is that of Heston (1993). The reason is that Heston assumes a degree of correlation between the returns of the stock underlying the option and the volatility itself, while other models do not. (see literature review). In this study we will not examine the closed-from solution given in Heston (1993), because our purpose is to use the basics of this model and valuate options via simulation. So, to begin with, Heston assumes that the asset price follows the usual stochastic process, written

$$dS_{(t)} = \mu S dt + \sqrt{v(t)} S dz_1(t) \qquad (1)$$

where v stands for the variance of the changes of the asset's price. Then he assumes an Ornstein-Uhlenbeck process for the volatility, which is

$$d\sqrt{v(t)} = -\beta\sqrt{v(t)}dt + \delta dz_2(t) \qquad (2)$$

and by using Ito's lemma, ends up in the following process

$$dv(t) = \left[\delta^2 - 2\beta v(t)\right]dt + \sigma\sqrt{v(t)dz_2(t)} \qquad (3)$$

which can be written as the square-root process of Cox, Ingersoll, Ross (1985)

$$dv(t) = \kappa\left[\theta - v(t)\right]dt + \sigma\sqrt{v(t)}dz_2(t) \qquad (4)$$

Equations (1) and (2) constitute the fundamental functions that characterize Heston's model of stochastic volatility. κ,θ,σ stand for mean reversion speed, long-run volatility and volatility of the volatility and are strictly positive constants. We also assume that $dz_1(t) \cdot dz_2(t) = \rho dt$, where ρ is the correlation constant in $[-1,1]$. There are two thinks we have to do to simulate this model:

- Sample from correlated Brownian motions, and
- Express the process followed by volatility, in its discrete form.

7.1.1 Sampling from correlated processes

We build two matrices of random normal variables, i.e. random1 and random2. A third matrix random3, will be a combination of the random1 and random2, written

$$Random_3 = \rho \cdot Random_1 + \sqrt{1-\rho^2} \cdot Random_2$$

where $Random_3$ and $Random_1$ are correlated with a degree ρ of correlation, exogenously defined. As in the other simulation schemes of this study, all the three matrices use the same number of positive and antithetic variables.

7.1.2 Discretization of the variance process

In order to have a discrete solution of (4), we must first define what is known as the price of volatility risk, namely λ. λ will depend on the stock's price, its variance and time t. We give Breeden's (1979) consumption-based model

$$\lambda(S,v,t)dt = \gamma Cov[dv, dC/C] \qquad (5)$$

where γ stands for the degree of risk aversion of the consumer. (5) can be written as a Cox, Ingersoll, and Ross mean-reverting model

$$dC(t) = \mu_c v(t)Cdt + \sigma_c \sqrt{v(t)}Cdz_3(t) \qquad (6)$$

where consumption growth has constant correlation with the asset return. This generates a risk premium proportional to variance, that is $\lambda(S,v,t)dt = \lambda v$. We finally derive the discrete form of (4), assuming a risk-neutral pricing framework, written

$$V(t+1) = [k^*\theta^* - k^*V(t)]\Delta t + \zeta \sqrt{v(t)}dz_4 \qquad (7)$$

where

$$k^* = k + \lambda \qquad \text{and} \qquad \theta^* = \kappa\theta/(\kappa + \lambda)$$

and ζ is the volatility of stochastic volatility. To conclude, the two SDEs of the Heston's stochastic volatility model, are

$$dS_{(t)} = \mu S dt + \sqrt{v(t)} S dz_1(t) \qquad (8)$$

$$V(t+1) = \left[k^*\theta^* - k^* V(t)\right]\Delta t + \zeta \sqrt{v(t)} dz_2 \qquad (9)$$

Remember matrices $random_1$, $random_2$ and $random_3$ from 6.1.1. We sample from $random_1$ for the simulation of the variance rate, and from $random_3$ to simulate the stock's price. We will use the LSM algorithm, so stock's price and stochastic volatility will be included in the regressors of the algorithm. As Longstaff and Schwartz state for the valuation of multifactor options, one has to include the cross-products, and the product of all the stochastic factors in the regression. We will include the product of volatility and stock's price.

7.2 Simulation issues and solutions

While (4) is guaranteed to be nonnegative, this is not the case for its discrete form. That makes an accurate valuation impossible. A lot of alternative discretizations have been proposed for (4). We will use the one introduced by Robel (2001). He starts with the introduction of an exponential Mean-Reverting Ornstein-Uhlenbeck process

$$X_t = e^{Y_t} \qquad (10),$$

where

$$dY_t = \lambda\left(\bar{Y} - Y_t\right)dt + \sigma dB_t \qquad (11)$$

where λ stands for mean-reversion speed and dB_t is a Brownian process. He then uses Ito's lemma to see how X_t is distributed, and if it reverts to a mean. We get the formula

$$dX_t = \lambda\left[\left(\bar{Y} + \frac{1}{2\lambda}\sigma^2\right) - \log(X_t)\right]X_t dt + \sigma X_t dB_t \qquad (12)$$

We see that X_t is lognormally distributed, with probability density function

$$f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\ln x-\mu)^2/2\sigma^2} \qquad (13).$$

The first two moments of the lognormal distribution are

$$E(X) = e^{\mu+\sigma^2/2} \qquad (14)$$

and

$$Var(X) = \left(e^{\sigma^2} - 1\right)e^{2\mu+\sigma^2} \qquad (15)$$

where μ and σ are the mean and the standard deviation of the variable's logarithm. We see from its probability density function in Figure 2, that the lognormal distribution cannot take negative values.

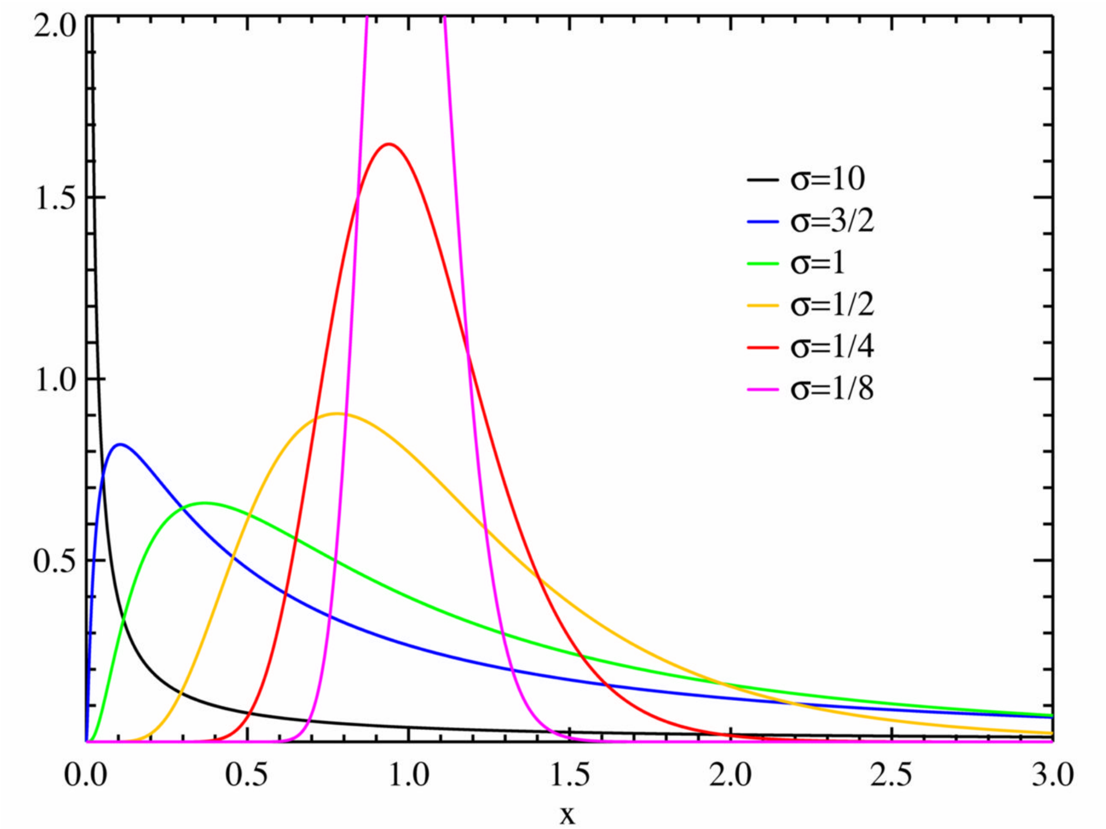

Figure 2. The probability density function of the lognormal distribution.

So, (12) is a suitable process to use for the simulation of the variance. The only think we have to do now is to find its discrete solution. An exact Euler discretization is

$$X_{t+\Delta t} = X_t e^{\lambda\left(\bar{Y} - \log(X_t)\right)\Delta t + \sigma\varepsilon_t\sqrt{\Delta t}} \qquad (16)$$

That's the process we adopt for the simulation of variance in this study.

Due to lack of computational means, we only used monomials on our simulation and we reached a number of three regressors. We have assumed a kappa=1, to make the process more volatile, θ=0.16 (it is the square of the volatility used in Black-Scholes framework), an initial variance v0=0.01, positive correlation ρ=0.5, and a volatility σ = ζ = 0.1 for the volatility parameter. Table 9 cites the valuation of a European option by using Boyle's (1976) simulation algorithm on our processes for the stochastic factors, and table 10 cites the results of a simulation done with LSM algorithm, on an American option. The options have the same parameters with the options evaluated using the methods cited earlier in this study.

Table 9

Simulation of a European option with stochastic volatility

M	Price	s.e.
10000	2.57897	0.009
20000	2.61583	0.010
30000	2.58286	0.010
40000	2.58129	0.010
50000	2.58453	0.010
60000	2.60632	0.010
70000	2.59685	0.010
80000	2.59943	0.010
90000	2.59687	0.010
100000	2.60348	0.010

Table 10

Simulation of an American option with stochastic volatility

	k=1		k=2		k=3	
M	price	s.e.	price	s.e.	price	s.e.
10000	2.56649	0.009	2.56945	0.009	2.57110	0.009
20000	2.61796	0.010	2.61872	0.010	2.61824	0.010
30000	2.60618	0.010	2.60626	0.010	2.60605	0.010
40000	2.61037	0.010	2.61133	0.010	2.61078	0.010
50000	2.61855	0.010	2.61875	0.010	2.61879	0.010
60000	2.60896	0.010	2.60902	0.010	2.60842	0.010
70000	2.61732	0.010	2.61736	0.010	2.61673	0.010
80000	2.61077	0.010	2.61094	0.010	2.61043	0.010
90000	2.61418	0.010	2.61419	0.010	2.61385	0.010
100000	2.59642	0.010	2.59642	0.010	2.59615	0.010

We note that the prices obtained by the process we adopted for the variance are significantly lower than the prices of options with constant volatility, valued with each of the methods proposed on this study. The two groups of prices are not comparable, because we assumed a process for the variance, different from the one proposed in Heston (who values European options only, and has results very close to the ones obtained by B-S pricing), in order to avoid negative prices.

8. Conclusion

This paper has covered a wide range of methods for option pricing. Beginning from the first analytical closed-form solution, the easiest to implement Black-Scholes method, we cover a part of lattices, both finite difference approaches and Monte Carlo simulations. What comes from this study is that we can avoid using simulation schemes when we have simple options in mind. Simulations come to cover the computational gap for multiple state options, like average options and options with stochastic volatility. The golden section haw to be found when using simulations, because we don't get more accurate results simply by increasing the domain space of the simulation. The use of more and more paths and most important of more and more complicated and numbered regressors should be sought. Future students can use this study as a base for implementing more complicated algorithms, i.e. add alternative discretizations for the process followed by the variance, or adapt stochastic interest rates. Time and will, shall be their outfits.

Bibliography

1) John Hull, "Options, Futures and Other Derivatives", Sixth Edition.

2) Γεώργιος Κ. Χρήστου, "Εισαγωγή στην Οικονομετρία".

3) Damodar N. Gujarati, "Basic Econometrics", Fourth Edition

4) Ηλίας Φλυτζάνης, "Ανώτερα Μαθηματικά".

5) Mark Broadie, Jerome Detemple. "Option Pricing: Valuation Models and Applications". Management Science, Vol. 50, No 9, September 2004, pp. 1145-1177

6) Phelim P. Boyle. "Options: A Monte Carlo Approach". Journal of Financial Economics 4 (1977) pp. 323-338.

7) Steven L. Heston. "A Closed-Form Solution for Option with Stochastic Volatility with Applications to Bond and Currency Options". The Review of Financial Studies, Vol. 6, No. 2. (1993), pp. 327-343.

8) Jerome Barraquand, Didier Martineau. "Numerical valuation of High Dimensional Multivariate American Securities". Journal of Financial and Quantitative Analysis, 30, pp. 383-405, 1994.

9) Francis A. Longstaff, Eduardo S. Schwartz. "Valuing American Options by Simulation: A Simple Least-Squares Approach". The Review of Financial Studies, Spring 2001, Vol. 14, No 1, pp. 113-147.

10) Gregory F. Robel. "Real Options and Mean-Reverting Prices". July 2001.

11) Jouni Kerman. "Numerical Methods for Option Pricing: Binomial and Finite-difference Approximations". Courant institute of Mathematical Sciences, January 2002.

12) Lars Stentoft. "Assessing the Least Squares Monte-Carlo Approach to American Option valuation". Review of Derivatives Research, Vol. 7, pp. 129-168, 2004.

13) Arthur Rodrigues, Manuel J. Rocha Armada. "The Valuation of Real Options with the Least Squares Monte Carlo Simulation Method". Social Science Research Network (SSRN). February 2006.