THE UNIVERSITY OF CALGARY

Advanced Monte Carlo Simulations and Option Pricing

by

Hilda Evangeline Wong

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

November , 2000

Abstract

Consider the problem of pricing options whose payoffs depend on multiple sources of risk (rainbow options). Generally, under well known risk neutrality assumptions, the prices of options are calculated to be the expected value of future cash flows, discounted with the appropriate risk-free interest rate. However, for many rainbow options, the derivation of close-form solutions do not exist. Therefore, there is a need to rely on numerical methods such as lattice and finite difference methods or Monte Carlo simulation.

This thesis deals with the use of Monte Carlo simulation of stochastic processes as applied to option pricing. We numerically develop higher order discretization methods for stochastic differential equations and compare their accuracy for high dimensional option pricing problems. Furthermore, a new quasi-random variance reduction technique, extending classical antithetic variates, is introduced to increase simulation efficiency. This is applied to rainbow options, up to 100 assets, and underlyings with stochastic volatility.

Acknowledgements

I would like to thank my supervisor, Dr. Ali Lari-Lavassani, for his ideas, support, generosity, patience and guidance and also Dr. Ali Sadeghi, for his constructive critiques and insights. Thanks also go to Dr. Anthony Ware, for demonstrating the vectorization powers of Matlab and providing computing tips to me. I would also like to thank the Department of Mathematics and Statistics, and the Mathematical and Computational Finance Laboratory for providing me with the funding, specialized library and computer facilities to carry out my research. Finally, I would like to express my gratitude to TransAlta Energy Marketing Corp and the Canadian Networks of Centres of Excellence, The Mathematics of Information Technology and Complex Systems (MITACS), for partial support.

Dedication

I dedicate this to my family, for all the support, encouragement and sense of humour to keep me going.

Table of Contents

List of Tables

List of Figures

CHAPTER 1
INTRODUCTION

1.1 Generalities

In their seminal work Black and Scholes applied stochastic processes, more precisely geometric Brownian motions, to model the random evolution of stock prices within an arbitrage free setting. Ever since, financial modelling via stochastic differential equations has become the norm in the practice of trading financial derivatives. Often, these complex derivatives are difficult to price and often case they do not have analytic solutions, and one must resort to numerical schemes. Even when analytical solutions exist, in dimensions greater than or equal to one, numerical methods must be employed to obtain a solution.

There are many challenging mathematical problems in modeling and implementing complex derivatives. In this work, we focus on the numerical implementation of the Monte Carlo method in option pricing.

As high performance computing becomes more affordable and the computing technology more attainable, Monte Carlo simulation is increasingly feasible and effective in pricing complex derivatives. For a general review on the application of Monte Carlo simulation to option pricing, we refer to (Boyle, Broadie and Glasserman, 1997) and (Monte Carlo: Methodologies and Applications for Pricing and Risk Management, 1998) and the references therein.

The two main issues of importance in simulation are those of accuracy and efficiency, and our principal goal is to numerically investigate these in this thesis.

The accuracy of a simulation method depends on one hand on the discretization techniques used as paths are numerically generated in discrete time. We focus on two time discrete approximations of stochastic processes, namely, Euler and Milstein which have different orders of convergence. In general, discretization schemes for stochastic processes vary from their deterministic counterparts as there is a need to simulate the additional stochastic component.

The Euler discretization scheme has been extensively discussed in traditional finance literature for option pricing. It is simple to use but practical applications of this method alone reveal numerical instabilities, due to its low order of convergence 0.5. The effect of these numerical instabilities decreases when more simulation paths are employed or if the number of discrete time-steps is selected to optimize the scheme. Therefore, the computational costs associated with these actions are often very high and might exceed one's computational capabilities. One way to overcome this problem is to resort to higher order stochastic Taylor expansions to discretize the processes involved.

Going beyond Euler's method, (Milstein, 1974) proposes a discretization of order 1, which in the multiasset or multifactor case, introduces the use of double Ito integrals. These double Ito integrals can be simplified for certain classes of stochastic differential equations used in finance, but for those that cannot, it is essential to develop efficient

numerical methods to handle these Ito integrals. This might be the principal reason why the Milstein approximation has not received the attention it deserves in financial literature. It is necessary to compare these schemes and the feasibility of their implementation for high dimensional option pricing problems. In the context of high dimensional option pricing, one should select a scheme solely based on its performance, that is, the highest level of accuracy and the highest level of efficiency.

Another main area of concern is the efficiency of the simulation scheme in terms of the variance of the final result. Reducing variance and hence increasing efficiency should take into account the additional computational cost it brings about. For such reasons, variance reduction techniques such as antithetic variates have been used with great success in option pricing. Note though, the classical form of antithetic variates is deemed to be not as effective as other forms of variance reduction methods such as control variates. However, these methods are case specific and tend not to be simple to implement numerically.

As Monte Carlo simulation for option pricing in its standard form is driven by random sampling, one possible variance reduction improvement to this is to utilize quasi-random sampling techniques. However, instead of changing the simulation algorithm, an effective non-random sampling algorithm should utilize previously generated random variables in an ordered manner to maximize the accuracy of the results. By doing so, one seeks to reflect the distribution of the random variables more accurately. Pure random sampling might lead to an over-representation of certain portions of the

distributions. By less random sampling under certain rules, one can force sampled points to be more evenly distributed. We propose one such method in this thesis.

In keeping with the objective of financial modelling, simulation methods in option pricing should yield numerical results which are useful for benchmarking purposes. Indeed, previous work in the area of large basket option pricing only reproduces some error plots of their algorithms without quoting actual values. We present full option values even for large baskets of up to 100 assets.

1.2 Literature Review

(Boyle, 1977) was the first to use Monte Carlo simulations in an option pricing context. Boyle demonstrated the effectiveness of Monte Carlo simulations by generating the process for the underlying asset under the risk-neutrality assumption to price an option. The major assumption in Boyle's work was that the returns (ratio of successive share prices) followed a lognormal distribution. Boyle showed that Monte Carlo simulations incorporating variance reduction techniques such as antithetic and control variates could be effectively and successfully used to price European call options paying dividends.

This work was updated in a more comprehensive manner twenty years later. (Boyle, Broadie and Glasserman, 1997) reported the tremendous flexibility and power of Monte Carlo simulations for less trivial option pricing cases. Further work was

conducted to extend variance reduction techniques and even provided some discussion on less random simulations such as Quasi-Monte Carlo. A brief summary of Monte Carlo simulations applied to path-dependant option pricing (such as American and Asian) and hedging was also covered (see Broadie and Glasserman, 1996).

During this time, there has been a great amount of published work describing the efficiency of Monte Carlo simulations for financial applications. (Clewlow and Strickland, 1999), Chapter 4, review the various capabilities of Monte Carlo simulations for multiasset models and models with stochastic volatilities. (Barraquand, 1995) uses Monte Carlo simulation in conjunction with a specific sampling method to price multidimensional European type options and stresses the relevance of Monte Carlo for efficient numerical methods in option pricing. (Bhansali, 1998), p. 174, further espouses Monte Carlo simulations as the "method of choice for multiasset cases with no early exercise features".

Since the major criticism of Monte Carlo simulation has been and is the apparent slow rate of convergence[1] (Kalos and Whitlock, 1986, p. 27), there is a recognition that the efficiency of Monte Carlo simulation is imperative for successful implementation in financial applications. For instance, (Duffie and Glynn, 1995) highlight the tenuous relationship between computing resources and efficiency. Other works which provide techniques on speeding up Monte Carlo simulations for acceptable levels of accuracy

[1] According to the Central Limit Theorem, the convergence rate is $O\left(1/\sqrt{p}\right)$, p being the number of paths.

include (Ninomiya and Tezuka, 1996), (Papageorgieou and Traub, 1996), (Paskov and Traub, 1995) and (Paskov, 1997). Recent efforts to improve the efficiency of Monte Carlo via less random algorithms are seen in literature such as (Joy et al., 1996) and (Morokoff and Caflisch, 1998). We finally note that Quasi-Monte Carlo is a fast growing area of research, we refer to (Lemieux and L'Ecuyer, 2000) and the references therein.

CHAPTER 2
SIMULATING STOCHASTIC DIFFERENTIAL EQUATIONS

2.1 Preliminaries

Before the problem is introduced, we review a few preliminary definitions and refer to (Kloeden and Platen, 1999) for full details. The setting within which we will work is that of a *filtered probability space*. This can be formally defined by $(\Omega, \mathfrak{A}, P)$, where Ω is an arbitrary nonempty set called the *sample space*, $\mathfrak{A}$ is a σ-algebra of subsets of Ω called *events* and P is the *probability measure* on $\mathfrak{A}$ which assigns to each event a *probability* between 0 and 1 and satisfies $p(A \cup B) = p(A) + p(B)$ if A and B are disjoint. We recall that a σ-algebra $\mathfrak{A}$ is a collection of subsets of Ω where

$$\begin{aligned} \Omega &\in \mathfrak{A} \\ A^c &\in \mathfrak{A} \quad \text{if } A \in \mathfrak{A} \\ \bigcup_{n=1}^{\infty} A_n &\in \mathfrak{A} \quad \text{if } A_1, A_2, ..., A_n, ... \in \mathfrak{A}. \end{aligned}$$

We also recall that $(\Omega, \mathfrak{A})$ is called a *measurable space*. We can now define a *random variable* or *measurable function*.

Definition 2.1 *A vector valued function $X = (X^k) : \Omega \to \Re^d$ is measurable or called*

a vector random variable if the set

$$\{\omega \mid X^k(\omega) \leq x_k,\ k = 1, ..., d\} \in \Omega,$$

for all $x_k \in \Re$.

Since this thesis is concerned with option pricing with assumed stochastic processes for the underlying assets, we now define a stochastic process.

Definition 2.2 *A stochastic process with index set* I *and state space* $\Re^d$ *is a family*

$$X = \{X(t), t \in I\}$$

of $\Re^d$*-valued random variables for each* t.

In this thesis, I will be $[0, T]$, with $T \in \Re$ and T will be henceforth termed as the time to expiry or time horizon.

Definition 2.3 *A sample path is a realization of a stochastic process for each* $\omega \in \Omega$, $X(., \omega) : [0, T] \rightarrow \Re^d$.

The probability distribution that we will focus on is the Gaussian distribution. A random variable X is called *Gaussian* or *normal* $N(\mu, \sigma^2)$, if its density is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\}.$$

The d-dimensional random variable $X = (X^1, ..., X^d)$ is said to have a multivariate Gaussian or normal $N(\mu, \sigma)$ distribution, with $\mu : \Re^d \rightarrow \Re^d$ and $\sigma = (\sigma^{kj}) : \Re^d \rightarrow \Re^{d \times d}$,

assuming symmetric matrix values with positive eigenvalues, with its d-dimensional density given by

$$f(x) = \frac{|\sigma|^{-1/2}}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(x-\mu)^T\sigma^{-1}(x-\mu)\right]$$

where $|\sigma|$ is the determinant of σ, σ^{-1} is the inverse of σ, and

$$(x-\mu)^T\sigma^{-1}(x-\mu) = \sum_{k=1}^{d}\sum_{j=1}^{d}\frac{1}{\sigma^{kj}}(x_k-\mu_k)(x_j-\mu_j).$$

Note that once again, the Gaussian distribution is uniquely determined by its first two moments, that is, μ and σ.

We can now define the Wiener process W by the following properties:

Definition 2.4 *A d-dimensional process $W(t) = (W_t^1, ..., W_t^d)$ measurable with respect to an increasing family or filtration of σ-algebras $\{\mathfrak{A}_t, t \geq 0\}$ is a d-dimensional Wiener process if and only if*

$$\begin{aligned} E(W_t^k|\mathfrak{A}_0) &= 0, \\ E(W_{t_2}^k - W_{t_1}^k|\mathfrak{A}_{t_1}) &= 0, \quad w.p.1 \text{ and} \\ E((W_{t_2}^{k_1} - W_{t_1}^{k_1})(W_{t_2}^{k_2} - W_{t_1}^{k_2})|\mathfrak{A}_{t_1}) &= (t_2 - t_1)\ \delta_{k_1k_2} \end{aligned}$$

for all $0 \leq t_1 \leq t_2$, $k = 1, ..., d$, and where

$$\delta_{k_1k_2} = \begin{cases} 1 & k_1 = k_2 \\ 0 & k_1 \neq k_2 \end{cases}, \quad \text{for } k_1, k_2 = 1, ..., d.$$

Note the important property

$$W_{t_2}^k - W_{t_1}^k = \sqrt{t_2 - t_1}\, Z^k \text{ for all } 0 \leq t_1 \leq t_2$$

where $Z^k \sim N(0,1)$.

Recall also that for $0 < T < \infty$, the class $\mathfrak{L}_T^2$ of functions $f : [0,T] \to \Omega$ is defined (Kloeden and Platen, 1999, p. 81) to satisfy these properties

$$f \text{ is jointly } \mathfrak{L} \times \mathfrak{A}\text{-measurable},$$

$$\int_0^T E(f(t,.)^2)dt < \infty,$$

$$E(f(t,.)^2) < \infty \quad \text{for each } 0 \leq t \leq T.$$

We next define Stochastic Differential Equations, denoted by (SDE) henceforth.

Definition 2.5 *Consider a d-dimensional vector function* $\mu = (\mu^k(S_t,t)) : [0,T] \times \Omega \to \Re^d$, *with* μ^k *satisfying* $\sqrt{\mu^k} \in \mathfrak{L}_T^2$, $k = 1,...,d$ *and a* $d \times d$*-matrix function* $\sigma = (\sigma^{kj}(S_t,t)) : [0,T] \times \Omega \to \Re^{d \times d}$ *with* $\sigma^{kj} \in \mathfrak{L}_T^2$, $k,j = 1,...,d$. *A* $d-$*dimensional stochastic process* $\{S_t, 0 \leq t \leq T\}$ *is said to satisfy a* <u>*stochastic differential equation*</u> *if*

$$dS_t = \mu(S_t,t)dt + \sigma(S_t,t)dW_t \tag{2.1}$$

where W *is a d-dimensional Wiener process.*

Equation (2.1) has a solution, in the sense explained by the following definition.

Definition 2.6 *With the previously defined notations, a strong solution to the SDE (2.1) is a stochastic process* $\{S_t, t \geq 0\}$ *that is* $\mathfrak{A}_t$*-measurable for any* t*, and such that the following conditions hold*

$$\int_0^t |\mu^k(S_v, v)| \; dv \; < +\infty,$$

$$\int_0^t |\sigma^{kj}(S_v, v)|^2 \; dW_v^j \; < +\infty, \; \textit{and}$$

$$S_t^k - S_0^k = \int_0^t \mu^k(S_v, v) \; dv + \sum_{j=1}^{d} \int_0^t \sigma^{kj}(S_v, v) \; dW_v^j \tag{2.2}$$

for all $t \geq 0$

Note that the integral form of equation (2.1) now becomes

$$S_t - S_0 = \underbrace{\int_0^t \mu(S_v, v) \; dv}_{\text{Lebesgue Integral}} + \underbrace{\int_0^t \sigma(S_v, v) dW_v}_{\text{Ito Integral}} \tag{2.3}$$

for any $0 \leq t \leq T$.

The sample paths of the Ito process inherit the characteristics of the sample paths of the driving Wiener process (see Figure 2.1) and one of the properties of a Wiener process is that its paths are not differentiable at any point, almost surely, therefore, the Ito integral seen in (2.3) is not a Lebesgue integral.

Definition 2.7 *A step function* $f \in \mathfrak{L}_T^2$ *is defined to be*

$$f(t, \omega) = f(t_j, \omega) \quad w.p.1, \textit{ for } t_j \leq t \leq t_{j+1}$$

for $j = 1, ..., n$. *Let* $\mathfrak{S}_T^2$ *be the subset of all step functions in* $\mathfrak{L}_T^2$.

Using this definition, we can define the Ito integral.

Definition 2.8 *The Ito integral* $I(f)$ *is defined as*

$$I(f) = \int_a^b f(t,\omega)dW_t = \sup \sum_{j=0}^{n-1} f(t_j,\omega)\left\{W_{t_{j+1}} - W_{t_j}\right\} \quad w.p.1. \tag{2.4}$$

where the supremum is taken over all possible discretizations of the interval $[a,b]$ *and* n *is the number of subintervals.*

In order to ensure the existence and uniqueness of a solution to equation (2.1), there have to be sufficient conditions on μ and σ as set out in Theorem 5.5 (Øksendal, 1995, p. 64) and Theorem 3.5.3 (Lamberton and Lapeyre, 1996, p. 50), which we recall now:

Theorem 2.9 *Consider the following measurable functions:*

$$\mu(.,.) : [0,T] \times \Re^d \to \Re^d$$

and

$$\sigma(.,.) : [0,T] \times \Re^d \to \Re^{d\times d}.$$

For any $0 \leq t \leq T$, *if*

$$|\mu(x,t) - \mu(y,t)| + |\sigma(x,t) - \sigma(y,t)| \leq K|x-y|, \quad x,y \in \Re^d \quad \textit{and}$$

$$|\mu(x,t)| + |\sigma(x,t)| \leq K(1+|x|)$$

for some constant K, $|\sigma|^2 = \sum_{k=1}^{d}\sum_{j=1}^{d}\left(\sigma^{kj}\right)^2$, *then (2.1) admits a unique strong solution for* $t \in [0,T]$. *This solution* $(S_t), 0 \leq t \leq T$, *also satisfies*

$$E\left(\sup_{0\leq t\leq T}|S_t|^2\right) < \infty.$$

Figure 2.1 provides some intuition into the behaviour of the Wiener process with relation to an underlying stochastic process. It compares a sample path from a lognormal model, governed by such a stochastic differential equation

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t, \tag{2.5}$$

with the associated Wiener process for that particular path. It is quite evident that the driving force in the random behaviour of the price path is the corresponding Wiener process. The peaks and troughs of the price path correspond exactly to that of the Wiener process.

2.2 Options

The no-arbitrage approach to pricing options was introduced by (Black and Scholes, 1973), and also discussed by (Cox and Ross, 1976). Its relationship with risk-neutral valuation was put in evidence by (Harrison and Pliska, 1981). We refer to (Bingham and Kiesel, 1998) for an excellent presentation on this material. Since the object of this thesis is the numerical valuation of options, we henceforth assume that all price

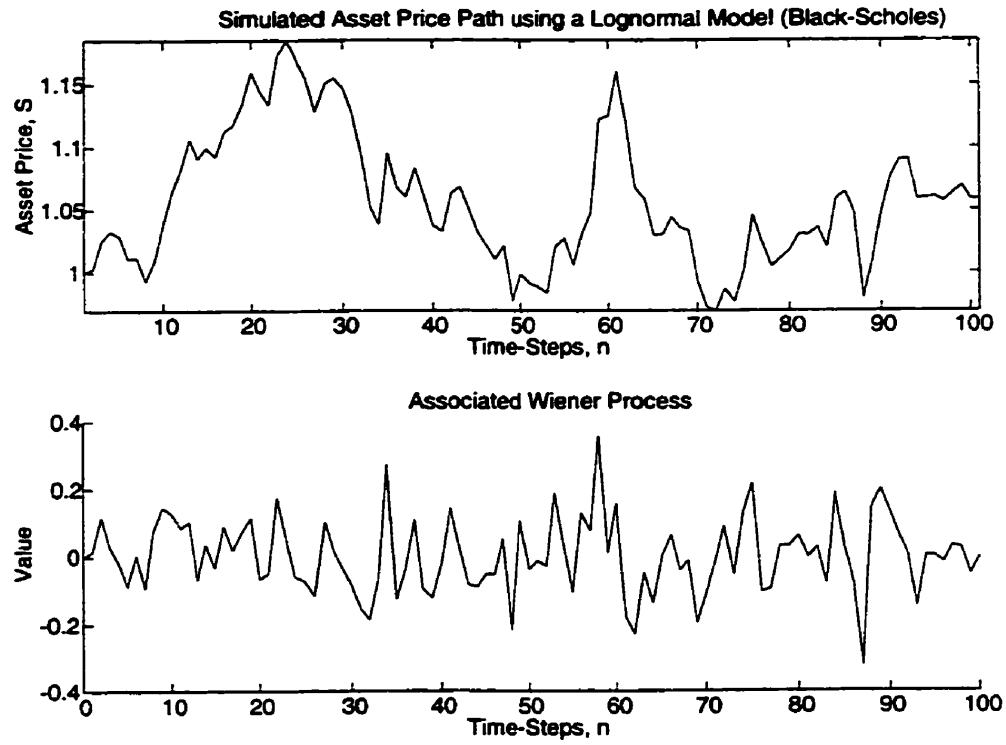

Figure 2.1: Sample Path and Corresponding Wiener Process

processes are risk neutral, and therefore the price of an option is equal to the expected payoff discounted at the risk-free interest rate.

We will seek to determine the expectation of a given functional f of S_t at a given time T, that is, to find $E(f(S_T)|S_0)$. In the context of option pricing, f defines the payoff of a given option on the d securities S_t^k, $k = 1, ..., d$, whose price paths are governed by a stochastic differential equation of the form (2.1). We can now formally define derivatives and options.

Definition 2.10 *A derivative asset is a security whose value explicitly depends on an*

underlying asset on which the derivative is written.

An option is a form of derivative, whose value depends on the underlying asset. This contingent claim could result in a potential gain for one party and also a concurrent potential liability for the opposing party. Options can be either call or put options.

Definition 2.11 *A call option gives its owner the right to buy an asset at a specified price (termed the strike or exercise price, K) on or before a stated date (termed maturity date or time to expiry, T).*

There is also the converse to the call, that is

Definition 2.12 *A put option gives the owner the right to sell the asset at a specified price on or before the stated date.*

In general, call and put options can be defined by the dependence of the option price on the path of the asset price.

Definition 2.13 *Path independent options have payoffs depending only on the events upon expiration, regardless of the route taken by the underlying process.*

Definition 2.14 *Path dependent options depend on at least one price path in addition to the terminal price.*

There are respectively American or European options which are defined as such:

Definition 2.15 *A European option does not permit the exercise of the said option before the expiry date, and has a payoff function of the form*

$$\max\left\{0, f\left(S_T^1, ..., S_T^d\right)\right\},$$

where in the case of a single security,

$$f\left(S_T\right) = \begin{cases} \max(S_T - K, 0), & \textit{for a call} \\ \max(K - S_T, 0), & \textit{for a put} \end{cases}.$$

Definition 2.16 *An American option can be exercised at any time between its start date to the terminal date and has a payoff function of the form*

$$\max\left\{0, f\left(S_\tau^1, ..., S_\tau^d\right)\right\},$$

where $\tau \leq T$, and in the case of a single security,

$$f\left(S_\tau\right) = \begin{cases} \max(S_\tau - K, 0), & \textit{for a call} \\ \max(K - S_\tau, 0), & \textit{for a put} \end{cases}$$

Figures 2.2 and 2.3 show the basic behaviour of standard European call and put options respectively. The truncated payoff is evident when the asset price is less than the exercise price for a call, therefore, the option owner is not obligated to exercise the option, limiting the losses to only the outlay for the option. The put option is the converse to this.

For more detailed expositions on basic and other exotic options, see (Wilmott, 1998), (Bhansali, 1998), (Jarrow and Turnbull, 1996), (Kolb, 1997), and (Taleb, 1997).

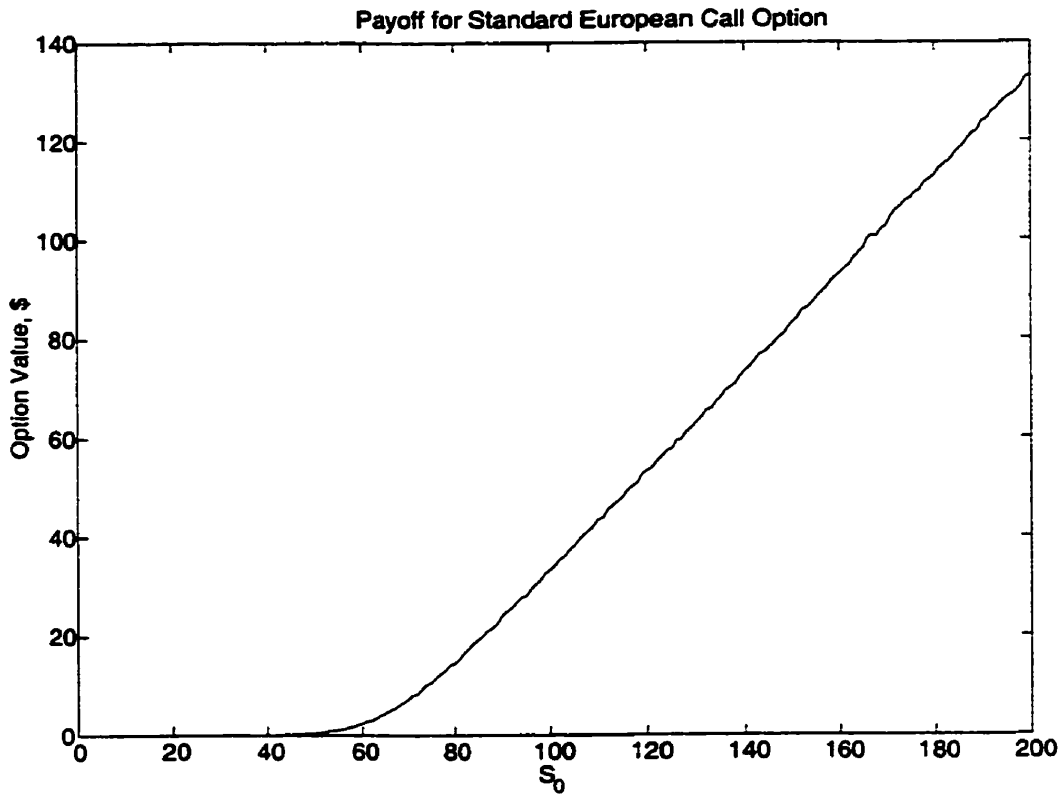

Figure 2.2: Payoff to the Holder of a Call Option at Maturity

We will concentrate on options on several risky assets, particularly European rainbow options such as basket/portfolio options, strikes and spreads. There will also be numerical focus on options with multiple underlying processes, such as those seen in the system (2.33).

To obtain the option price, one will discount the expectation $E(f(S_T)|S_0)$ at time T with respect to the risk-free interest rate, r, i.e.

$$e^{-rT}E(f(S_T)|S_0).$$

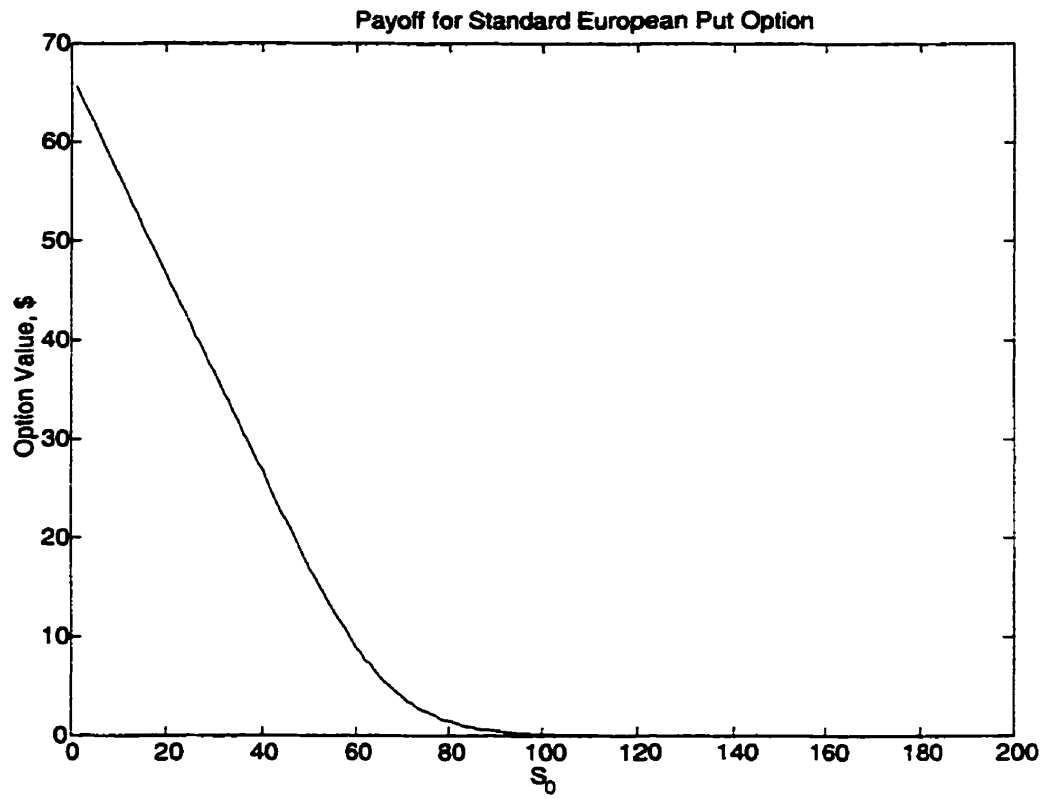

Figure 2.3: Payoff to the Holder of a Put Option at Maturity

Since the options that will be focused on do not have exact solutions, numerical methods have to be used to compute $E(f(S_T)|S_0)$. These methods include Monte Carlo, lattice or tree methods, and finite differences. Lattice, or tree, and finite difference methods are useful whenever there are path dependent features in an option, and they are numerically efficient for dimensions up to two. However, lattice and finite difference methods are not flexible, as dimensions increase, and as the SDE for the underlying or the payoffs for different options change, the algorithms have to be completely revised. Hull and White (1988) note that if the number of nodes at a point in time is

unnecessarily large, the lattice methods become inefficient as the computational time is proportional to the total number of nodes.

On the other hand, the Monte Carlo method is extremely robust and flexible, as the same algorithm can be used for any dimension theoretically. It also has the distinct advantage that for path independent options, the same algorithm can be used again. Additionally, changes in the payoffs require minimal additional effort. Monte Carlo also deals easily with multiple random factors such as stochastic volatility. Furthermore, it is able to incorporate more realistic price processes such as jumps in asset prices. Monte Carlo is also relatively cheap to perform as the computational cost of simulations increases linearly with the number of variables, so the method's competitiveness increases for options with multifactor models and large number of assets. The major drawback of standard Monte Carlo is the speed of convergence and the apparent inapplicability to path dependent options[1]. Finally, large numbers of simulations are required for convergence if standard Monte Carlo is used.

To summarize, these are the following steps in obtaining an option price numerically by Monte Carlo:

1. Generate the sample paths of the underlying assets over the given time horizon under the risk neutral measure;

2. Evaluate the discounted payoffs of the option for each path; and

[1](Broadie and Glasserman, 1997) use a revised form of Monte Carlo for American option pricing.

3. Average these payoffs over all paths.

2.3 Monte Carlo Simulation

2.3.1 Monte Carlo Methodology

Suppose to estimate the quantity θ, we have to use an estimator, $\widehat{m}_p$. By using Monte Carlo simulation one generates an arbitrary number of independent samples $\widehat{\theta}_i, i = 1, ..., p$, in a sample space Ω, where for each $\widehat{\theta}_i$,

$$E\left(\widehat{\theta}_i\right) = \theta \tag{2.6}$$

and

$$var\left(\widehat{\theta}_i\right) = \sigma^2. \tag{2.7}$$

Set

$$\widehat{m}_p = \frac{1}{p}\sum_{i=1}^{p}\widehat{\theta}_i, \tag{2.8}$$

the sample mean, to be the estimator of θ. Also, note that for any p,

$$E\left(\widehat{m}_p\right) = \frac{1}{p}\sum_{i=1}^{p}E\left(\widehat{\theta}_i\right) = \theta \tag{2.9}$$

and

$$\begin{aligned} var\left(\widehat{m}_p\right) &= var\left(\frac{1}{p}\sum_{i=1}^{p}\widehat{\theta}_i\right) \\ &= \frac{1}{p^2}\sum_{i=1}^{p} var\left(\widehat{\theta}_i\right) \\ &= \frac{\sigma^2}{p}. \end{aligned} \tag{2.10}$$

According to (Lehmann and Casella, 1998, Chapter 2, pp. 83-100) and (Bartoszyński and Niewiadomska-Bugaj, 1996, p. 474), a good estimator $\widehat{m}_p$ should become more precise with increasing numbers of samples, p and it should enforce some measure of impartiality. This leads us to defining the following characteristics of a good estimator.

Definition 2.17 *An estimator $\widehat{m}_p$ of θ is weakly consistent if for every $\varepsilon > 0$*

$$\lim_{p\to\infty} P\left(|\widehat{m}_p - \theta| \leq \varepsilon\right) = 1. \tag{2.11}$$

The estimator $\widehat{m}_p$ is strongly consistent if

$$P\left(\lim_{p\to\infty} \widehat{m}_p = \theta\right) = 1. \tag{2.12}$$

Definition 2.18 *An estimator $\widehat{m}_p$ of θ is unbiased if*

$$E\left(\widehat{m}_p\right) = \theta \quad \text{for all } p. \tag{2.13}$$

Additionally, $\widehat{m}_p$ is asymptotically unbiased if

$$\lim_{p\to\infty} E\left(\widehat{m}_p\right) = \theta. \tag{2.14}$$

Unbiasedness of an estimator will ensure that the amount of times the estimator $\widehat{m}_p$ over- and underestimates θ will cancel each other out in the long run. Note from (2.9) that $\widehat{m}_p$ is unbiased.

The characteristics of consistency and unbiasedness of the estimator $\widehat{m}_p$ are justified by the Central Limit Theorem (CLT).

Theorem 2.19 *Let X_i, $i = 1, ..., p, ...$ be independent identically distributed (iid) random variables with mean $E(X_i) = \mu$ and $var(X_i) = \sigma^2 < \infty$. Then letting $\overline{X} = \frac{1}{p}\sum_{i=1}^{p} X_i$, we have*

$$\frac{(\overline{X} - \mu)}{\sigma/\sqrt{p}} \rightarrow N(0,1) \tag{2.15}$$

in distribution as $p \rightarrow \infty$.

Using the CLT, we can define

$$X_i = \widehat{\theta}_i$$

and thus estimate θ by $\overline{X} = \frac{1}{p}\sum_{i=1}^{p} X_i$, or using the terminology introduced in (2.8), we can use

$$\widehat{m}_p = \frac{1}{p}\sum_{i=1}^{p} \widehat{\theta}_i.$$

The CLT ensures that an asymptotic confidence interval exists for θ. Let us define the sample variance as

$$s^2_{\widehat{m}_p} = \sum_{i=1}^{p} \frac{\left(\widehat{\theta}_i - \theta\right)^2}{p-1}.$$

Then using (2.15), with $X_i = \widehat{\theta}_i$, $\mu = \theta$, for large p, the $100\,(1-\alpha)\,\%$ confidence interval for θ is

$$\theta \in \left[\widehat{m}_p - \frac{\sigma}{\sqrt{p}}\, z_{\frac{\alpha}{2}},\ \widehat{m}_p + \frac{\sigma}{\sqrt{p}}\, z_{\frac{\alpha}{2}}\right] \tag{2.16}$$

where $z_{\alpha/2}$ is the $1 - \frac{\alpha}{2}$ quantile of the standard normal distribution.

From this, we can see that $\widehat{m}_p$ is strongly consistent and asymptotically unbiased because for fixed and known σ,

$$\lim_{p\to\infty} \frac{\sigma}{\sqrt{p}} \to 0,$$

implying, from (2.16) and (2.9),

$$\lim_{p\to\infty} \widehat{m}_p = \theta$$

and also

$$P\left(\lim_{p\to\infty} \widehat{m}_p = \theta\right) = 1 \quad \text{almost surely,}$$

thereby, arriving at the conditions defined in (2.12) and (2.14).

Consider again the confidence interval in (2.16), it is obvious that to change the width of the interval, one only needs to be concerned with $\frac{\sigma}{\sqrt{p}}$ since $z_{\frac{\alpha}{2}}$ is an arbitrary constant. By reducing the width of the interval, the reliability of $\widehat{m}_p$ as an estimator is increased. One does so by either increasing p or decreasing σ. Decreasing σ by a factor of some constant K has the same effect as increasing p by a factor of K^2.

Obviously, both methods increase the computational cost, with varying p being the more expensive of the two. However, since we seek to minimize computational cost for an acceptable level of accuracy, we choose to use variance reduction techniques for this work.

2.4 Measures of Performance

Consider two unbiased estimators $\hat{\theta}^1$ and $\hat{\theta}^2$ of an option price θ, where one has a lower variance,

$$\sigma_1^2 < \sigma_2^2.$$

Both are unbiased, hence from (2.9),

$$E\left[\hat{\theta}^1\right] = E\left[\hat{\theta}^2\right] = \theta.$$

Also, the variance of the estimator $\widehat{\theta}^k$, $k = 1, 2$, is known from (2.10) to be

$$var\left[\hat{\theta}^k\right] = \frac{\sigma_k^2}{p}.$$

We need to choose the best estimator based on the criterion of variance levels and computational costs involved.

2.4.1 Level of Variance

Based on the confidence interval in (2.16), we know that lowering variance will bring the estimator $\widehat{m}_p$ closer to θ. So, for some fixed p, using variance as a measure of

performance, one should select $\hat{\theta}^1$ as

$$\begin{aligned} \frac{\sigma_1^2}{p} &< \frac{\sigma_2^2}{p} \\ &\Longrightarrow \sigma_1^2 < \sigma_2^2. \end{aligned} \tag{2.17}$$

If the number of simulations p is fixed for both estimators, it would be inefficient to use $\hat{\theta}^2$ as one gets a more variant estimator for the same amount of work. However, if p_k, $k = 1, 2$, is the required simulation to generate each $\hat{\theta}^k$ then there should be a way to measure the work each p_k entails. The *level of variance* as a measure of performance then fails to reflect the possible differences in computational effort.

2.4.2 Level of Efficiency

As an additional measure of performance that incorporates computational effort, we will use the *level of efficiency*. We choose to use the measure used by (Boyle et al, 1997), discussed in (Bratley et al, 1987), (Hammersley and Handscomb, 1964) and further extended by (Glynn and Whitt, 1992). We will now review the measure of (Boyle et al, 1997).

Consider the standard error term from (2.16) again:

$$\frac{\sigma_k}{\sqrt{p}}, \quad k = 1, 2.$$

(Boyle et al, 1997) accounts for the computing time, t and the work required to generate

a single path, F. Setting

$$p = \left\lfloor \frac{t_k}{F_k} \right\rfloor,$$

assuming $\left\lfloor \frac{t_k}{b_k} \right\rfloor$ is a positive integer, the estimators introduced in section 2.4 become

$$\hat{\theta}^k = \frac{F_k}{t_k} \sum_{i=1}^{p} \hat{\theta}_i^k,$$

and the standard error term is revised to

$$\frac{\sigma_k}{\sqrt{\frac{t_k}{F_k}}} = \sigma_k \sqrt{\frac{F_k}{t_k}}.$$

For the same amount of computing time, $t_k = t$, the decision rule is to choose $\hat{\theta}^1$ over $\hat{\theta}^2$ if

$$\begin{aligned} \sigma_1\sqrt{F_1} &< \sigma_2\sqrt{F_2} \\ &\Longrightarrow \sigma_1^2 F_1 < \sigma_2^2 F_2. \end{aligned} \tag{2.18}$$

2.4.3 Pricing Options

Consider now the case of option pricing, with the underlying vector of securities $\left(S_t^k\right)$, $k = 1, ..., d$. The stochastic differential equation in (2.1) governs the evolution of S_t^k. Recall that for a payoff function, f, we can price an option under the risk-neutral measure by

$$\theta = e^{-rT} E\left(f\left(S_T\right) | S_0\right).$$

This would require the integration of $f(S_T)$ over its state space, however, if the exact distribution of S_T is not known, then this would be impossible.

So, using the methodology of Monte Carlo discussed in section 2.3.1, we sample p independently generated prices, $S_T^{k,i}$ of S_T^k, $k = 1, ..., d$, $i = 1, ..., p$. Using the arguments presented in section 2.3.1, we can use this sample average to estimate the option price to be

$$\frac{e^{-rT}}{p} \sum_{i=1}^{p} f\left(S_T^{1,i}, ..., S_T^{d,i}\right),$$

where r is the risk-free interest rate.

To generate the sample values, $S_T^{k,i}$ of the asset price S_T^k, one needs to simulate the paths of the stochastic processes dictated by the SDE in (2.1). We choose to use various time discrete approximations to the SDE over the horizon period, $[0, T]$ which converge to the continuous time stochastic processes if certain criterion are met. This choice is mainly motivated by the fact that sample paths of Ito processes receive the nondifferentiability characteristic of the corresponding Wiener processes. Therefore, time discrete approximations will seek to approximate these paths.

2.5 Discretization Schemes

2.5.1 Convergence

For the discretization schemes, we partition the time interval $[0, T]$ into n equal subintervals. A time discrete approximation $S_t^{\Delta t}$ of (2.1) reduces it to a stochastic difference equation, characterized by the mesh $\Delta t = \frac{T}{n}$. In order to consider the effectiveness of $S_t^{\Delta t}$ as a good pathwise approximator of S_t, we need to introduce the *absolute error* measure

$$E(|S_T - S_T^{\Delta t}|),$$

which is the expected value of the magnitude of the difference between the Ito process S_t and the approximation $S_t^{\Delta t}$.

One then says that $S_t^{\Delta t}$ *converges strongly* to S_t *with order* $\gamma > 0$ at time T, if

$$E(|S_T - S_T^{\Delta t}|) \leq C(\Delta t)^{\gamma} \tag{2.19}$$

where C is a positive constant independent of Δt.

We should note that generally,

$$\frac{1}{p}\sum_{i=1}^{p} f(S_T^{1,i}, ..., S_T^{d,i})$$

is a biased estimator of

$$E(f(S_T)|S_0).$$

However, if $S_t^{\Delta t}$ converges strongly, then it is asymptotically unbiased (please see (2.14) for elaboration). Therefore, we need to choose n to be sufficiently large to ensure that the expectation of $\frac{1}{p}\sum_{i=1}^{p} f(S_T^{1,i}, ..., S_T^{d,i})$ is reasonably close to $E(f(S_T)|S_0)$. This is seen in Figures 2.4 and 2.5 where increasing the time-steps, n, improves the pathwise approximation of the discretization scheme. Henceforth, the distinction between $S_t^{\Delta t}$ and S_t is dropped to simplify notations.

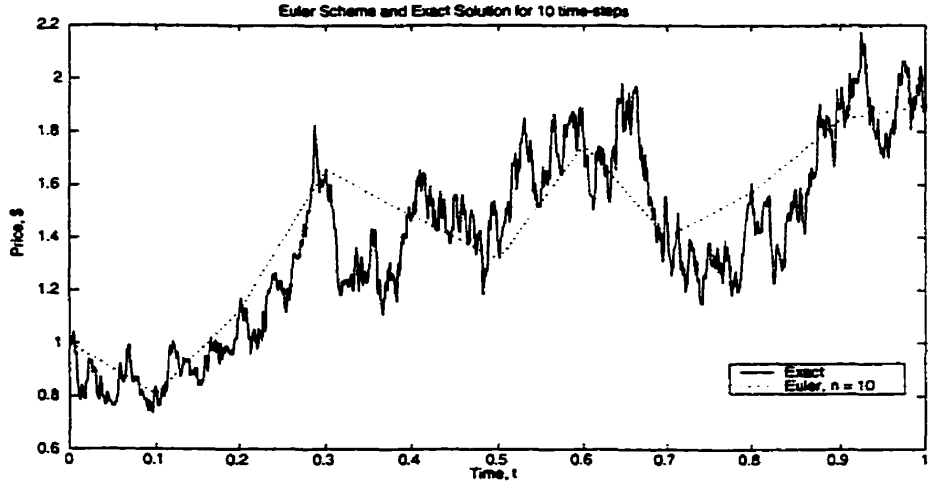

Figure 2.4: Euler Scheme with 10 Time-Steps

We recall that in pricing an option, we require the expectation $\frac{1}{p}\sum_{i=1}^{p} f(S_T^{1,i}, ..., S_T^{d,i})$ to approximate $E(f(S_T)|S_0)$. For this case, it is also possible to approximate the probability distribution of S_T instead of the actual path of S_T, since we only require the expected value of the terminal distribution. In such a case, we do not require such a

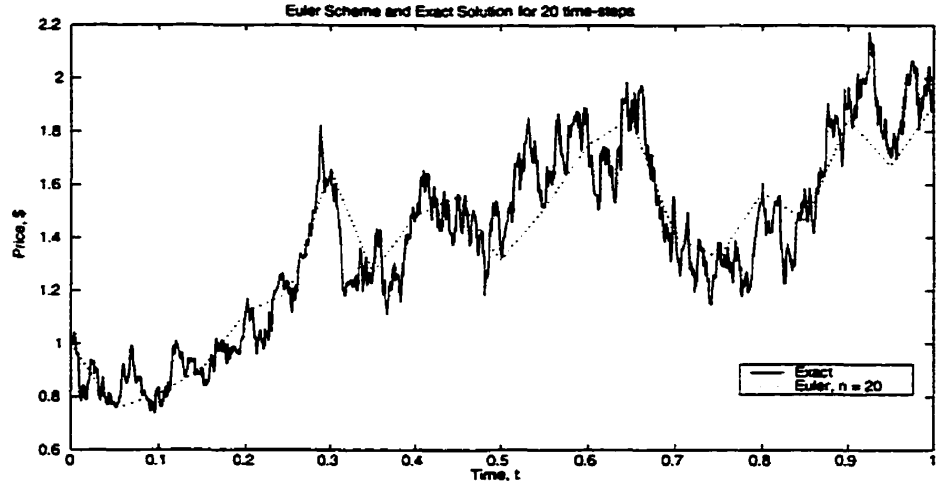

Figure 2.5: Euler Scheme with 20 Time-Steps

strong form of convergence, and we can define the *mean error* as

$$E\left(f(S_T)\right) - E(f(S_T^{\Delta t}))\,.$$

The discrete scheme $S_t^{\Delta t}$ is said to *converge weakly with order* $\beta > 0$ at time T, if for each $f \in C^{2(\beta+1)}$, which, along with its derivatives up to $[2(\beta+1)]$, have polynomial growth, we have

$$\left|E\left(f\left(S_T\right)\right) - E\left(f\left(S_T^{\Delta t}\right)\right)\right| \leq C\left(\Delta t\right)^{\beta} \tag{2.20}$$

where $C^{2(\beta+1)}$ is the space of continuously differentiable functions up to order $[2(\beta+1)]$.

We will focus on discrete schemes with strong order of convergence γ because we require the simulation models to be applicable to different types of options with various payoffs. Therefore, a good pathwise approximation of the Ito process is considered more essential and from this point onwards, we only discuss time discrete approximations

with respect to the strong convergence criterion in (2.19).

2.5.2 Euler Discretization

Consider an initial value problem

$$\begin{aligned} \frac{ds}{dt} &= \mu(s,t) \\ s(t_0) &= s_0 \end{aligned}$$

The deterministic Euler method is used to approximate the continuous time differential equation. Using the same notations introduced in the previous section and also considering t as an integer representing $t \cdot \frac{T}{n}$, we have

$$s_{t+1} = s_t + \mu(s_t, t)\Delta t$$

where $t = 0, ...n$. In a stochastic context, the differential equation becomes an SDE, for example, of the form (2.1), with an additional Ito integral term dW_t^j.

Therefore, the stochastic Euler discretization is a direct extension of the well known deterministic Euler method used above. For the SDE, we need to approximate the additional Ito term

$$\Delta W_t^j = W_{t+\Delta t}^j - W_t^j.$$

Since Wiener processes have independent normally distributed increments, then

$$\Delta W_t^j \sim N(0, \Delta t)$$

which, using the definition of Wiener processes (see definition (2.4)), also means

$$\Delta W_t^j = \sqrt{\Delta t} Z_t^j,$$

where Z_t^j denotes, here and throughout this thesis, a standard Gaussian random variable.

This implies that

$$dW_t^j \approx \sqrt{\Delta t} Z_t^j.$$

Hence, the Euler scheme for (2.1) becomes

$$S_{t+1}^k = \underbrace{S_t^k + \mu^k(S_t, t)\Delta t}_{\text{Deterministic Euler}} + \underbrace{\sum_{j=1}^{d} \sigma^{k,j}(S_t, t)\sqrt{\Delta t} Z_t^j}_{\text{Stochastic Term}} \tag{2.21}$$

for $k = 1, ..., d$, $t = 0, ..., n$. The discretized Euler scheme (and other time discrete schemes), as a method of approximating the Ito process at discretization points in time, lends itself very well for implementation on a digital computer. Furthermore, Ito processes inherit the irregularity properties of the associated Wiener process, as seen in Figure 2.1, so by discretizing these schemes, these irregularities will not be as prominent.

Although the Euler scheme is of order 1 in the deterministic context, its strong order of convergence in the stochastic context is only 0.5 since it ignores some first order terms in the stochastic Taylor expansion (see Kloeden and Platen, 1999, Chapter 5, p. 182) of $S_{t+\Delta t}$ around S_t. See Figures 2.4 and 2.5 for some examples of the Euler scheme approximating an Ito process with different time-steps.

2.5.3 Milstein Discretization

The Milstein scheme incorporates all the first order terms of the stochastic Taylor expansion. It has strong order of convergence equal to 1 but extends the Euler scheme. For equation (2.1), the Milstein scheme is written as, see e.g. (Kloeden and Platen, 1999), p. 346,

$$S_{t+1}^k = \underbrace{S_t^k + \mu^k(S_t,t)\Delta t + \sum_{j=1}^{d} \sigma^{kj}(S_t,t)\sqrt{\Delta t}(Z_t^j) +}_{\text{Euler Scheme}} \tag{2.22}$$
$$\underbrace{\sum_{j_1=1}^{d}\sum_{j_2=1}^{d}\sum_{l=1}^{d} \sigma^{lj_1}(S_t,t)\frac{\partial \sigma^{kj_2}}{\partial S_t^l}(S_t,t)I_{j_1 j_2}}_{\text{Additional 1st Order Term}}$$

with

$$I_{j_1 j_2} = \int_t^{t+\Delta t}\int_t^s dW_\tau^{j_1} dW_s^{j_2}. \tag{2.23}$$

Figure 2.6 illustrates the differences in the Euler and Milstein discretization schemes for a sample path of equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

with the following parameters: $S_0 = 1.0$, $\mu = 1.5$, $\sigma = 1.0$. The explicit solution to this is well known, see e.g. (Lamberton and Lapeyre, 1997) Chapter 3, Section 3.4.3, p. 35,

$$S_t = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right).$$

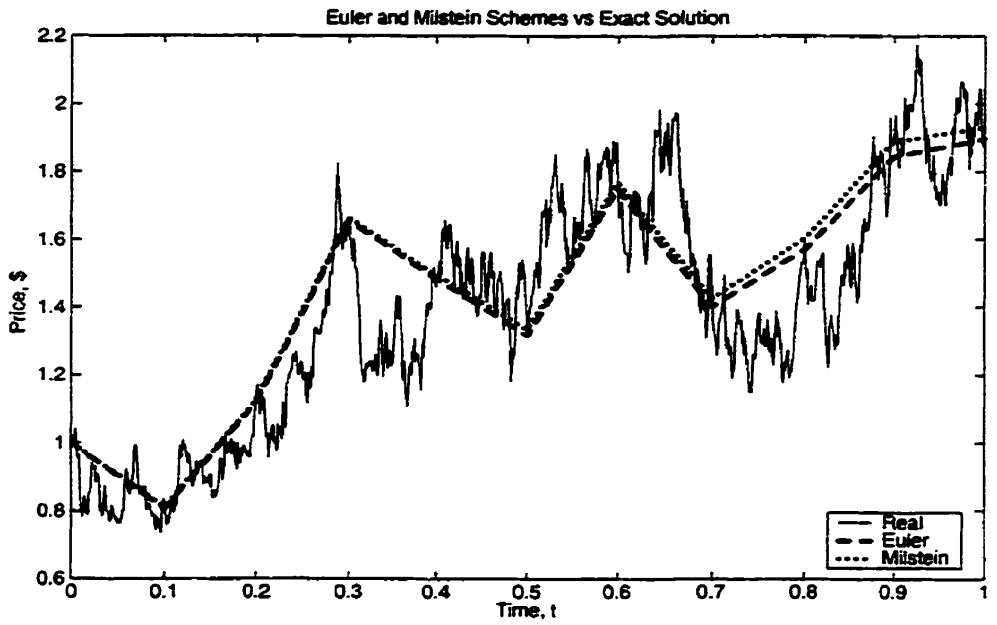

Figure 2.6: Discretization Schemes and the Exact Solution with $n = 10$ time-steps

The effect from the additional double Ito integral term is obvious for t getting close to the time to expiry, T.

The main difficulty arising here not evident in the Euler scheme is the computation of the double Ito integral term, I_{j_1,j_2} and this issue is addressed in the next sections.

2.5.3.0.1 Simplifying Cases There exists three simplifying cases in which the computation of I_{j_1,j_2} for the Milstein discretization scheme becomes feasible and simple. The following cases are particularly relevant in finance.

Case 2.20 *Additive Volatilities*

Suppose the volatility matrix is only a function of time, with

$$\sigma(S_t, t) = \sigma(t)$$

in (2.1). Then obviously the partial derivatives,

$$\frac{\partial \sigma^{kj_2}}{\partial S_t^l}(S_t, t) = 0 \quad \text{for all } k, l$$

in the Milstein expression (2.22), thereby reducing it to the Euler scheme (2.21). In this particular case, the Euler scheme has order of convergence 1.

Case 2.21 *Uncorrelated Assets*

Consider the system (2.1) governing the dynamics of the securities S_t^k. If the volatility matrix is only a function of S_t^k itself and the dW_t^k are uncorrelated, the system becomes

$$dS_t^k = \mu(S_t^1, \cdots, S_t^d)dt + \sigma^k(S_t^k)dW_t^k, \quad k = 1, \cdots, d. \tag{2.24}$$

Then, the partial derivatives

$$\frac{\partial \sigma^{kj_2}}{\partial S_t^l}(S_t, t) = 0 \quad \text{for all } k \neq l,$$

resulting in the volatility matrix $\sigma(S_t, t)$ in (2.1) being diagonal with each diagonal element as $\left\{\sigma^k(S_t^k)\right\}$. As a result of this, the Milstein scheme's (2.22) additional terms will only include $I_{j_1 j_1}$. To compute this double integral, we first give the following.

Theorem 2.22 (Chebyshev's Inequality) *For any random variable X with*

$$E(X) = \mu, \quad \text{and}$$

$$var(X) = \sigma^2,$$

for some constant k,

$$P\{|X-\mu| \geq k\} \leq \frac{\sigma^2}{k^2}.$$

Proposition 2.23 *Consider $\int_a^b \int_a^s dW_\tau dW_s$, then we have*

$$\int_a^b \int_a^s dW_\tau dW_s = \frac{b-a}{2}(Z^2-1) \tag{2.25}$$

where $Z \sim N(0,1)$.

Proof. *First, we have*

$$\begin{aligned}\int_a^b \int_a^s dW_\tau dW_s &= \int_a^b (W_s - W_a) dW_s \\ &= \int_a^b W_s dW_s - W_a(W_b - W_a),\end{aligned}$$

using the definition of the Ito integral in (2.4), we get

$$\int_a^b W_s dW_s \approx \sum_{j=0}^{n-1} W_{s_j}(W_{s_{j+1}} - W_{s_j})$$

where $s_{j+1} = a + \frac{j(b-a)}{n}$, $j = 0, ..., n-1$, then,

$$\begin{aligned}\sum_{j=0}^{n-1} W_{s_j}(W_{s_{j+1}} - W_{s_j}) &= \frac{1}{2}\sum_{j=0}^{n-1} 2W_{s_j}W_{s_{j+1}} - 2W_{s_j}^2 \\ &= \frac{1}{2}\sum_{j=0}^{n-1} 2W_{s_j}W_{s_{j+1}} - W_{s_j}^2 - W_{s_j}^2 + W_{s_{j+1}}^2 - W_{s_{j+1}}^2\end{aligned}$$

Grouping terms to factorize,

$$\begin{aligned} &= \frac{1}{2}\sum_{j=0}^{n-1}(W_{s_{j+1}}^2 - W_{s_j}^2) - \left(W_{s_{j+1}}^2 - 2W_{s_j}W_{s_{j+1}} + W_{s_j}^2\right) \\ &= \frac{1}{2}\sum_{j=0}^{n-1}(W_{s_{j+1}}^2 - W_{s_j}^2) - \left(W_{s_{j+1}} - W_{s_j}\right)^2, \end{aligned}$$

and expanding,

$$\begin{aligned} &= \frac{1}{2}[(W_{s_1}^2 - W_{s_0}^2) - (W_{s_1} - W_{s_0})^2 + \\ &\quad (W_{s_2}^2 - W_{s_1}^2) - (W_{s_2} - W_{s_1})^2 + \ldots + \\ &\quad (W_{s_n}^2 - W_{s_{n-1}}^2) - \left(W_{s_n} - W_{s_{n-1}}\right)^2] \\ &= \frac{1}{2}[W^2(b) - W^2(a)] - \frac{1}{2}\sum_{j=0}^{n-1}\left(W_{s_{j+1}} - W_{s_j}\right)^2. \end{aligned}$$

First, let

$$\Delta W_j = W_{s_{j+1}} - W_{s_j}$$

and set

$$\varepsilon_n = \sum_{j=0}^{n-1}(\Delta W_j)^2.$$

Recall that $\Delta W_j \sim N(0, s_{j+1} - s_j)$, *so*

$$E(\varepsilon_n) = \sum_{j=0}^{n-1} E\,(\Delta W_j)^2 = \sum_{j=0}^{n-1} var\,(\Delta W_j) = \sum_{j=0}^{n-1}(s_{j+1} - s_j).$$

Being a telescoping sum,

$$E(\varepsilon_n) = b - a.$$

Now, the variance is

$$\begin{aligned} var\,(\varepsilon_n) &= \sum_{j=0}^{n-1} var\left[(\Delta W_j)^2\right] \\ &= \sum_{j=0}^{n-1} \left\{ E\left[(\Delta W_j)^2\right]^2 - \left[E\,(\Delta W_j)^2\right] \right\} \\ &\leq \sum_{j=0}^{n-1} E\,(\Delta W_j)^4 . \end{aligned}$$

Recall that the moment generating function for ΔW_j *is*

$$m_{\Delta W_j}(u) = e^{\frac{(s_{j+1}-s_j)^2}{2}u^2}$$

and the fourth derivative of this evaluated at $u = 0$ *will yield*

$$E(\Delta W_j)^4 = m_{\Delta W_j}(0) = 3\,(s_{j+1} - s_j)^2 ,$$

so,

$$var\,(\varepsilon_n) \leq 3 \sum_{j=0}^{n-1} (s_{j+1} - s_j)^2 .$$

Now, using Chebyshev's Inequality in Theorem (2.22), with $X = \varepsilon_n$,

$$\begin{aligned} P\left\{|\varepsilon_n - E(\varepsilon_n)| \geq k\right\} &\leq \frac{var\,(\varepsilon_n)}{k^2} \\ &\leq \frac{3}{k^2} \sum_{j=0}^{n-1} (s_{j+1} - s_j)^2 , \end{aligned}$$

but since the size of the equal increments is

$$s_{j+1} - s_j = \frac{b-a}{n},$$

then

$$\sum_{j=0}^{n-1}(s_{j+1}-s_j)^2=\sum_{j=0}^{n-1}\left(\frac{b-a}{n}\right)^2=n\cdot\frac{(b-a)^2}{n^2}=\frac{(b-a)^2}{n}.$$

Then, as $n\longrightarrow\infty$,

$$\frac{3}{k^2}\sum_{j=0}^{n-1}(s_{j+1}-s_j)^2=\frac{3}{k^2}\frac{(b-a)^2}{n}\longrightarrow 0,$$

thus $P\{|\varepsilon_n-E(\varepsilon_n)|\geq k\}\longrightarrow 0$, *and so*

$$\varepsilon_n\longrightarrow E(\varepsilon_n)=b-a$$

in probability. So, finally,

$$\begin{aligned}\int_a^b\int_a^s dW_\tau dW_s &= \int_a^b W_s dW_s - W_a(W_b-W_a),\\ &\simeq \left\{\frac{1}{2}[W^2(b)-W^2(a)]-\frac{1}{2}(b-a)\right\}-W_a(W_b-W_a)\\ &= \left\{\frac{1}{2}W^2(b)+\frac{1}{2}W^2(a)-W(a)W(b)\right\}-\frac{1}{2}(b-a)\\ &= \frac{1}{2}[W(b)-W(a)]^2-\frac{1}{2}(b-a)\end{aligned}$$

but $W(b)-W(a)\sim N(0,b-a)$, *then*

$$\begin{aligned}\int_a^b\int_a^s dW_\tau dW_s &= \frac{1}{2}[\sqrt{b-a}Z]^2-\frac{1}{2}(b-a)\\ &= \frac{b-a}{2}\left(Z^2-1\right).\end{aligned}$$

■

Now, regarding the calculation of $I_{j_1j_1}$, (2.25) yields

Corollary 2.24 *In the above notations*

$$\begin{aligned} I_{j_1 j_1} &= \int_t^{t_{n+1}} \int_t^s dW_\tau^{j_1} dW_s^{j_1} \\ &= \frac{\Delta t}{2}\left(\left(Z_t^{j_1}\right)^2 - 1\right) \end{aligned}$$

where $\Delta t = t_{n+1} - t_n$.

Proposition 2.25 *Commutative Systems*

The SDE (2.1) is said to be "commutative", if for all j_1, j_2 and k,

$$\sum_{l=1}^{d} \sigma^{lj_1}(S_t, t) \frac{\partial \sigma^{kj_2}}{\partial S_t^l}(S_t, t) = \sum_{l=1}^{d} \sigma^{lj_2}(S_t, t) \frac{\partial \sigma^{kj_1}}{\partial S_t^l}(S_t, t). \tag{2.26}$$

In this case the last term in (2.22) can be split into the three forms of the Ito integral,

$I_{j_1 j_2}$

$$\begin{aligned} &\sum_{j_1=1}^{d}\sum_{j_2=1}^{d}\sum_{l=1}^{d} \sigma^{lj_1}(S_t, t) \frac{\partial \sigma^{kj_2}}{\partial S_t^l}(S_t, t) I_{j_1 j_2} \\ = &\sum_{j_1=2}^{d}\sum_{j_2=1}^{j_1-1}\sum_{l=1}^{d} \sigma^{lj_1}(S_t, t) \frac{\partial \sigma^{kj_2}}{\partial S_t^l}(S_t, t) I_{j_1 j_2} + \\ &\sum_{j_1=2}^{d}\sum_{j_2=1}^{j_1-1}\sum_{l=1}^{d} \sigma^{lj_1}(S_t, t) \frac{\partial \sigma^{kj_2}}{\partial S_t^l}(S_t, t) I_{j_2 j_1} + \\ &\sum_{j_1=1}^{d}\sum_{l=1}^{d} \sigma^{lj_1}(S_t, t) \frac{\partial \sigma^{kj_1}}{\partial S_t^l}(S_t, t) I_{j_1 j_1}. \end{aligned} \tag{2.27}$$

$I_{j_1 j_2}$ for $j_1 \neq j_2$ cannot be expressed as easily as $I_{j_1 j_2}$ for $j_1 = j_2$, so we use an identity from (Kloeden and Platen, 1999), (3.15), p. 348

$$I_{j_1 j_2} + I_{j_2 j_1} = \Delta W^{j_1} \Delta W^{j_2},$$

using (2.27) we get, after factorizing,

$$\begin{aligned}
&\sum_{j_1=1}^{d}\sum_{j_2=1}^{d}\sum_{l=1}^{d}\sigma^{lj_1}(S_t,t)\frac{\partial\sigma^{kj_2}}{\partial S_t^l}(S_t,t)I_{j_1j_2} \\
= \;& \Delta t\sum_{j_1=2}^{d}\sum_{j_2=1}^{j_1-1}\sum_{l=1}^{d}\sigma^{lj_1}(S_t,t)\frac{\partial\sigma^{kj_2}}{\partial S_t^l}(S_t,t)(Z_t^{j_1}Z_t^{j_2}) + \\
& \frac{\Delta t}{2}\sum_{j_1=1}^{d}\sum_{l=1}^{d}\sigma^{lj_1}(S_t,t)\frac{\partial\sigma^{kj_1}}{\partial S_t^l}(S_t,t)((Z_t^{j_1})^2-1). \qquad (2.28)
\end{aligned}$$

One typical example of this case in finance is when the kth security S_t^k follows the dynamics:

$$dS_t^k = \mu(S_t^1,\cdots,S_t^d,t)dt + \xi^k(S_t^k)dW_t^k \quad \text{for } k=1,\cdots,d \qquad (2.29)$$

where the W_t^k are correlated standard Wiener processes. As such, it is necessary to simulate a multivariate normal distribution, discussed next.

2.5.3.1 Simulating Multivariate Normal Distributions

Suppose $\mathbf{Z} = (Z^1,...,Z^d)$ where

$$Z^k \overset{iid}{\sim} N(0,1), \qquad (2.30)$$

for $k = 1,...,d$ and *iid* denotes identically and independently distributed, then let

$$\mathbf{W} = \mathbf{AZ} \qquad (2.31)$$

where $\mathbf{W} = (W^1, ..., W^d)$ and $\mathbf{A}$ is a nonsingular $d \times d$ matrix. The linear transformation in (2.31) has

$$\begin{aligned} E(\mathbf{W}) &= E(\mathbf{AZ}) \\ &= \mathbf{A}E(\mathbf{Z}) \\ &= 0 \end{aligned}$$

and the variance-covariance matrix as

$$\begin{aligned} \sum_{\mathbf{W}} &= E(\mathbf{WW}^T) \\ &= E\left[(\mathbf{AZ})(\mathbf{AZ})^T\right] \\ &= E\left[\mathbf{AZZ}^T\mathbf{A}^T\right] \\ &= \mathbf{A}E(\mathbf{ZZ}^T)\mathbf{A}^T \end{aligned}$$

but

$$\mathbf{ZZ}^T = \begin{pmatrix} Z^1 \\ \vdots \\ Z^d \end{pmatrix} \begin{pmatrix} Z^1 \\ \vdots \\ Z^d \end{pmatrix}^T = \begin{pmatrix} (Z^1)^2 & Z^1Z^2 & \cdots & Z^1Z^d \\ Z^1Z^2 & (Z^2)^2 & & Z^2Z^d \\ \vdots & & \ddots & \vdots \\ Z^1Z^d & Z^2Z^d & \cdots & (Z^d)^2 \end{pmatrix}.$$

So

$$E(\mathbf{ZZ}^T) = \begin{pmatrix} E(Z^1)^2 & E(Z^1Z^2) & \cdots & E(Z^1Z^d) \\ E(Z^1Z^2) & E(Z^2)^2 & & E(Z^2Z^d) \\ \vdots & & \ddots & \vdots \\ E(Z^1Z^d) & E(Z^2Z^d) & \cdots & E(Z^d)^2 \end{pmatrix},$$

and using the distribution of Z^k in (2.30), we know that

$$E\left(Z^{k_1}Z^{k_2}\right) = \left\{ \begin{array}{ll} 1, & \text{for } k_1 \neq k_2 \\ 0, & \text{for } k_1 = k_2 \end{array} \right\},$$

getting

$$E\left(\mathbf{Z}\mathbf{Z}^T\right) = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} = \mathbf{I}, \quad \mathbf{I} \text{ being the Identity matrix.}$$

Therefore,

$$\sum_{\mathbf{W}} = \mathbf{A}\mathbf{I}\mathbf{A}^T = \mathbf{A}\mathbf{A}^T.$$

Furthermore, it is easily shown (Lehmann and Casella, 1998, p. 20) with the usual formula for density transformations that the distribution of $\mathbf{W}$ is

$$\begin{aligned} f_{\mathbf{W}}(w) &= f_{\mathbf{Z}}(w) \cdot \left\|\mathbf{A}^{-1}\right\| \\ &= \frac{\left\|\mathbf{A}^{-1}\right\|}{(2\pi)^d} e^{-\frac{1}{2}w^T(\mathbf{A}\mathbf{A}^T)w} \end{aligned}$$

where f denotes the distribution function and $\|.\|$ denotes the Jacobian of the linear transformation (2.31), that is,

$$\mathbf{W} \sim N\left(0, \mathbf{A}\mathbf{A}^T\right).$$

Hence, to simulate random variables $\mathbf{W}$, one needs to find $\mathbf{A}$ for which $\sum_{\mathbf{W}} = \mathbf{A}\mathbf{A}^T$, and a Cholesky Decomposition can be used for this purpose.

By using Cholesky decomposition, the commutative system (2.29) is transformed to that of the form seen in (2.1).

2.5.3.2 Cholesky Decomposition

For details on this method, see (Conte and de Boor, 1972, p. 142) and (Morgan, 1984, p. 282). The goal of this method is to produce an upper triangular matrix **A** through recursive computation by columns. Let the correlation matrix be $\sum = \{\rho_{kj}\}$, $k = 1, ..., d$ and $j = 1, ..., d$,

$$\mathbf{A} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ a_{d1} & a_{d2} & \cdots & a_{dd} \end{pmatrix},$$

so set

$$\sum = \mathbf{A}\mathbf{A}^T,$$

getting

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ a_{d1} & a_{d2} & \cdots & a_{dd} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{d1} \\ 0 & a_{22} & & a_{d2} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{dd} \end{pmatrix}.$$

Solving the system of equations from the matrix multiplication, we get

$$a_{k1} = \begin{cases} 1 & \text{for } k = 1 \\ \rho_{k1} & \text{for } k > 1 \end{cases}, \quad \text{and}$$

$$a_{kj} = \begin{cases} \sqrt{1 - \sum_{c=1}^{j-1} a_{jc}^2} & \text{for } k = j \\ \rho_{kj} - \sum_{c=1}^{j-1} a_{kc} a_{jc} & \text{for } k > j \end{cases}$$

It can be easily shown next that the system (2.29) is commutative.

To account for the correlation in the processes, the volatility matrix for equation (2.29) is as follows

$$\left[\sigma^{kj}\right] = \underbrace{\begin{pmatrix} \xi^1(S_t^1) & 0 & \cdots & 0 \\ 0 & \xi^2(S_t^2) & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \xi^d(S_t^d) \end{pmatrix}}_{\textit{diffusion}} \times \underbrace{\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ a_{d1} & a_{d2} & \cdots & a_{dd} \end{pmatrix}}_{\textit{Cholesky Decomposition of Correlation Matrix}},$$

where a_{kj} are constants. Then dropping (S_t^k) for simpler notation, we get

$$\left[\sigma^{kj}\right] = \begin{pmatrix} a_{11}\xi^1 & 0 & \cdots & 0 \\ a_{21}\xi^2 & a_{22}\xi^2 & & 0 \\ \vdots & & \ddots & \vdots \\ a_{d1}\xi^d & a_{d2}\xi^d & \cdots & a_{dd}\xi^d \end{pmatrix}.$$

Consider the commutativity condition in (2.26), for $k \neq l$,

$$\frac{\partial \sigma^{kj}}{\partial S_t^l} = 0 \quad for\ all\ j = j_1, j_2.$$

Thus (2.26) simplifies to

$$\sigma^{kj_1} \frac{\partial \sigma^{kj_2}}{\partial S_t^k} = \sigma^{kj_2} \frac{\partial \sigma^{kj_1}}{\partial S_t^k}.. \tag{2.32}$$

Since $a_{kj} = 0$ when $j > k$ $(j = j_1, j_2)$, then for $j_1, j_2 = 1, ..., k$, the left-hand side of (2.32) becomes

$$a_{kj_1}\ \xi^k\ a_{kj_2}\ \frac{\partial \xi^k}{\partial S_t^k}$$

while the right-hand side becomes

$$a_{kj_2}\ \xi^k\ a_{kj_1}\ \frac{\partial \xi^k}{\partial S_t^k}..$$

So the revised commutativity condition in (2.32) is met for this case.

Note that if ξ^k is a function of any other security besides S_t^k then the resulting system would not be commutative. In such cases the stochastic Ito integrals $I_{j_1 j_2}$ have to be simulated. This issue is addressed in the next section.

2.5.3.3 Simulating the Multiple Ito Integrals

There are some types of option pricing problems which do not fall into the above mentioned categories. An example is mentioned wherein the volatility of a security depends on other factors, thereby necessitating the simulation of the Ito integral $I_{j_1 j_2}$.

Example 2.26 *Stochastic volatility.*

Consider the system:

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{\sigma_t} S_t dW_t^1 \\ d\sigma_t &= \alpha(\beta - \sigma_t)dt + \gamma\sqrt{\sigma_t} dW_t^2 \end{aligned} \tag{2.33}$$

where W_t^1 and W_t^2 are standard Wiener processes correlated by correlation coefficient ρ ($\rho = 0$ or $\neq 0$), and the identity

$$\gamma^2 \leq 2\alpha\beta$$

must hold to ensure positive real values for σ_t. We first verify that this system is not commutative for the general case with correlation ρ:

Note that the volatility matrix for the system (2.33) above is

$$[\sigma^{kj}] = \begin{pmatrix} \sqrt{\sigma_t} S_t & 0 \\ \gamma\rho\sqrt{\sigma_t} & \gamma\sqrt{\sigma_t(1-\rho)} \end{pmatrix}.$$

Using the commutativity condition (2.26), for $k = 1$, $d = 2$, $j_1 = 1$, $j_2 = 2$, we get

$$\sum_{l=1}^{2} \sigma^{l1}(S_t, t) \frac{\partial \sigma^{12}}{\partial S_t^l}(S_t, t) = 0,$$

while

$$\sum_{l=1}^{2} \sigma^{l2}(S_t, t) \frac{\partial \sigma^{11}}{\partial S_t^l}(S_t, t) = \frac{1}{2}\gamma\sqrt{1-\rho} S_t,$$

so the commutativity condition is not met.

The problem above requires the direct computation of the multiple Ito integral (2.23) and to compute the double Ito integral, one traditional approach is first reviewed and then a direct and efficient numerical scheme is proposed. An analytic method is proposed in (Kloeden and Platen, 1999) Chapter 5, by expanding the Ito integral as a Fourier series and using a truncation as an approximation. This technique is used to approximate another form of integrals termed the Stratonovich integrals in Section 8, Chapter 5 of the above reference. The approximation, c.f. the above reference p. 347, is given as

$$\begin{aligned} I_{j_1 j_2} \approx I^q_{j_1 j_2} = (\frac{1}{2} Z_t^{j_1} Z_t^{j_2} + \sqrt{\rho_q}(Y^{j_1} Z_t^{j_2} - Y^{j_2} Z_t^{j_1}))\Delta t + \\ \frac{\Delta t}{2\pi} \sum_{r=1}^{q} \frac{1}{r} (X_r^{j_1}(\sqrt{2} Z_t^{j_2} + W_r^{j_2}) - X_r^{j_2}(\sqrt{2} Z_t^{j_1} + W_r^{j_1})) \end{aligned} \tag{2.34}$$

where

$$\rho_q = \frac{1}{12} - \frac{1}{2\pi^2} \sum_{r=1}^{q} \frac{1}{r^2}, \tag{2.35}$$

and disregarding the indices, W, X and Y, Z are respectively $2q$ and 2 standard Gaussian random variables. We note that larger values of q give more accurate approximations. For a given q, formula (2.34) requires the generation of $4(1+q)$ standard Wiener processes. The value of q must be chosen (see above reference) so that

$$q = q\,(\Delta t) \geq \frac{K}{\Delta t}$$

for some positive constant K. This condition ensures a strong order of convergence $\gamma = 1$ for the Milstein scheme using the approximation of (2.34).

Another approach presented in (Lari-Lavassani et al, 2000) evaluates the multiple Ito integrals (2.23) by going back to the very definition of the Ito integral. Recall that

$$\int_a^b f(t,\omega)dW_t = \sup \sum_{i=0}^{k-1} f(t_i,\omega)(W_{t_{i+1}} - W_{t_i}),$$

where the supremum is taken over all possible discretizations of the interval $[a,b]$ and k is the number of subintervals.

Proposition 2.27 *Consider the integral*

$$I_{j_1 j_2} = \int_t^{t+\Delta t} \int_t^s dW_\tau^{j_1} dW_s^{j_2},$$

with $t \leq t_i \leq t + \Delta t$ *where*

$$t_i = t + \frac{i\Delta t}{k}, i = 0, 1, ..., k.$$

Let X^i, Y^i *be* $2k+2$ *uncorrelated standard Gaussian random variables. Then, we have*

$$I_{j_1 j_2} \approx \frac{\Delta t}{k} \sum_{i=0}^{k-1} (\sum_{j=0}^{i} X^j) Y^i. \tag{2.36}$$

Proof. First, we recall that

$$\begin{aligned} \int_t^s dW_\tau^{j_1} &\approx \sum_{j=0}^{i-1} (W_{t_{j+1}}^{j_1} - W_{t_j}^{j_1}) \\ &= \sqrt{\frac{\Delta t}{k}} \sum_{j=0}^{i-1} X^j \end{aligned}$$

where t_i will be the nearest point to s in the discretization. Therefore, reapplying this for the second integral, we get

$$\begin{aligned} I_{j_1 j_2} &\approx \sum_{i=0}^{k-1}(\sqrt{\frac{\Delta t}{k}}\sum_{j=0}^{i-1} X^j)(W_{t_{j+1}}^{j_1} - W_{t_j}^{j_1}) \\ &= \sum_{i=0}^{k-1}(\sqrt{\frac{\Delta t}{k}}\sum_{j=0}^{i-1} X^j)\sqrt{\frac{\Delta t}{k}}Y^i \\ &= \frac{\Delta t}{k}\sum_{i=0}^{k-1}(\sum_{j=0}^{i-1} X^j)Y^i. \end{aligned}$$

■

Remark 1 *To use (2.36) for computing (2.23) the following additional identities must be used to be consistent with the Milstein scheme in (2.22)*

$$\begin{aligned} Z_t^{j_1} &= \frac{1}{\sqrt{k}}\sum_{i=0}^{k-1} X^i \text{ and} \\ Z_t^{j_2} &= \frac{1}{\sqrt{k}}\sum_{i=0}^{k-1} Y^i. \end{aligned}$$

Remark 2 *Note that to implement (2.36), for a given number of discretizations k, only $2k$ standard Wiener processes are required, and it is now shown to be numerically more efficient than (2.34).*

We now compare the numerical efficiency of both formulas (2.34) and (2.36) in estimating $I_{j_1 j_2}$. The measure of efficiency here will be the computational cost represented by Flops. The time interval $[t, t+\Delta t]$ is considered to be $[0, 1]$, i.e, $\Delta t = 1$, and we take $k = q = 1000$ to ensure a high level of accuracy for both formulas.

This estimation process is repeated 50000 times to produce two distributions for $I_{j_1 j_2}$. Having done so, we compare these two distributions to see if there is a significant difference in the two methods. This is done by plotting a histogram seen in Figure 2.7 of the differences in values that are less than or equal to 0.0001.

The histogram shows that almost all of the 50000 samples of the differences have no significant differences between them, showing that the two distributions are almost identical as almost all of the 50000 samples have differences very close to 0. The evidence from the graph is also validated by the numerical values of the first four moments whose values are presented in Table 2.1. First, we note that the moments are very close to each other for both methods. Therefore, we can conclude that both formulas approximate $I_{j_1 j_2}$ well, while their major difference is seen in the computational cost column.

Our formula (2.36), relying on the definition of the Ito integral, only requires 4004 Flops to approximate $I_{j_1 j_2}$ whereas formula (2.34) requires 11024 Flops. This means that numerically, formula (2.36) is much more efficient as it requires 64% less computational effort compared to formula (2.34). This is principally due to the number of random variables each formula has to generate in order to approximate $I_{j_1 j_2}$. Formula (2.34) requires $4\,(1+q)$ standard Wiener processes, while formula (2.36) only requires $2k$. For high levels of accuracy, i.e. $q \approx k$, formula (2.34) requires less than half of the number of standard Wiener processes of formula (2.36).

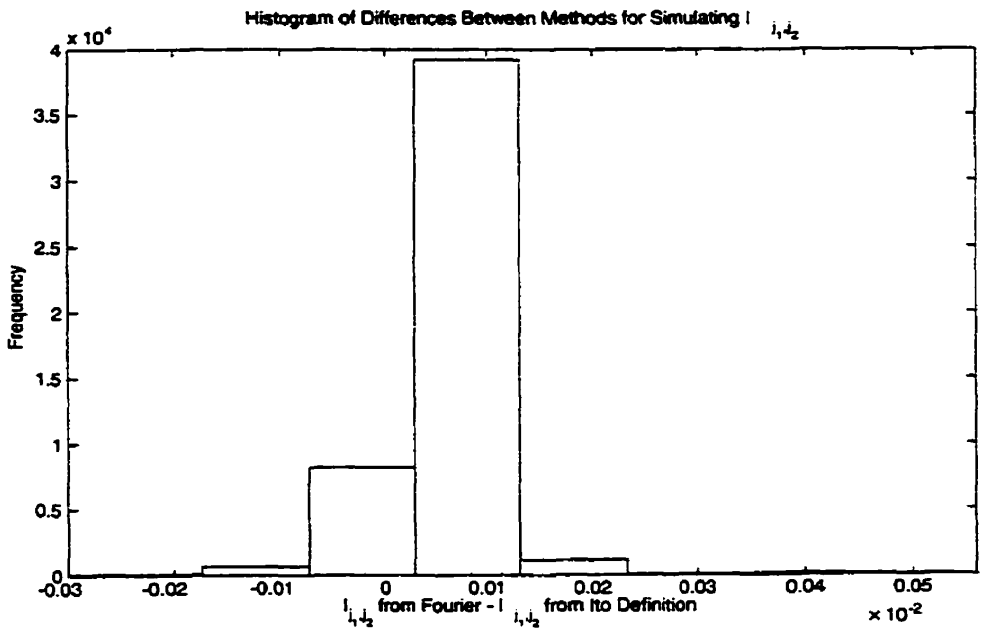

Figure 2.7: Distribution of the Differences of the Double Ito Integral

$I_{j_1 j_2}$ Evaluated via the Formula	Mean $\times 10^{-3}$	Volatility	Skewness	Kurtosis	Cost in Flops $\times 10^7$
(2.34)	1.6694	0.5025	-0.0829	6.7533	11024
(2.36)	2.9077	0.5108	-0.0208	7.0188	4004

Table 2.1: The Four Moments of the Distributions of the Double Ito Integral.

CHAPTER 3
VARIANCE REDUCTION WITH ANTITHETIC VARIATES

Variance reduction techniques improve the efficiency of Monte Carlo methods. They are discussed in (Boyle et al, 1997) or (Clewlow and Carverhill, 1994), among others. We briefly motivated the need for variance reduction in the context of the Central Limit Theorem as a cheaper alternative to increasing the number of simulations in section 2.3.1.

To simplify notations, we still consider t as an integer representing in fact $t \cdot \frac{T}{n}$. For a single path of S_t , we add an additional label i, to be the index for the number of simulations, and note that to simulate this path using the time discrete schemes of (2.21) or (2.22) requires $n \times d$ standard Gaussian random variables.

These $n \times d$ Gaussian random variables define the matrix

$$Z^i = (Z_t^{j,i}),$$

where $j = 1, \ldots, d$ and $t = 1, \ldots, n$. Each Z^i will provide a vector, of length d, of random variables, where each $Z_t^{j,i}$ is a Gaussian random variable at time t, for asset or dimension j along the path indexed by i. In terms of the Monte Carlo method, to improve efficiency, all Z^i's are produced at each instant of time, in matrix form before passing them through the discretization schemes of (2.21) or (2.22), and at each instant of time, the new Z^i's overwrite the previous ones. We choose to do so for memory

management and also to preserve the assumption of independence between random variables.

As a result of this, there is a one-to-one relationship between the matrix Z^i and a sample path of the discretization of (2.1). Hence our attention can be shifted to producing additional matrices Z^i, while preserving the Gaussian symmetry. That is, our emphasis will be on one specific form of variance reduction, namely, antithetic variates.

3.1 Classical Antithetic Variates (CAV)

This particular technique of variance reduction exploits the reduction in variance as a result of random variables being negatively correlated. The reduction in variance arises from the averaging effect of combining results obtained from negatively correlated random variables. For instance, larger than usual estimates of the option price are paired with smaller than average values, the final average of these values will drive the option price closer to the actual mean of the integrand.

The idea of classical antithetic variates is to draw p samples $Z^i = (Z_t^{j,i})$, $i = 1, \ldots, p$, and pairing them with their reflection, $-Z^i$, to obtain a second path. The random inputs from the p pairings $\{(Z^i, -Z^i)\}$ are more regularly distributed than the $2p$ independent samples of Z^i and $-Z^i$ and this arises from the symmetric property of the normal distribution. Theoretically, the sample average of the random samples Z^i is 0. However, due to the randomness of the sample, or obtaining samples which are

less representative of the population, the average value of the Z^i will not usually be zero in actual simulations. We, however, note that the sample mean for the antithetic pairs will always be 0.

Any set of p sample paths $Z^i = (Z_t^{j,i})$, $i = 1, \ldots, p$ yields p samples of $(S_T^{k,i})$, $k = 1, ..., d$. Standard Monte Carlo method then estimates the price of a European option, under the risk-neutral measure, with payoff $f(S_T^1, \ldots, S_T^d)$ by

$$\hat{\theta} = \frac{e^{-rT}}{p} \sum_{i=1}^{p} f(S_T^{1,i}, \ldots, S_T^{d,i}), \quad r = \text{risk-free interest rate.} \tag{3.1}$$

However, with classical antithetic variates (CAV), the price of the option is derived from two sources

$$\hat{\theta}_{CAV} = \frac{e^{-rT}}{p} \sum_{i=1}^{p} \frac{f(S_T^{1,i}, \ldots, S_T^{d,i}) + f(\tilde{S}_T^{1,i}, ..., \tilde{S}_T^{d,i})}{2}, \tag{3.2}$$

where $S_T^{k,i}$, $k = 1, ..., d$ is obtained using (2.21) or (2.22), and $\tilde{S}_T^{k,i}$, $k = 1, ..., d$ is obtained in the same manner except all the Gaussian random variables are the opposite to that of $S_T^{k,i}$. A major advantage of this method is the fact that these Gaussian random variables are generated once but result in two paths.

The averaging effect from the use of antithetic variates is seen in Figure 3.1, which shows the sensitivity of the payoff of a standard European call option with respect to changing initial prices. We see that even for low levels of simulations and time-steps, the averaging effect is dramatic.

The SDE used to price this option is a standard lognormal as in (2.5). The

option has the following parameters:

$$\begin{aligned} \text{Initial price, } S_0 &= 10j, \quad j = 1, ..., 20 \\ \text{Strike price, } K &= 35, \\ \text{Dividends, } \delta &= 0, \\ \text{Volatility, } \sigma &= 0.2. \end{aligned}$$

It is priced using the Euler scheme of (2.21) under the risk-neutral measure, $r = 0.05$, with simulation parameters set at:

$$\begin{aligned} \text{Simulation paths, } p &= 50 \quad \text{and} \\ \text{Number of time-steps, } n &= 30. \end{aligned}$$

3.2 $Z_2 \oplus Z_2$-Symmetric Antithetic Variates

A symmetric extension of the classical antithetic variates technique is proposed in (Lari-Lavassani et al, 2000). We now review this construction. The idea is to produce from a given simulated path, three additional paths while maintaining the symmetry of the Gaussian distributions involved. At the same time, the method seeks to reduce the global variance in a cost effective manner. It extends the CAV method.

This method proceeds as follows: given a discretization scheme, e.g. (2.21) or (2.22), with the matrix of Gaussian variables $Z_t^{j,i}$ as defined in a section 3.1, a second matrix is produced using $(-Z_t^{j,i})$. Both $Z_t^{j,i}$ and $(-Z_t^{j,i})$ have the same distribution.

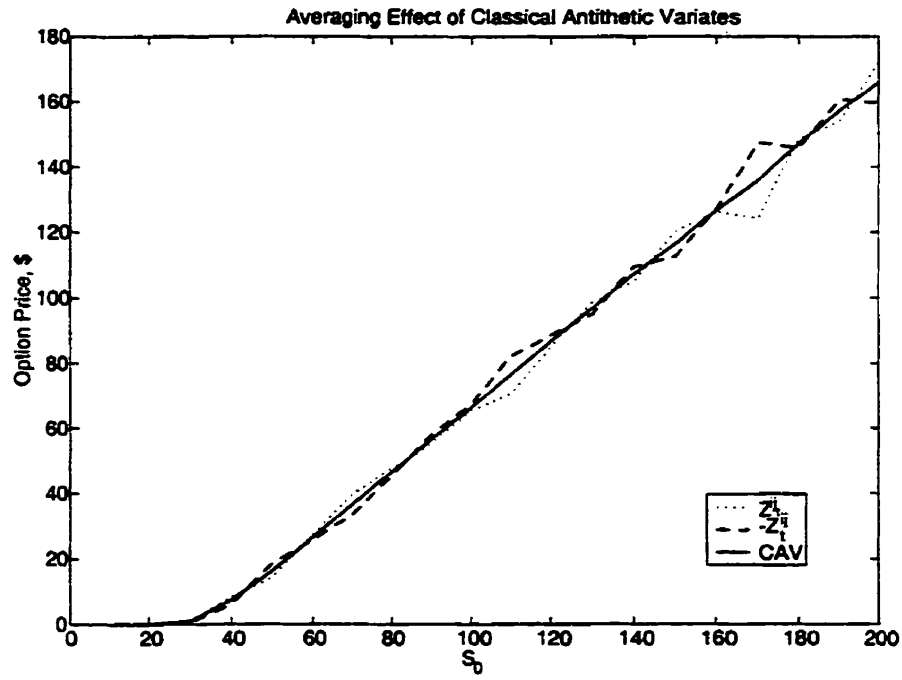

Figure 3.1: Effect of CAV for the payoff of the option

We now give an algebriac interpretation of the CAV method:

Let $Z_2 = \{-1, 1\}$ be the group on two elements, and consider its action on the discrete space $\{Z_t^{j,i}, t = 0, \ldots, n,\ j = 0, \ldots, d\}$ defined by:

$$1 \cdot Z_t^{j,i} = Z_t^{j,i} \quad \text{and}$$

$$(-1) \cdot Z_t^{j,i} = -Z_t^{j,i}.$$

The above action extends naturally to Z^i component-wise. Note that CAV uses the entire orbit of a path under the above action, consisting of Z^i and its opposite path $-Z^i$. The above Z_2-symmetry action leaves the distribution invariant within a given orbit. This is a restatement of the fact that CAV preserves the symmetry of the

Gaussian distributions.

The generalization of the above method consists of enlarging this symmetry group, by considering the four element group $Z_2 \oplus Z_2$ that can be identified with the set of two by two matrices

$$\Gamma_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \Gamma_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \ \Gamma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \ \Gamma_4 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The new action is then defined on the two dimensional grid $\{(Z_{t-1}^{j,i}, Z_t^{j,i})\}$ where t assumes all odd values from 1 to n by:

$$\begin{aligned}
\Gamma_1 \cdot (Z_{t-1}^{j,i}, Z_t^{j,i}) &= (Z_{t-1}^{j,i}, Z_t^{j,i}), \\
\Gamma_2 \cdot (Z_{t-1}^{j,i}, Z_t^{j,i}) &= (-Z_{t-1}^{j,i}, -Z_t^{j,i}), \\
\Gamma_3 \cdot (Z_{t-1}^{j,i}, Z_t^{j,i}) &= (Z_{t-1}^{j,i}, -Z_t^{j,i}), \text{ and} \\
\Gamma_4 \cdot (Z_{t-1}^{j,i}, Z_t^{j,i}) &= (-Z_{t-1}^{j,i}, Z_t^{j,i}).
\end{aligned}$$

which can also be written as

$$\begin{aligned}
\Gamma_1 \cdot Z_t^{j,i} &= Z_t^{j,i}, \\
\Gamma_2 \cdot Z_t^{j,i} &= -Z_t^{j,i}, \\
\Gamma_3 \cdot Z_t^{j,i} &= (-1)^{t+1} Z_t^{j,i}, \text{ and} \\
\Gamma_4 \cdot Z_t^{j,i} &= (-1)^t Z_t^{j,i},
\end{aligned}$$

for all $t = 0, \ldots, n$. This action naturally extends to the matrix path $Z^i = (Z_t^{j,i})$ component-wise.

This is depicted in Figure 3.2. Geometrically, given a path $\Gamma_1 \cdot Z_t^{j,i}$, the second path $\Gamma_2 \cdot Z^i$ moves in the opposite direction of Z^i at all times, the third path $\Gamma_3 \cdot Z^i$ moves in the same direction as Z^i for odd time-steps and in opposite directions for even time-steps, and the fourth path $\Gamma_4 \cdot Z^i$ moves in the opposite direction of the third at all times. All four paths make an orbit under this $Z_2 \oplus Z_2$ action and have the same Gaussian distribution.

It is noted in (Lari-Lavassani et al, 2000) that there are three copies of Z_2 as subgroups of $Z_2 \oplus Z_2$, namely $\{\Gamma_1, \Gamma_2\}$, $\{\Gamma_1, \Gamma_3\}$ and $\{\Gamma_1, \Gamma_4\}$. The case CAV corresponds to $\{\Gamma_1, \Gamma_2\}$, and it is noted that from a variance reduction viewpoint, the other two subgroups alone are not quite as effective. We will see numerically why this is so in Table 4.6. Indeed, $\{\Gamma_1, \Gamma_2\}$ has paths which are highly negatively correlated as one is just the opposite to the other. However, $\{\Gamma_1, \Gamma_3\}$ and $\{\Gamma_1, \Gamma_4\}$ do not share this quality so strongly as at every other time-step, the paths are correlated.

We establish below that variance becomes smaller as the symmetry groups become larger. From this point forth, we label the $Z_2 \oplus Z_2$-symmetric antithetic variates EAV4 (Extended Antithetic Variates with 4 paths).

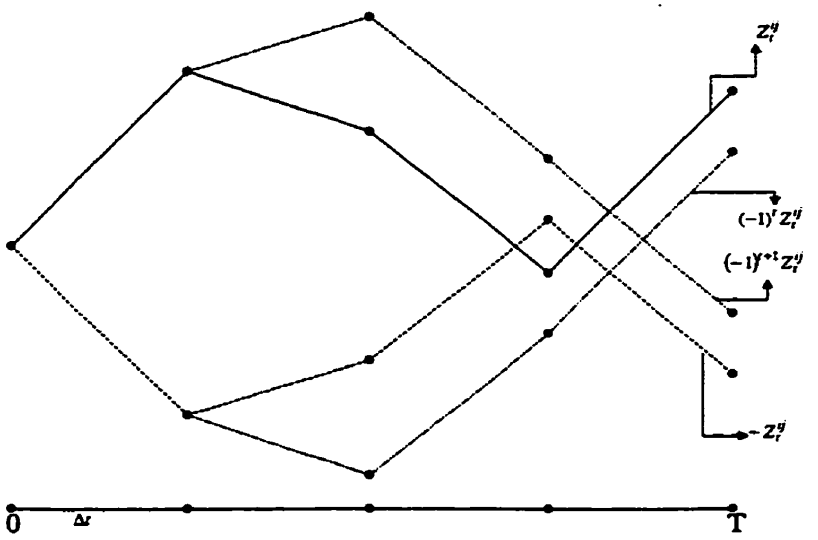

Figure 3.2: Example of the Four $Z_2 \oplus Z_2$ Symmetric Paths

3.3 Implementation and Efficiency of EAV

Once all relevant parameters are fixed,

$$\frac{e^{-rT}}{p} \sum_{i=1}^{p} f(S_T^{1,i}, \ldots, S_T^{d,i})$$

is a function of the random matrix Z^i, $i = 1, \ldots p$. We let

$$g(Z^i) = g(Z_t^{j,i}) := e^{-rT} f(S_T^{1,i}, \ldots, S_T^{d,i})$$

and define

$$\begin{aligned} \hat{\theta}^1 &= \frac{1}{p}\sum_{i=1}^{p} g(Z_t^{j,i}), \\ \hat{\theta}^2 &= \frac{1}{p}\sum_{i=1}^{p} g(-Z_t^{j,i}), \text{ and} \\ \hat{\theta}^k &= \frac{1}{p}\sum_{i=1}^{p} g((-1)^{t+k} Z_t^{j,i}), \text{ for } k = 3, 4 \end{aligned} \tag{3.3}$$

where $\hat{\theta}^1$ is the option price derived from the original path with $Z_t^{j,i}$, $\hat{\theta}^2$ is the price from $(-Z_t^{j,i})$ that moves in an opposite direction to $\hat{\theta}^1$, $\hat{\theta}^3$ derives its price from the original path with $(-1)^{t+3} Z_t^{j,i}$, and similarly for $\hat{\theta}^4$.

Combining the above option prices in various ways, we can now define three additional estimators for the price of the option

$$\begin{aligned} \hat{\theta}_{CAV} &= \frac{\hat{\theta}^1 + \hat{\theta}^2}{2}, \\ \hat{\theta}_{EAV3} &= \frac{\hat{\theta}^1 + \hat{\theta}^2 + \hat{\theta}^3}{3}, \text{ and} \\ \hat{\theta}_{EAV4} &= \frac{\hat{\theta}^1 + \hat{\theta}^2 + \hat{\theta}^3 + \hat{\theta}^4}{4}. \end{aligned}$$

This is done to be consistent with the main idea of antithetic variates, that is, to reduce variance by taking groups of random variables and obtaining a suitable average. We choose to also include $\hat{\theta}_{EAV3}$ to demonstrate that significant efficiency is only gained for estimators associated with subgroups.

Recall that the subgroup $\{\Gamma_1, \Gamma_2\}$ corresponds to $\hat{\theta}_{CAV}$ and $Z_2 \oplus Z_2$ to $\hat{\theta}_{EAV4}$, while $\hat{\theta}_{EAV3}$ does not correspond to any subgroup. More precisely, it can be numerically verified in the Test Case Option, see Figure 3.3 (refer to section 4.3 for details), that

$\hat{\theta}_{EAV3}$ does not perform any better than $\hat{\theta}_{CAV}$, but $\hat{\theta}_{EAV4}$ does so significantly. That is why $\hat{\theta}_{EAV3}$ is eliminated in future cases as it does not consistently reduce variance in comparison to other subgroups. In Figure 3.3, $\hat{\theta}_{EAV3}$ does not perform well for a given level of computational work, especially between 10000 and 20000 Flops. The standard error, in this case is taken to be

$$\sum_{b=1}^{15} \frac{\left(\hat{\theta}_b^k - \hat{\theta}\right)}{\sqrt{15}} \tag{3.4}$$

shows the option price with EAV4 reaching a stable variance very quickly, for only 10000 simulation paths, while CAV shows the same pattern, but at a higher level of variance. The standard Monte Carlo scheme has the worst result, with variance even increasing from 15000 to 30000 paths. EAV3 performs almost as badly as the standard scheme at 15000 paths, furthermore, it does not have the steady behaviour as seen for CAV or EAV4. This therefore justifies the elimination of EAV3 as a variance reduced estimator.

An alternative interpretation to Figure 3.3 is that for a given level of computational cost, EAV4 has the lowest variance while EAV3 is inconclusive.

3.4 Justification of Variance Reduction of EAV

As discussed in section 2.4, we need to justify the extra computational cost involved in variance reduction techniques according to some measures of performance, namely, level of variance and level of efficiency. The following series of propositions from (Lari-

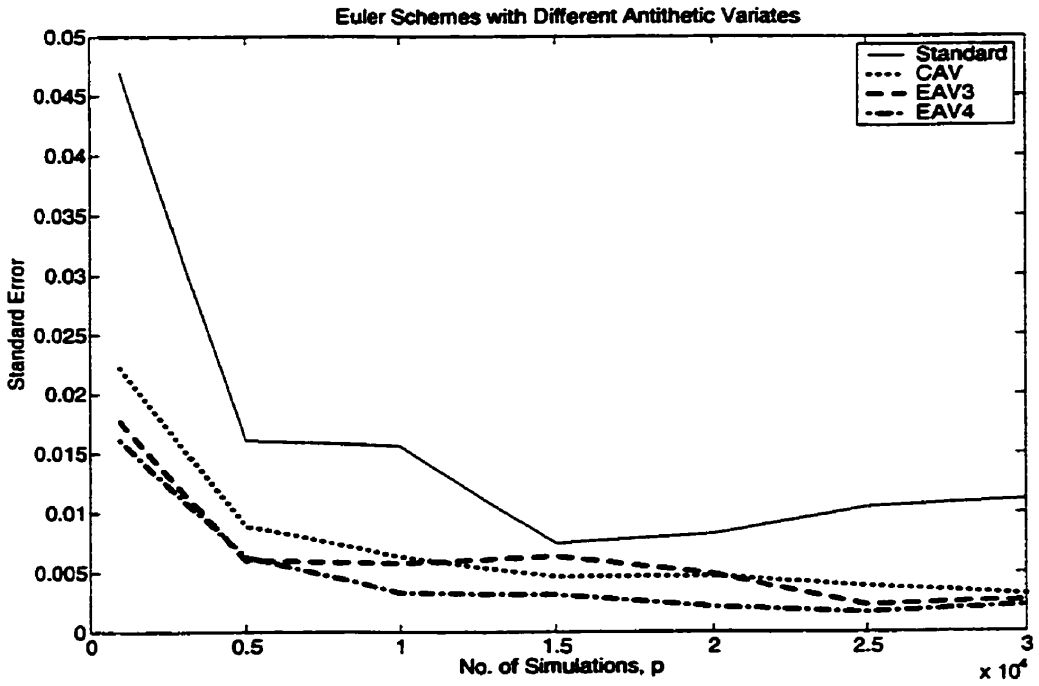

Figure 3.3: Comparison of Euler Schemes with Various Antithetic Variates

Lavassani et al, 2000) explain the benefits in terms of reducing variance and provide an indicator on how effective or costly an estimator is.

We first need to show that by using the various forms of antithetic variates, global variance is actually reduced, and the following proposition does this:

Proposition 3.1

i) *Suppose*

$$cov(g(Z_t^{j,i}), g(-Z_t^{j,i})) \leq var(g(Z_t^{j,i})).$$

Then we have

$$var\left[\hat{\theta}_{CAV}\right] \leq var\left[\hat{\theta}^{(1)}\right].$$

ii) *Suppose*

$$2cov\left(g(Z_t^{j,i}), g((-1)^{t+1} Z_t^{j,i})\right) \leq var\left(g(Z_t^{j,i})\right) + cov\left(g(Z_t^{j,i}), g(-Z_t^{j,i})\right).$$

Then we have

$$var\left[\hat{\theta}_{EAV4}\right] \leq var\left[\hat{\theta}_{CAV}\right] \leq var\left[\hat{\theta}^{(1)}\right].$$

Proof. We have

$$E\left[\hat{\theta}^k\right] = E(g(Z_t^{j,i})),\ k = 1, 2, 3, 4$$

due to the symmetry of the Gaussian distribution, and

$$var\left[g(Z_t^{j,i})\right] = var\left[g(-Z_t^{j,i})\right] = var\left[g((-1)^t Z_t^{j,i})\right] = var\left[g((-1)^{t+1} Z_t^{j,i})\right].$$

Therefore,

$$E\left[\hat{\theta}_{CAV}\right] = \theta$$

and

$$\begin{aligned} var\left[\hat{\theta}_{CAV}\right] &= var\left[\frac{1}{p}\sum_{i=1}^{p}\frac{g(Z_t^{j,i}) + g(-Z_t^{j,i})}{2}\right] \\ &= \frac{1}{4p^2}\sum_{i=1}^{p}\left\{var\left[g(Z_t^{j,i})\right] + var\left[g(-Z_t^{j,i})\right] + 2cov(g(Z_t^{j,i}), g(-Z_t^{j,i}))\right\}, \end{aligned}$$

now since

$$var\left[g(Z^i)\right] = var\left[g(-Z^i)\right],$$

we have

$$var\left[\hat{\theta}_{CAV}\right] = \frac{1}{2p^2}\sum_{i=1}^{p}\left\{var\left[g(Z_t^{j,i})\right] + cov\left(g(Z_t^{j,i}), g(-Z_t^{j,i})\right)\right\}$$

but

$$cov\left(g(Z_t^{j,i}), g(-Z_t^{j,i})\right) \leq var\left[g(Z_t^{j,i})\right]$$

so, we get

$$var\left[\hat{\theta}_{CAV}\right] \leq \frac{1}{p^2}\sum_{i=1}^{p}\left\{var\left[g(Z_t^{j,i})\right]\right\} = var[\hat{\theta}^1]. \tag{3.5}$$

As for the variance of $\hat{\theta}_{EAV4}$, we have

$$\begin{aligned} var\left[\hat{\theta}_{EAV4}\right] &= var[\frac{1}{4p}\sum_{i=1}^{p} g(Z_t^{j,i}) + g(-Z_t^{j,i}) + \\ &\qquad g\left((-1)^t Z_t^{j,i}\right) + g\left((-1)^{t+1} Z_t^{j,i}\right)] \\ &= \frac{1}{(4p)^2}\sum_{i=1}^{p}\{4var\left(g(Z_t^{j,i})\right) + 4\ cov\left(g(Z_t^{j,i}), g(-Z_t^{j,i})\right) \\ &\quad + 8\ cov\left(g(Z_t^{j,i}), g((-1)^{t+1} Z_t^{j,i})\right)\} \end{aligned}$$

because there are 2 paths which move in completely opposite directions to each other, while the others have directions based on whether the time-step is even or not, and using the assumption,

$$2cov\left(g(Z_t^{j,i}), g((-1)^{t+1} Z_t^{j,i})\right) \leq var\left(g(Z_t^{j,i})\right) + cov\left(g(Z_t^{j,i}), g(-Z_t^{j,i})\right),$$

it implies,

$$\begin{aligned} var\left[\hat{\theta}_{EAV4}\right] &\leq var\left[\hat{\theta}_{CAV}\right] \text{ and} \\ var\left[\hat{\theta}_{EAV4}\right] &\leq var\left[\hat{\theta}_{CAV}\right] \leq var(\hat{\theta}^1) \end{aligned}$$

■

To use the level of efficiency as a measure of performance, we need to account for the extra effort required to generate $\hat{\theta}_{EAV4}$ and $\hat{\theta}_{CAV}$. One would expect the effort required for replicating samples of $\hat{\theta}_{EAV4}$ to be at most twice that of $\hat{\theta}_{CAV}$, which is itself at most twice that of $\widehat{\theta}^1$. The following proposition shows that even after accounting for the extra computational cost, $\hat{\theta}_{EAV4}$ is still the best estimator in terms of variance reduction, followed by $\hat{\theta}_{CAV}$ and then the standard Monte Carlo option price, $\widehat{\theta}^1$. Using the level of efficiency we choose to compare

$$\sigma_1 F_1 < \sigma_2 F_2$$

for two estimators where σ_1 is the variance of the estimator and F_1 is the work required to generate the option price.

Proposition 3.2 *Suppose*

$$cov(g(Z_t^{j,i}), g(-Z_t^{j,i})) \leq 0$$

and

$$cov\left(g(Z_t^{j,i}), g((-1)^{t+1} Z_t^{j,i})\right) \leq \varepsilon$$

for some $\epsilon \geq 0$. *Then we have*

$$\begin{aligned} 4var\left[\hat{\theta}_{EAV4}\right] &\leq 2var[\hat{\theta}_{CAV}] + \frac{2\varepsilon}{p} \text{ and} \\ 2var\left[\hat{\theta}_{CAV}\right] &\leq var\left[\hat{\theta}^{(1)}\right] \end{aligned}$$

Proof. Using (3.5) we have

$$\begin{aligned} 2var\left[\hat{\theta}_{CAV}\right] &= \frac{1}{p^2}\sum_{i=1}^{p}\left\{var\left[g(Z_t^{j,i})\right] + cov\left(g(Z_t^{j,i}), g(-Z_t^{j,i})\right)\right\} \\ &\leq \frac{1}{p^2}\sum_{i=1}^{p}\left\{var\left[g(Z_t^{j,i})\right]\right\} = var(\hat{\theta}^{(1)}). \end{aligned}$$

Similarly, this leads to

$$\begin{aligned} 4var\left[\hat{\theta}_{EAV4}\right] &= \frac{1}{p^2}\sum_{i=1}^{p}\{var\left(g(Z_t^{j,i})\right) + cov\left(g(Z_t^{j,i}), g(-Z_t^{j,i})\right) \\ &\quad +2\ cov\left(g(Z_t^{j,i}), g((-1)^{t+1}\ Z_t^{j,i})\right)\} \\ &= 2\ var\left[\hat{\theta}_{CAV}\right] + \frac{2}{p^2}\sum_{i=1}^{p}\left\{cov\left(g(Z_t^{j,i}), g((-1)^{t+1}\ Z_t^{j,i})\right)\right\} \\ &\leq 2var\left[\hat{\theta}_{CAV}\right] + \frac{2}{p}\varepsilon. \end{aligned}$$

■

In the previous proposition, we assume that $cov\left(g(Z_t^{j,i}), g(-Z_t^{j,i})\right) \leq 0$. This is the necessary condition for antithetic variates to help in variance reduction. If this assumption is not valid, then the variance reduction properties of the antithetic variates will not exist and the extra computational effort in generating these variates will be an unjustified cost. Numerically, this is simple to verify as the covariance between the different option prices can be computed, as it will be seen in Section 4.3, Table 4.6.

(Barlow and Proschan, 1975, Section 2.2, p. 29) show that for increasing functions G and H, we have

$$cov\,(G, H) \geq 0.$$

Using this, we next show that

$$cov\left(g(Z_t^{j,i}), g(-Z_t^{j,i})\right) \leq 0.$$

Proposition 3.3

(i) *Suppose the drift, $\mu(.,t)$ and the volatility matrix, $\sigma(.,t)$ are increasing and the payoff of the option $f(S_T^1, ..., S_T^d)$ is monotone with respect to all S_T^j, $j = 1, ..., d$. Then, $g(Z_t^{j,i})$ is monotone with respect to all $Z_t^{j,i}$.*

(ii) *If the function $g(Z_t^{j,i})$ is increasing, then we have*

$$cov\left(g(Z_t^{j,i}), g(-Z_t^{j,i})\right) \leq 0.$$

Proof. Consider two random matrices, with the notation as defined in earlier sections, $V_t^{j,i}$ and $Z_t^{j,i}$ with the following characteristics

$$V_\vartheta^{\rho,i} > Z_\vartheta^{\rho,i} \quad \text{for } (t,j) = (\vartheta,\rho), \quad \text{and} \tag{3.6}$$

$$V_t^{j,i} = Z_t^{j,i} \quad \text{for } (t,j) \neq (\vartheta,\rho) \tag{3.7}$$

Recall that the asset price S_t^j is actually a function of its random matrix, i.e. we introduce the notation

$$S_t^j(Z_t^{j,i})$$

to imply the asset price S_t^j for asset j as a function of the random matrix $Z_t^{j,i}$.

Then (3.7) yields

$$S_{\vartheta-1}^j(V_{\vartheta-1}^{j,i}) = S_{\vartheta-1}^j(Z_{\vartheta-1}^{j,i}) \quad \text{for all } j.$$

Also, at time ϑ, both (3.7) and (3.6) respectively lead to

$$\begin{aligned} S_\vartheta^j(V_\vartheta^{j,i}) &= S_\vartheta^j(Z_\vartheta^{j,i}) \quad \text{for } j \neq \rho, \text{ and} \\ S_\vartheta^\rho(V_\vartheta^{\rho,i}) &> S_\vartheta^\rho(Z_\vartheta^{\rho,i}) \quad \text{for } j = \rho. \end{aligned}$$

Therefore for the next time instant $\vartheta + 1$, we can assume that

$$S_{\vartheta+1}^j(V_{\vartheta+1}^{j,i}) \geq S_{\vartheta+1}^j(Z_{\vartheta+1}^{j,i}),$$

which, in turn, implies

$$S_{\vartheta+2}^j(V_{\vartheta+2}^{j,i}) \geq S_{\vartheta+2}^j(Z_{\vartheta+2}^{j,i}) \quad \text{for all } i.$$

Repeating this until we arrive at time T, we get

$$S_T^j(V_T^{j,i}) \geq S_T^j(Z_T^{j,i}) \text{ for all } i$$

and in general

$$S_T^\varrho(V_T^{j,i}) > S_T^\varrho(Z_T^{j,i}).$$

Therefore, we conclude from this that the payoff $f(S_T^1, ..., S_T^d)$ is monotone with respect to all S_T^j, and since

$$g\left(Z_t^{j,i}\right) = e^{-rT} f(S_T^{1,i}, ..., S_T^{d,i}),$$

it is thus monotone with respect to all $Z_t^{j,i}$.

Now, we know that $-Z_t^{j,i}$ moves in the opposite direction to $Z_t^{j,i}$, so, $g\left(-Z_t^{j,i}\right)$ is a decreasing function, and $-g\left(-Z_t^{j,i}\right)$ must then be increasing. Therefore, using the result from (Barlow and Proschan, 1975), we have

$$cov(g\left(Z_t^{j,i}\right), -g\left(-Z_t^{j,i}\right)) \geq 0.$$

In order to reverse the direction of the inequality, we need to compare the covariances for functions with opposing directions, that is, $g\left(Z_t^{j,i}\right)$, which is increasing, and $g\left(-Z_t^{j,i}\right)$, which is decreasing. If so, the result is

$$cov(g\left(Z_t^{j,i}\right), g\left(-Z_t^{j,i}\right)) \leq 0.$$

■

CHAPTER 4
SIMULATION ACCURACY AND EFFICIENCY

In this chapter we consider a Test Case Option to illustrate the issues of efficiency and accuracy. We measure accuracy by the distance between the estimated option price and the true value, that is, whether an estimate has converged to the true price or not. Whereas efficiency is measured in terms of the effort it requires to yield an option price with low variance.

The Test Case Option is a standard European call option on an underlying with a lognormal process of this form

$$dS_t = (r - \delta)\ S_t\ dt + \sigma\ S_t\ dW_t.$$

The option parameters are provided by Table 4.1.

As a reminder, S_0 denotes the initial asset price, K is the strike or exercise price, σ is the volatility, δ is the continuous dividend rate, T is the time to expiry and r is the risk-free interest rate. Note also that CAV denotes classical antithetic variates and EAV4 the $Z_2 \oplus Z_2$-symmetric antithetic variates. The "true" option value of \$3.985 is determined by a Monte Carlo simulation with the Milstein scheme using EAV4 at $n = 200$ and $p = 200000$. This option value is taken to be the actual option price from this point on. This particular option is also priced in (Rubinstein, 1991), p. 21, using a binomial tree, yielding \$3.99.

Option Parameters	*Values*
S_0	100
K	100
σ	0.1
δ	0.05
T	0.5
r	0.1

Table 4.1: Parameters for Test Case Option

To gauge the effectiveness and accuracy of the Monte Carlo simulations, we consider two issues, namely, the payoff between the reduction of variance and numerical efficiency, and the relationships between convergence schemes and numerical accuracy. As a measure of accuracy, we thus define the *relative error* as

$$Relative\ Error = \widehat{\theta} - \text{“}true\ value\text{”} \tag{4.1}$$

and we also define the standard error as

$$Standard\ Error = \frac{1}{\sqrt{batch}} \sum_{b=1}^{batch} \left(\widehat{\theta}_b^k - \widehat{\theta}\right)^2 \tag{4.2}$$

where $\widehat{\theta} = \sum_{b=1}^{batch} \frac{\widehat{\theta}_b^k}{batch}$ and the batch is the number of times the simulation is repeated.

4.1 Selection of Time-Step

Recall that there exists two forms of convergence, one strong and the other weak (see (2.19) and (3.5)). Both convergence schemes are affected by their time-step size, that is, $\Delta t = T/n$. Therefore, as mentioned earlier, a suitable n has to be selected to ensure the discretization schemes converge to the continuous process. We select an optimal n via the application of the Milstein algorithm with EAV4 to the Test Case Option.

We compute the magnitude of the relative error to be consistent with the strong convergence requirement of (2.19) to determine the impact of varying n. We then plot the results in Figure 4.1. We note the steady decline in the magnitude of the relative error for increasing number of time-steps (which corresponds to a decreasing time-step increment Δt). However, after a threshold point, the magnitude of the relative error seems to increase again. To determine an optimal n, we fix an acceptable error bound, e.g. 0.005, and the best values for n are those that lead to the relative error falling under this acceptable error bound line. For this case, we can see $n \in [30, 100]$ as a possible optimal range. This corresponds to $\Delta t \in [T/100, T/30] \simeq [0.005, 0.02]$. As noted before, after $n = 65$, the curve rises again and although the graph falls into the acceptable error region, one can obtain the same result with less computational effort for $n \in [30, 65]$. Therefore, the most optimal range for n is

$$\begin{aligned} n &\in [30, 65], \text{ or} \\ \Delta t &\in [T/65, T/30] \simeq [0.008, 0.02]. \end{aligned}$$

For the rest of the numerical examples, we will choose $n = 30$.

Next, we also show that increasing the number of time-steps does not improve the volatility of a Monte Carlo simulation. Figure 4.2 shows that for increasing values of n, the volatility does not converge, in fact, it oscillates. This is consistent with the strong convergence criteria that relates only the relative error to the time-step size. Furthermore, in the discussions of variance reduction in section 2.3.1, we note that to reduce the standard error, the two alternatives are variance reduction or increasing the number of simulations, p.

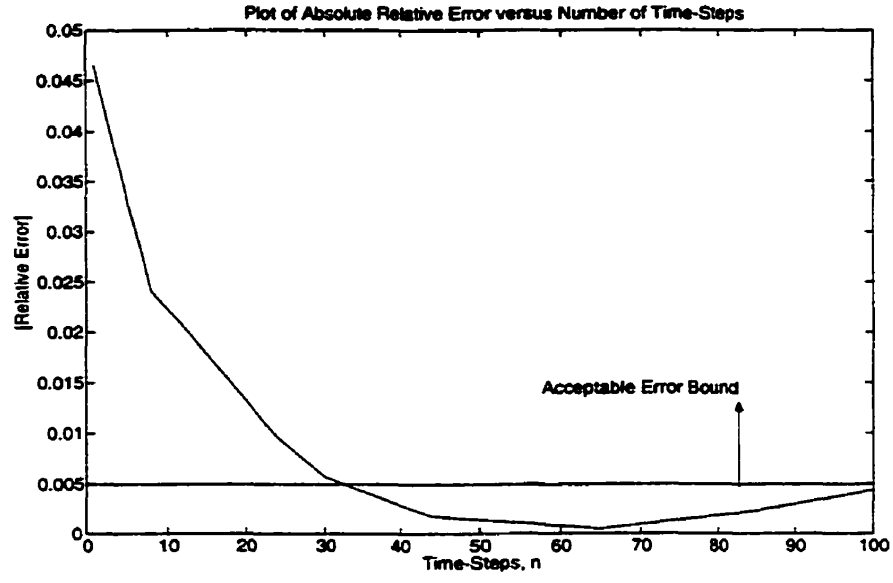

Figure 4.1: Convergence of Option Price with Respect to Number of Time-Steps

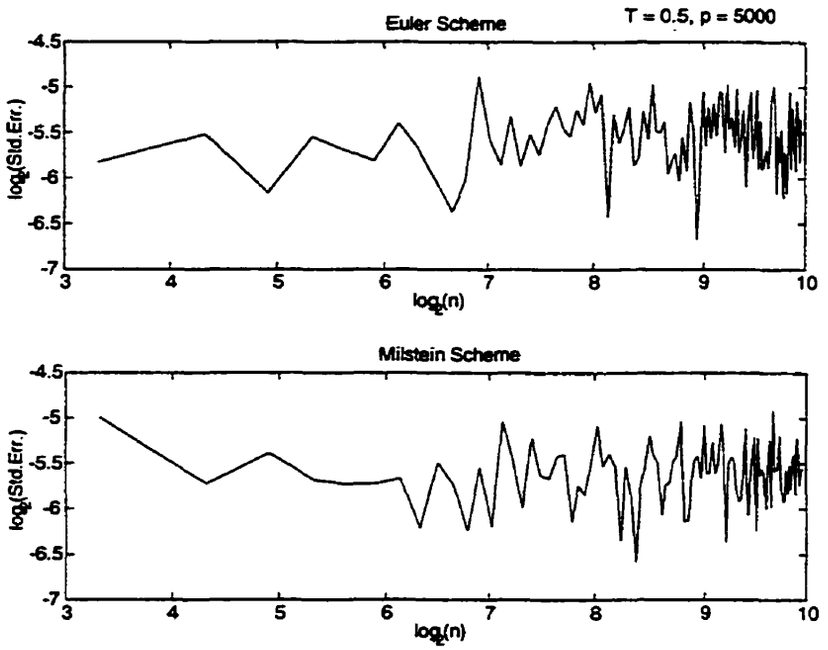

Figure 4.2: Impact of Number of Time-Steps on Volatility

4.2 Convergence Schemes and Numerical Accuracy

We now consider the issue of numerical accuracy. Based on the previous section, we have opted to use $n = 30$ for the numerical results in Tables 4.2 and 4.3. The tables compare the Euler and Milstein schemes with different levels of antithetic variates for the Test Case Option. In Table 4.2, it is seen that for increasing simulation paths, p, the option values start converging to the true value of \$3.985. The exception to this is the Euler scheme with CAV where at 100000 simulation paths, the option value is furthest away from the true value. This is demonstrated in the relative error columns where there is a steady decrease, tending to 0. For Euler, EAV4 performs the best

with a relative error of 0.0001 for $p = 100000$.

In Table 4.3, the results using Milstein also lead to similar conclusions. Milstein with EAV4 outperforms the other versions. In fact, Milstein EAV4 shows very steady convergence with increasing number of simulation paths. This is also seen in the relative error column whereas in the Euler case, the magnitude is only 0.0001. In general Milstein outperforms Euler. We should note also that Milstein produces results which are less variant, as seen in the proximity of the results to each other between successive increases in p. The prices using Milstein hover around 3.98 in all but 3 cases whereas for Euler, there are 6 cases which deviate significantly from 3.98.

The conclusions from the numerical results also provide evidence that by increasing the number of simulations, the width of the confidence interval (2.16) shrinks, so the estimate of the option price will tend to the actual option price for large p.

$n = 30$	*No VR*			*CAV*			*EAV4*		
Paths p	*Value*	*Rel. Error*	*Flops* $\times 10^7$	*Value*	*Rel. Error*	*Flops* $\times 10^7$	*Value*	*Rel. Error*	*Flops* $\times 10^7$
1000	3.9691	-0.0159	0.0153	3.9711	-0.0139	0.0277	3.9829	-0.0021	0.0583
5000	4.0004	0.0150	0.0765	3.9865	0.0150	0.1385	3.9857	0.0007	0.2915
15000	3.9880	0.0030	0.2295	3.9875	0.0025	0.4155	3.9826	-0.0024	0.8745
30000	3.9985	0.0135	0.4590	3.9878	0.0028	0.8310	3.9863	0.0013	1.7490
60000	3.9788	-0.0062	0.9180	3.9870	0.0020	1.6620	3.9860	0.0010	3.4980
100000	3.9872	0.0022	1.5300	3.9903	0.0053	2.7700	3.9851	0.0001	5.8300

Table 4.2: The relative errors and the computations costs in Flops for Euler discretization.

Since Milstein has strong order of convergence 1, we expect Milstein to outperform Euler. This is further seen in Tables 4.4 and 4.5. The convergence patterns are obvious in the tables. For Euler, it is evident that the prices have not converged for this level of p and oscillate, indicating instability. For Milstein, there is steady convergence to 3.985 particularly for the EAV4 case (note that Euler with EAV4 also performs well).

$n = 30$	*No VR*			*CAV*			*EAV4*		
Paths p	*Value*	*Rel. Error*	*Flops* $\times 10^7$	*Value*	*Rel. Error*	*Flops* $\times 10^7$	*Value*	*Rel. Error*	*Flops* $\times 10^7$
1000	3.9689	-0.0161	0.0273	4.0096	0.0246	0.0517	3.9772	-0.0078	0.1063
5000	3.9823	-0.0027	0.1365	3.9781	-0.0069	0.2585	3.9875	0.0025	0.5315
15000	3.9804	-0.0046	0.4095	3.9822	-0.0028	0.7755	3.9874	0.0024	1.5945
30000	3.9860	0.0010	0.8190	3.9858	0.0008	1.5510	3.9882	0.0032	3.1890
60000	3.9843	-0.0007	1.6380	3.9883	0.0033	3.1020	3.9848	-0.0002	6.3780
100000	3.9938	0.0088	2.7300	3.9867	0.0017	5.1700	3.9849	-0.0001	10.6300

Table 4.3: The relative errors and the computations costs in Flops for Milstein discretization.

We next attempt to relate convergence to the number of time-steps, n and the number of simulations, p. This phenomena is seen in the 3 dimensional plot of Figure 4.3. The effect of increasing the number of simulations is more dominant than increasing the number of time-steps. For instance, at the simulation level of approximately 1000 paths, the relative error of the price is very erratic and does not improve even with increasing time-steps. However, for increasing simulation paths, the price gets less erratic, also for increasing time-steps, with the best results at $p = 15000$.

The observations from the numerical results in the previous tables are also verified graphically. Figure 4.5 plots the relative error against the number of simulations

	Standard			*CAV*			*EAV4*		
Time Steps	*Value*	*Rel. Error*	*Flops* $\times 10^7$	*Value*	*Rel. Error*	*Flops* $\times 10^7$	*Value*	*Rel. Error*	*Flops* $\times 10^7$
10	3.9794	-0.0056	0.1590	3.9774	-0.0076	0.2910	3.9734	-0.0116	0.6090
30	3.9876	0.0026	0.4590	3.9835	-0.0015	0.8310	3.9712	-0.0138	1.7490
50	3.9879	0.0029	0.7590	3.9824	-0.0026	1.3710	3.9849	-0.0001	2.8890
80	3.9836	-0.0014	1.2090	3.9847	-0.0003	2.1810	3.9846	-0.0004	4.5990

Table 4.4: Relative Error of Euler With Different Variance Reductions, p = 30000.

	Standard			*CAV*			*EAV4*		
Time Steps	*Value*	*Rel. Error*	*Flops* $\times 10^7$	*Value*	*Rel. Error*	*Flops* $\times 10^7$	*Value*	*Rel. Error*	*Flops* $\times 10^7$
10	3.9837	-0.0013	0.2790	3.9826	-0.0024	0.5310	3.9867	0.0017	1.0890
30	3.9774	-0.0076	0.8190	3.9887	0.0037	1.5510	3.9869	0.0019	3.1890
50	3.9906	0.0056	1.3590	3.9883	0.0033	2.5710	3.9845	-0.0005	5.2890
80	3.9885	0.0035	2.1690	3.9842	-0.0008	4.1010	3.9852	0.0001	8.4390

Table 4.5: Relative Error of Milstein With Different Variance Reductions, p = 30000

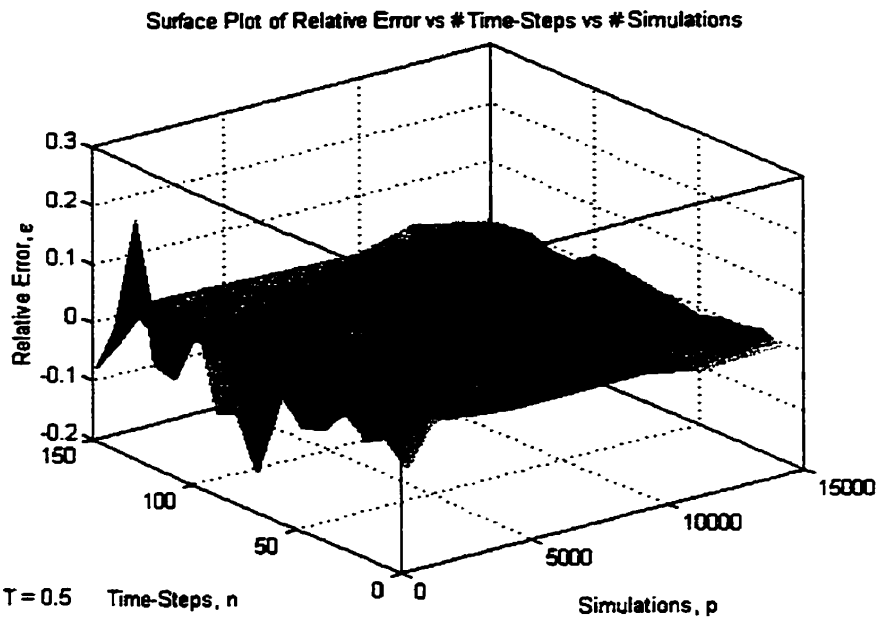

Figure 4.3: The Relative Error versus the Number of Discretizations and the Number of Simulations for Standard Euler

for Euler and Milstein with different kinds of antithetic variates. Standard Euler and standard Milstein do not perform well, displaying the erratic and oscillatory behaviour as demonstrated in Tables 4.4 and 4.5. In fact, these two schemes are furthest away from zero. Note the steep drop in relative error for standard Milstein from 1000 to 5000 simulation paths. For Euler, it moves in the opposite direction. The addition of CAV improves both standard Euler and Milstein dramatically, with both being closer to zero and showing convergence behaviour to zero. However, the effect of EAV4 for relative error is the most dramatic, with Milstein EAV4 performing the best. After 15000 simulation paths, Milstein with EAV4 is the only scheme that has fully converged

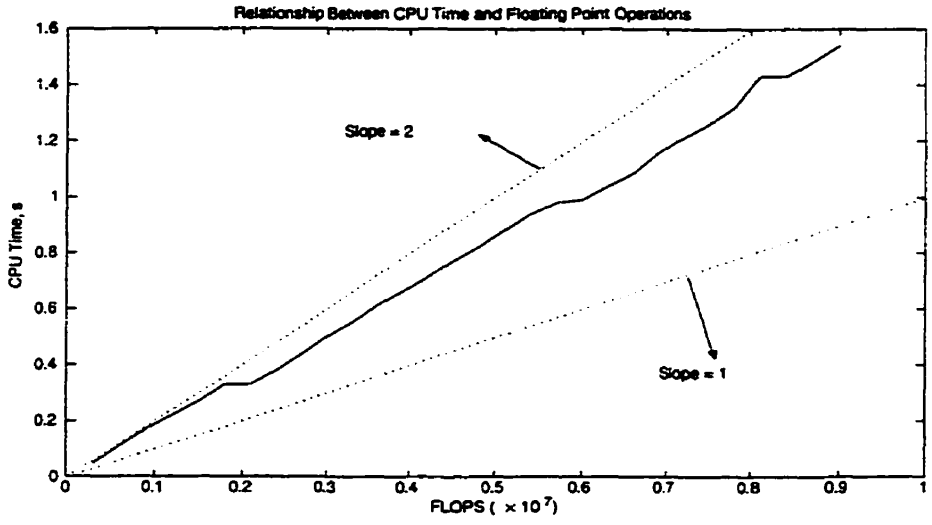

Figure 4.4: Relationship between CPU Time and Computational Costs in Flops.

as its relative error is not deviating from zero anymore. Euler with EAV4 also performs well, but not as well as Milstein with EAV4.

In section 2.4.2, we also stress the importance of accounting for the computational costs, F associated with the respective estimators as in reality, 1 simulation path for standard schemes is equivalent to 2 paths for CAV and 4 paths for EAV4. The computational cost is represented by Flops, the number of floating point operations, automatically generated by Matlab. Figure 4.4 shows the approximate relationship between Flops and CPU time. To put it into a time perspective, it is approximately 10^7 Flops per 1.7 seconds of CPU time on a Pentium III machine with 128 MB RAM.

The associated costs are shown in Figure 4.6. Comparing convergence patterns with respect to computational costs in Flops, Milstein with EAV4 dominates again.

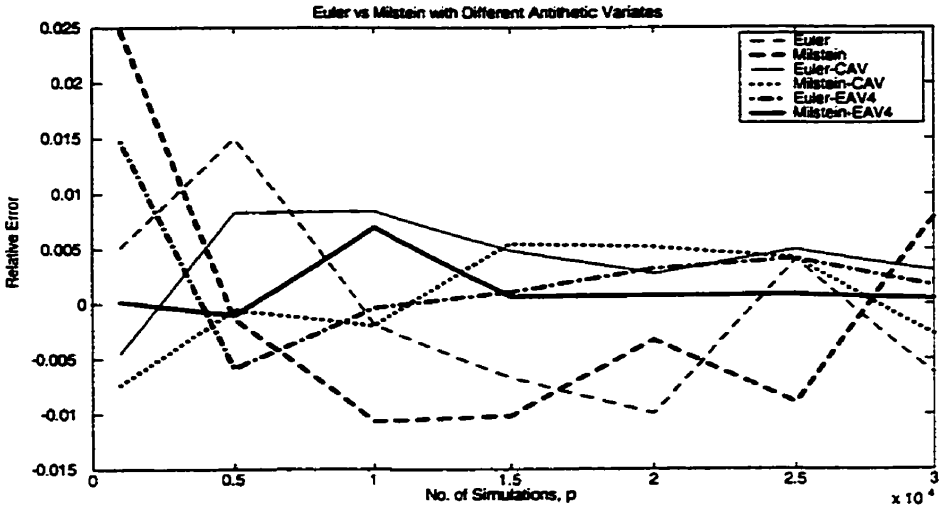

Figure 4.5: Relative Error of Euler and Milstein with Different Antithetic Variates

4.3 Variance Reduction and Numerical Efficiency

Consider the Test Case Option again. Since we are now focusing on the reduction of variance, the first thing we have to check is whether the assumptions of the propositions in Section 3.4, that is,

$$cov(g(Z_t^{ji}), g(-Z_t^{ji})) \leq 0$$

and

$$cov\left(g(Z_t^{j,i}), g((-1)^{t+1} Z_t^{j,i})\right) \leq \varepsilon,$$

hold numerically. This is reported in Table 4.6, which shows the correlations between option prices using different kinds of antithetic variates. The correlation matrix shows that the assumptions are met. Therefore, variance is reduced by using antithetic vari-

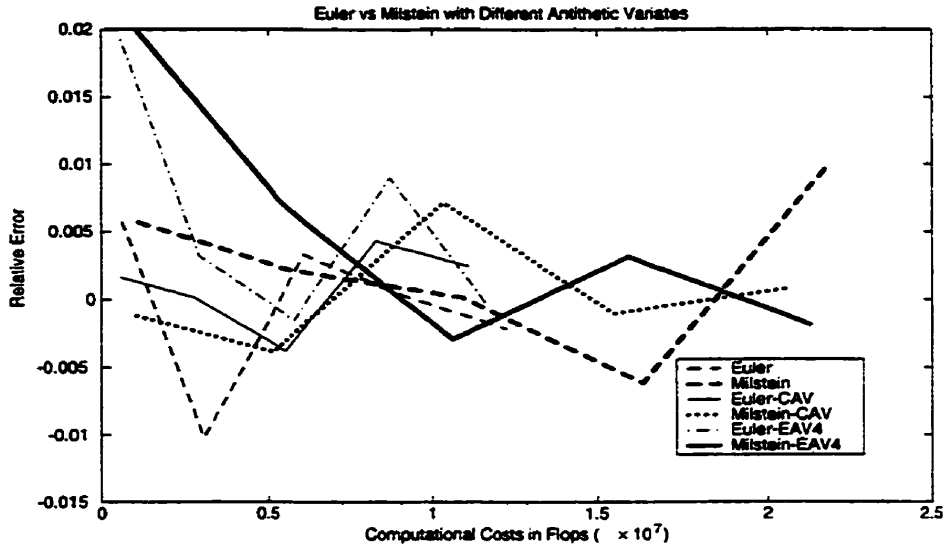

Figure 4.6: Euler and Milstein with Different Antithetic Variates for Relative Error versus Computational Cost.

is reduced by using antithetic variates, and furthermore, EAV4 reduces variance more than CAV, which in turn reduces variance more than standard schemes even after accounting for differences in computational costs.

Here the numerical efficiency of the different methods discussed in the previous sections is presented. We choose to use graphical means for this purpose. The standard error values are computed from a batch of 15. Figure 4.7 illustrates the various efficiencies of different schemes in achieving low variant option prices. The conclusions are similar to that of the previous section, in that standard Euler and Milstein are not very efficient as they require a large number of simulations to obtain a reduction in volatility. We note that Milstein falls to a much less variant level after

	$g(W)$	$g(-W)$	$g(-1^t W)$	$g((-1)^{t+1}W)$
$g(W)$	1	- 0.6285	- 0.0013	- 0.0059
$g(-W)$	-0.6285	1	0.0007	0.0010
$g(-1^t W)$	- 0.0013	0.0007	1	- 0.6277
$g((-1)^{t+1}W)$	- 0.0059	0.0010	- 0.6277	1

Table 4.6: Correlations of the Four Simulated Paths

the 20000 simulation mark. CAV schemes are more efficient compared to the standard schemes because they require only 10000 simulations to reach a plateau in the volatility level. It should be pointed out that the difference in volatility between Milstein CAV and standard Milstein from 1000 to 5000 simulations is approximately double that of Euler CAV and standard Euler, reflecting the efficiency of Milstein over Euler. Both schemes with EAV4 are the most efficient, with Milstein EAV4 only requiring 5000 simulations to attain a very stable result. Euler with EAV4 requires about 10000 simulations. Neither of the EAV4 schemes display any oscillatory behaviour. From a perspective of computational cost, the results are similar for different methods as illustrated in Figure 4.8. The conclusion is that for a low level of work, 0.5×10^7 Flops, EAV4 in both schemes reach their plateaus where any extra work beyond this point would imply a waste of resources.

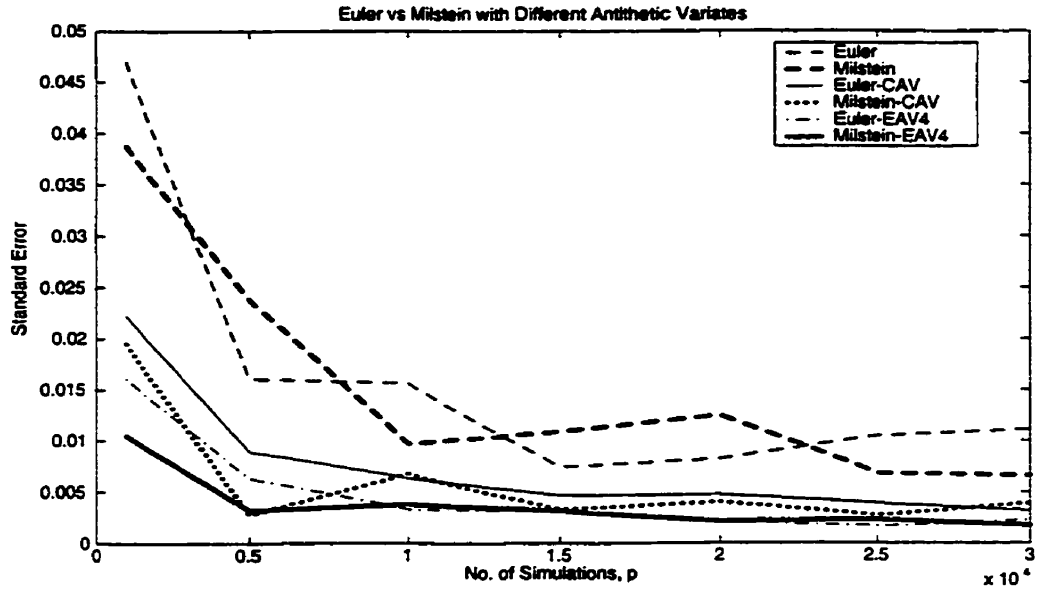

Figure 4.7: Variance of Schemes with Different Antithetic Variates Versus the Number of Simulations

Finally, we again relate variance to the number of simulations and the number of time-steps simultaneously. Figure 4.9 is a surface plot of volatility versus the number of time-steps n versus the number of simulations p. Variance is again seen to mainly depend on the number of simulations and not as much on the number of time steps. The impact of the number of simulations on the volatility of the option price is greater for increasing p, as the surface is less jagged. It has a similar surface to Figure 4.3 with the same conclusions.

Therefore, implementing Milstein with EAV4 with a reasonable number of paths is the optimal choice in that it is the most cost effective in producing minimum variance with the highest accuracy as seen in the previous section. In other words, the trade-

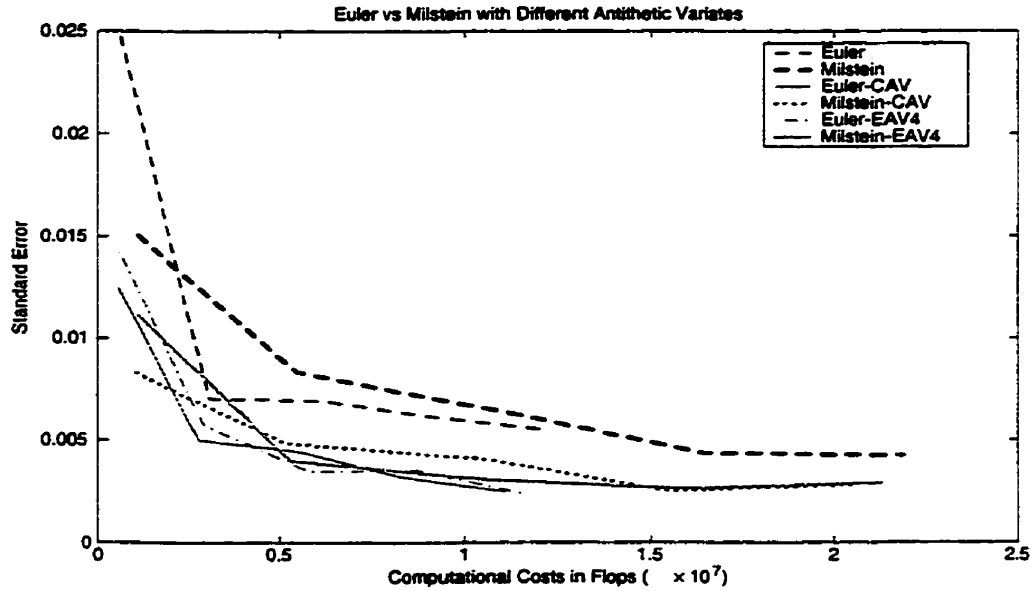

Figure 4.8: Variance of Schemes with Different Antithetic Variates versus the Computational Costs

off between variance reduction and the extra computational cost involved in using the Milstein scheme with EAV4 is positive and it would be the scheme of choice.

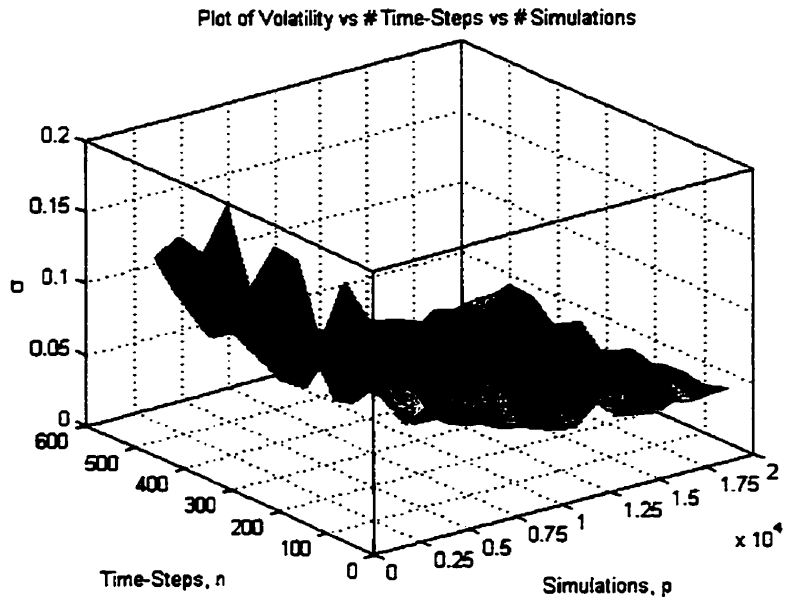

Figure 4.9: Surface Plot of Volatility versus Number of Time-Steps versus Simulation Paths

CHAPTER 5
NUMERICAL RESULTS

5.1 Multiasset Options

Here we put the previous discussions on Monte Carlo simulations and antithetic variates to practice in various option pricing problems. There will be particular emphasis on options with multiple sources of risk, that is, rainbow options. We focus on options with multiple assets (up to 100 assets) and underlying factors. Baskets or portfolio options, spreads and strikes are priced using the two discretization schemes and Monte Carlo simulations with different variance reduction techniques discussed. Numerical treatments of these kinds of options are also found in (Lari-Lavassani et al, 2000).

Several European multiassets or Rainbow call options on d securities are examined. These options have dynamics which are governed, over the horizon $[0, T]$, by

$$dS_t^k = (r - \delta^k)S_t^k dt + \sigma_k S_t^k dW_t^k, \ k = 1, ..., d \tag{5.1}$$

where $W_t^{k_2}, W_t^{k_2}$ are correlated standard Wiener processes for $k_1 \neq k_2$, $k_1, k_2 = 1, ..., d$.

The options considered do not have closed-form solutions, hence we use Monte Carlo simulations to find a price for each type of option considered. For other numerical treatments on these options, see the work of (Boyle et al, 1989), (Hua He, 1990) and (Lari-Lavassani et al, 2000).

The option prices are computed using both the Euler and Milstein schemes without and with CAV and EAV4 variance reductions. Since the system (5.1) is of the form of formula (2.29), it is commutative. Therefore when the Milstein scheme is implemented, formula (2.28) is used. All options are priced with the optimal time-step $n = 30$ and $p = 20000$ simulation runs. The numerical results are seen in Table 5.22.

5.1.1 Basket Options

The payoff for a European basket option is

$$\max\left\{\left(\sum_{k=1}^{d} v_k S_T^k\right) - K, 0\right\},$$

where K is the strike price and v_k is the weight or number of units of S_t^k in the portfolio. The behaviour of a basket option is very similar to that of a standard European call option. The payoff function is essentially that of the standard call, where the terminal asset price of one asset S_t^1 is replaced with a weighted average of the final prices, S_t^k.

Using previously defined notation, consider a 2 asset basket option with the following parameters, for $k = 1, 2$, as seen in Table 5.1.

Note that the dividend and the risk-free interest rate are continuously compounded, so one has to account for this by using

$$\text{revised parameter} = log\ (1 + \text{parameter}).$$

We price the option using Milstein with EAV4 with $n = 30$ and $p = 10000$.

Option Parameters	*Value*
S_0^k	$80 \rightarrow 120$
K	200
σ_k	0.1
δ_k	0.05
ρ	0.5
w_k	1
r	0.1

Table 5.1: Sample Option Parameters

Figure 5.1 shows the similarities between the 2 asset basket option in Table 5.1 and a standard call option. Referring back to Figure 2.2, we note that the 2 asset basket still has the same payoff structure for increasing initial values S_0, truncated portion is to the left and the option only has a value for prices greater than the exercise. This payoff structure is also evident for each instant of time. We can also see from Figure 5.1 that for increasing time to expiries, the basket option value also increases. This is explained by the fact that a longer term option allows the holder of the option the privilege of having a longer waiting time to exercise. Now, the longer the time to expiry, the greater the possibility of a higher terminal asset price while the exercise price stays constant, thereby representing a greater potential benefit to the owner. Another feature of the basket option is the smoothness of the surface, typifying a characteristic

independent option.

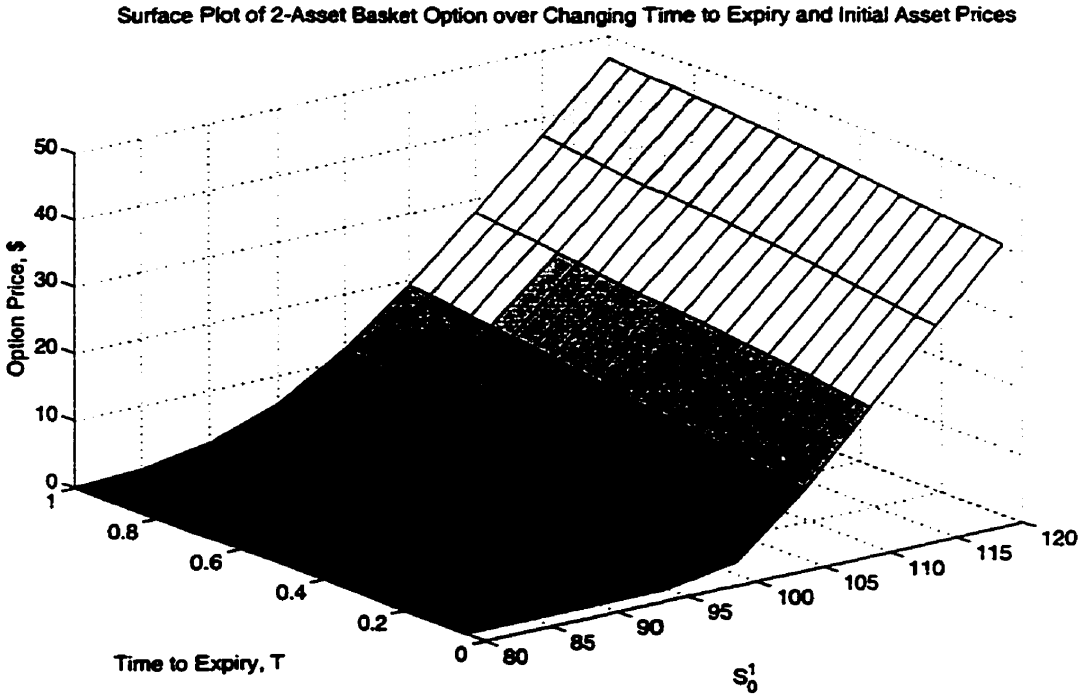

Figure 5.1: Sensitivity of 2-Asset Basket Option to Varying Expiry Times and Initial Asset Prices

In terms of the accuracy and efficiency of the various Monte Carlo schemes, we refer to Table 5.22, where it is shown that Milstein with EAV4 yields values closest to the "true" value in all cases considered. Additionally, the savings in computational cost is evident for increasingly large number of assets. This is seen in the ratio of the computational cost between Euler with no variance reduction and Milstein with EAV4. Milstein with EAV4 increases the computational cost from standard Euler by a factor of 4.74 for the 2 asset case, and only by 1.15 for the 100 asset case. This is

Option Parameters	*Asset 1*	*Asset 2*
S_0	100	100
w_k	1	1
σ_k	0.1	0.1
δ_k	0.05	0
r	0.1	
ρ	0	
T	0.5	
K	200	

Table 5.2: 2 Asset Basket Option Parameters

a clear indication that the relative efficiency of Milstein with EAV4 compared to the least expensive method, the standard Euler, increases with the number of assets.

5.1.1.1 Portfolio on Two Assets

Consider now basket options on 2 assets given by equation (5.1) with correlation ρ. Some results from (Rubinstein, 1991) and (Ware and Lari-Lavassani, 2000) are reproduced. The two assets satisfy the parameters set out in Table 5.2.

The dividend and risk-free interest rate are continuously compounded again. The true value of \$8.2615 in Table 5.22 is computed with $n = 200$ time steps and $p = 200000$ simulation runs using the Milstein scheme with EAV4. It is consistent

Absolute Relative Errors	*Euler*	*Milstein*
Std.	0.1113	0.0605
CAV	0.0130	0.0051
EAV4	0.0107	0.0006

Table 5.3: Absolute Relative Errors for 2 Asset Basket

with the results from (Rubinstein, 1991) and (Ware and Lari-Lavassani, 2000), who respectively use a quasi-binomial method and a high-order Gauss-Hermite integration method. The respective values obtained are \$8.26 and \$8.2612.

Comparing the numerical results in Table 5.22 for this portfolio on two assets, we note that for the chosen levels of n and p, Euler has not converged to the true value yet. Standard Euler performs the worst, being the furthest from the true value. Even Euler with EAV4 has not converged to the true value yet. Milstein consistently outperforms the Euler scheme for this particular option in terms of numerical accuracy. We can see this in Table 5.3, which shows the absolute value of the relative error defined in (4.1). This Milstein scheme converges, with the use of more antithetic variates, to almost the true value, while Euler is not as quick in convergence. Furthermore, there is a very dramatic decrease in the absolute relative error with the addition of antithetic variates, implying that antithetic variates are very effective for both cases. It is also evident that Milstein's numerical accuracy is about double that of Euler for the standard and CAV cases.

Ratio of Computational Costs	*Euler*	*Milstein*
Std.	1.00	1.49
CAV	1.51	2.50
EAV4	1.83	4.74

Table 5.4: Ratios of Computational Costs for 2 Asset Basket

In terms of computational costs, we set standard Euler to be the base case with a ratio of 1. The other computational costs will be measures relative to this base case and the following table has the results. Milstein is more expensive than Euler for all cases with standard Milstein being approximately the same as Euler with CAV. Milstein with EAV4 is the most expensive. The results are gathered in Table 5.4.

We conclude that for a basket on 2 assets, Milstein outperforms Euler consistently, and antithetic variates are very effective. However, the cost of Milstein is relatively high.

5.1.1.2 Portfolio On Seven Assets

We next consider a portfolio on 7 assets. This option pricing problem was originally presented by (Milevsky and Posner, 1998) and reexamined by (Ware and Lari-Lavassani, 2000). The portfolio option is embedded within an index-linked guaranteed investment certificate with the interest rate, $r = 0.063$ and the other parameters are given in Tables 5.5 and 5.6.

Index	S_0 (\$)	*Weight (%)*	*Volatility (%)*	*Dividend (%)*
TSE 100	1	10	11.55	1.69
CAC 40	1	15	20.68	2.39
DAX	1	15	14.53	1.36
MIB30	1	5	17.99	1.92
Nikkei 225	1	20	15.59	0.81
FTSE 100	1	10	14.62	3.62
S&P 500	1	25	15.68	1.66

Table 5.5: Index Linked GIC Option Pricing Parameters

	TSE 100	*CAC 40*	*DAX*	*MIB 30*	*Nikkei 225*	*FTSE 100*	*S&P 500*
TSE 100	1	0.35	0.1	0.27	0.04	0.17	0.71
CAC 40	0.35	1	0.39	0.27	0.5	-0.08	0.15
DAX	0.1	0.39	1	0.53	0.7	-0.23	0.09
MIB30	0.27	0.27	0.53	1	0.46	-0.22	0.32
Nikkei 225	0.04	0.5	0.7	0.46	1	-0.29	0.13
FTSE 100	0.17	-0.08	-0.23	-0.22	-0.29	1	-0.03
S&P 500	0.71	0.15	0.09	0.32	0.13	-0.03	1

Table 5.6: Correlations Between Indices

Absolute Relative Errors	*Euler*	*Milstein*
Std.	0.0007	0.0002
CAV	0.0003	0.0002
EAV4	0.0001	0.0000

Table 5.7: Absolute Relative Errors for 7 Asset Basket

The true value in Table 5.22 is a result of the Milstein scheme with EAV4 with $n = 200$ and $p = 200000$ and is consistent with the value of \$0.0622 produced using high-order Gauss-Hermite integration in (Ware and Lari-Lavassani, 2000). Table 5.7 compares the distances of the various estimates from the true value for this option. It is evident that all schemes converge towards the true value, with Milstein consistently outperforming the Euler schemes for all cases. In fact, Milstien with EAV4 converges exactly to the true value. The addition of antithetic variates is seen to also improve the values, with a rapid decrease in absolute relative errors with the addition of each antithetic variate method. This is more dramatic in the Euler case.

Table 5.8 compares relative computational costs with standard Euler as the base case again. Milstein is more expensive than Euler for all cases with standard Milstein being approximately the same as Euler with CAV. Milstein with EAV4 is the most expensive. Note however that the Milstein schemes are not as expensive relative to standard Euler in the 2 asset basket case. In fact, we see that Milstein with EAV4 only requires 2.67 times the computational cost of standard Euler; while for the 2 asset case,

Ratio of Computational Costs	*Euler*	*Milstein*
Std.	1.00	1.22
CAV	1.23	1.67
EAV4	1.46	2.67

Table 5.8: Relative Computational Costs for 7 Asset Basket

this value was 4.74. Therefore, with increasing number of assets for the options, the antithetic variates are more efficient as it is relatively cheap to obtain highly accurate results with respect to standard Euler.

5.1.1.3 Portfolios on Fifty and Hundred Assets

We consider next portfolios on large numbers of assets N with the following parameters set out in Table 5.9.

The true value for the 50 asset portfolio is produced from the Milstein EAV4 scheme with $n = 200$, $p = 200000$ and the true value for the 100 asset portfolio is from the Milstein EAV4 scheme with $n = 50$, $p = 800000$. Table 5.10 compares the distances of the various estimates from the true value for this option again. We note that all schemes are converging towards the true value, with different rates of convergence. We note on one hand the general trend wherein with the addition of antithetic variates, the absolute relative error is dramatically decreased, and on the other hand that Milstein outperforms Euler in all cases in terms of numerical accuracy. The best results are

Option Parameters	*50 Asset*	*100 Asset*
S_0	$70 + 0.5j, \quad j = 0, ..., 49$	$70 + 0.5j, \quad j = 0, ..., 99$
w_k	1	1
σ_k	0.2	0.1
δ_k	0.1/50	0.1/100
$\rho(j_1, j_2)$	$\frac{j_2}{50} - 0.05$ for $j_1 > j_2$	$\frac{j_2}{100} - 0.05$ for $j_1 > j_2$
r	0.1	
T	1	
K	60	

Table 5.9: Large Asset Basket Option Parameters

noted for Milstein EAV4 for both 50 and 100 asset cases.

The next table, Table 5.11, compares the relative computational costs with standard Euler as the base case. It is now very obvious that the efficiency increases with the number of assets. The computational costs for the 50 asset case are very similar, with Milstein using EAV4 only requiring 1.37 times more Flops than standard Euler. This is a trend noted from the 2 asset case, where Milstein EAV4 was 4.74 more expensive than standard Euler. This trend is also more evident for the 100 asset case where Milstein EAV4 only requires 1.15 times more Flops than standard Euler.

We therefore conclude that for options on large numbers of assets, the antithetic variates become more efficient as it becomes relatively cheaper to obtain highly accurate

	50 Asset Case		*100 Asset Case*	
Absolute Relative Errors	*Euler*	*Milstein*	*Euler*	*Milstein*
Std.	0.0943	0.0385	0.0363	0.0023
CAV	0.0127	0.0008	0.0037	0.0025
EAV4	0.0049	0.0007	0.0027	0.0002

Table 5.10: Absolute Relative Errors for Large Asset Baskets

	50 Asset Case		*100 Asset Case*	
Ratio of Computational Costs	*Euler*	*Milstein*	*Euler*	*Milstein*
Std.	1.00	1.04	1.00	1.02
CAV	1.04	1.12	1.02	1.06
EAV4	1.14	1.37	1.07	1.15

Table 5.11: Ratios of Computational Costs for Large Asset Baskets

results with respect to standard Euler. In other words, extended antithetic variates is particularly useful in cases of high-dimensions; Finally the effect of EAV4 is more prominent with the higher order scheme of Milstein.

5.1.2 Spread Options

A dual call option on a spread has a payoff of

$$\max\left\{S_T^1 - S_T^2 - K, 0\right\}.$$

Option Parameters	*Value*
S_0^k	$80 \longrightarrow 120$
w_k	1
σ_k	0.1
δ_k	0.05
r	0.1
ρ	0.5
T	0.5
K	2

Table 5.12: Dual Spread Option Parameters

Figure 5.2 shows the change in option value for a 2 asset spread option with the following parameters set out in Table 5.12.

The dividend and the risk-free interest rate are continuously compounded. It is priced using Milstein with EAV4. The increasing option value with increasing time to expiry is also evident for the spread, however, unlike the basket option, there is no truncated curve for increasing initial prices as the exercise price is only \$2. The surface is also showing a more prominent increase over time to expiry. The effect of increasing initial prices of S_0 is not as dramatic as in the basket option case. In fact, the rate of increase over varying initial prices is quite slow. Overall, the surface for the spread option corresponds to the surface area of the basket option that is at an angle to the

T versus S_0 plane, but high values of S_0 at higher time to expiries are more dramatic than the basket option.

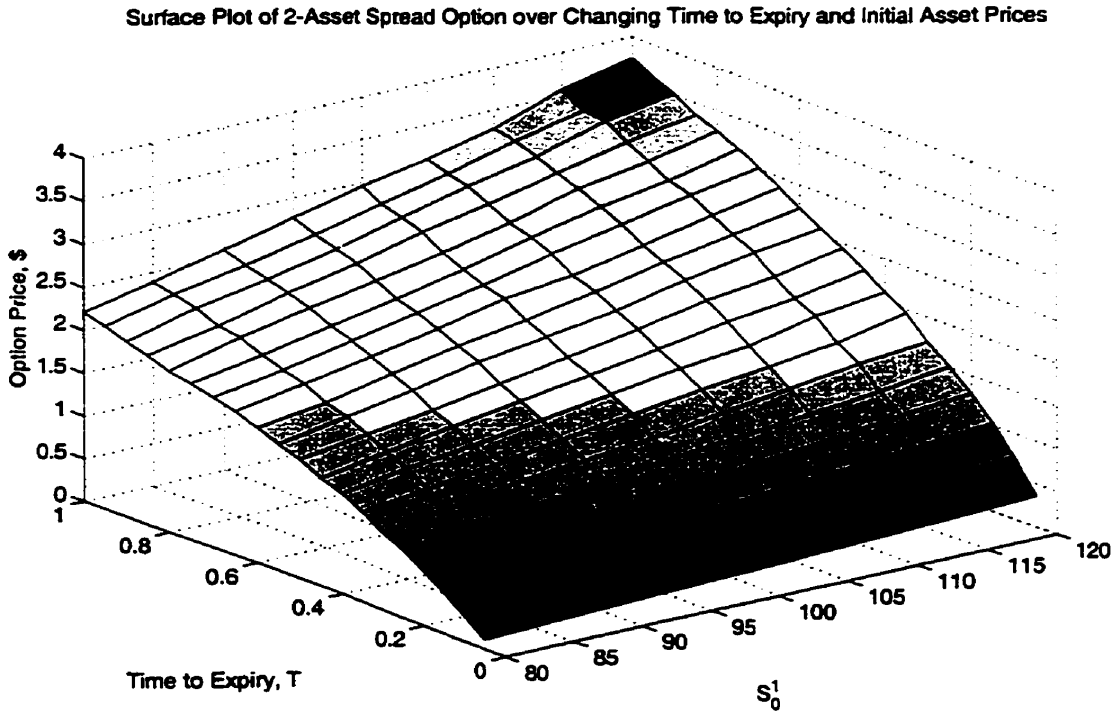

Figure 5.2: Sensitivity of 2-Asset Spread to Varying Expiry Times and Initial Asset Prices

Consider now the dual spread option prices computed using both the Euler and Milstein schemes in Table 5.22. The parameters for the dual spread option are as seen in Table 5.13.

The dividends and interest rate are continuously compounded.

From Table 5.22, the Milstein scheme generally outperforms the Euler scheme again in terms of numerical accuracy, with Milstein with EAV4 producing an option

Option Parameters	*Asset 1*	*Asset 2*
S_0	100	100
w_k	1	1
σ_k	0.1	0.1
δ_k	0.05	0.05
r	0.1	
ρ	0.5	
T	0.5	
K	2	

Table 5.13: 2 Asset Spread Option Parameters

price closest to the "true" value of \$5.2795. Note that the quasi-binomial solution (Rubinstein, 1991) for the same spread is \$5.28. In terms of the magnitude of relative errors, Milstein with EAV4 outperforms all other schemes and the ratios of computational costs yield results very similar to that of the 2 asset basket (see Tables 5.14 and 5.15).

5.1.3 Strike Options

A dual strike option has the following payoff

$$\max\left\{S_T^1 - K_1, S_T^2 - K_2, 0\right\}.$$

Absolute Relative Errors	*Euler*	*Milstein*
Std.	0.0529	0.0642
CAV	0.0369	0.0113
EAV4	0.0388	0.0020

Table 5.14: Absolute Relative Errors for 2 Asset Spread

Ratio of Computational Costs	*Euler*	*Milstein*
Std.	1.00	1.50
CAV	1.51	2.49
EAV4	2.76	4.75

Table 5.15: Ratios of Computational Costs for 2 Asset Spread

Figure 5.2 shows the change in option value for a dual strike option with the following parameters in Table 5.16.

The dividend and the risk-free interest rate are continuously compounded. The strike option is priced using Milstein with EAV4. The behaviour of the surface is very similar to that of a 2 asset basket option, however, the surface is not as uniform for the region from $S_0^1 = 100$ to \$120 as there is a bigger dip in option value at low time to expiry.

The parameters for the dual strike option are as set out in Table 5.17 where the dividends and interest rate are continuously compounded. The quasi-binomial solution from (Rubinstein, 1991) for this strike is \$17.70, while the "true" value generated by

Option Parameters	*Value*
S_0^k	$80 \rightarrow 120$
w_k	1
σ_k	0.1
δ_k	0.05
r	0.1
ρ	0.5
T	0.5
K	100

Table 5.16: Dual Strike Option Parameters

Milstein with EAV4 is \$17.6891. The same conclusion is made for the result as in the spread option case, in that, Milstein outperforms Euler and the addition of antithetic variates improves the results, especially the addition of EAV4. We also note that for 2 assets, the cost of EAV4 is high. The convergence pattern and relative computational costs also yield similar results to that of the 2 asset basket and spread as demonstrated in Tables 5.18 and 5.19.

Option Parameters	*Asset 1*	*Asset 2*
S_0	100	100
w_k	1	1
σ_k	0.1	0.1
δ_k	0.05	0.05
r	0.1	
ρ	0.5	
T	0.5	
K	90	

Table 5.17: 2 Asset Strike Option Parameters

Absolute Relative Errors	*Euler*	*Milstein*
Std.	0.1577	0.0033
CAV	0.0746	0.0187
EAV4	0.0047	0.0044

Table 5.18: Absolute Relative Errors for 2 Asset Strike

Ratio of Computational Costs	*Euler*	*Milstein*
Std.	1.00	1.50
CAV	1.51	2.49
EAV4	2.76	4.75

Table 5.19: Ratios of Computational Costs for 2 Asset Strike

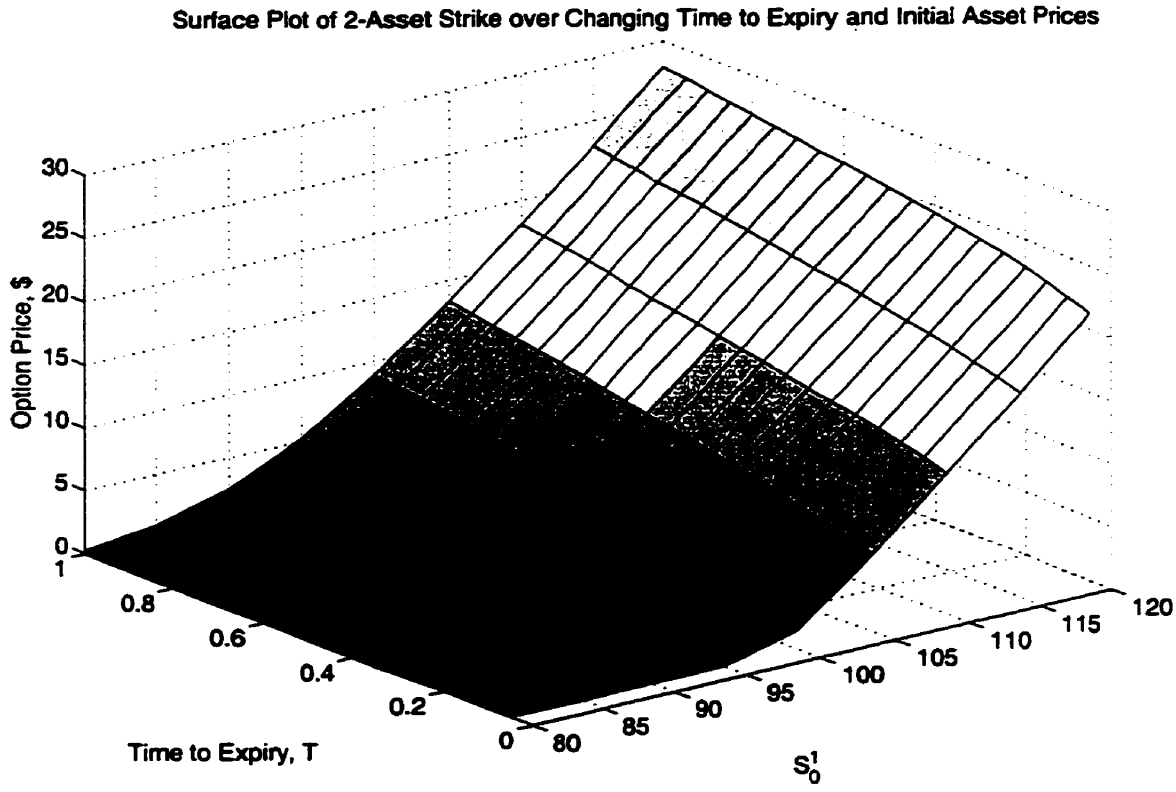

Figure 5.3: Sensitivity of Dual Strike for Varying Expiry Times and Initial Asset Prices

5.2 Multifactor Models

5.2.1 Stochastic Volatility

Consider here the system (2.33), where the processes are uncorrelated, i.e., $\rho = 0$. Note that in this case the techniques developed for the estimation of the double Ito integral must be used as it was shown that the system does not commute. A European put option is priced with coefficients set out in Table 5.20.

In order to avoid negative values of σ so that the system stays in the real plane,

Option Parameters	*Values*
α	4
β	0.09
γ	0.4
$r = \mu$	0
T	0.5
σ_0	0.09
S_0	80
K	100

Table 5.20: Parameters for Stochastic Volatility

one has to choose a relatively large number of time discretizations. Therefore, the interval [0,0.5] is divided into $n = 200$ time steps. We choose to compare their resulting accuracies and variances for the same computational cost so that the efficiency and effectiveness of the various methods are more obvious. This is achieved by maintaining the corresponding Flops in the same range for each case.

The results are gathered in Table 5.21, where the Milstein scheme implements formula (2.22) with (2.34) and (2.36) in computation of the double Ito integral, I_{j_1,j_2}. The quoted standard deviations are obtained from a batch of 50 experiments.

This option pricing case is also treated in (Clarke and Parrott, 1999), using a multigrid finite difference approach, leading to the option price of \$21.417. The

analytic solution, derived from a Power series, is known from (Ball and Roma, 1994) to be \$21.430. We note first that Milstein with EAV4 using the formula (2.36) clearly outperforms the other methods. It consistently yields an option value in the range of the analytic solution of \$21.43. The results of Milstein using formula (2.34) are not as good, similarly for Euler. We also note that with the addition of antithetic variates, the volatility of the estimates drops dramatically, with the best variance result for EAV4 in all schemes. Since the computational costs are all in the same range, we can say that in general, Milstein is more efficient and accurate compared to Euler. Furthermore, Milstein with formula (2.36) is also superior to Milstein using (2.34) in these two aspects.

Finally, we can conclude that antithetic variates help in reducing volatility and improving convergence, additionally, EAV4 is superior to CAV which is superior to no antithetic variates. In addition, as in the multiasset case, for options with multifactor models, the higher order scheme of Milstein is superior to Euler; and EAV4, used in conjunction with Milstein, is very effective.

Scheme	*Value*	*SD*	*Flops* ($\times 10^7$)	p	k or ς
Finite Difference	21.4170				
Power Series	21.4300				
Euler					
Std.	21.4160	0.0921	12.2	32000	
CAV	21.4345	0.0193	14.7	21000	
EAV4	21.4277	0.0195	14.7	11000	
Milstein (2.34)					
Std.	21.4635	0.2393	12.5	3000	9
CAV	21.4393	0.0540	14.9	3000	9
EAV4	21.4430	0.0488	14.6	2000	5
Milstein (2.36)					
Std.	21.4389	0.3039	12.1	3000	20
CAV	21.4337	0.0533	14.8	3000	20
EAV4	21.4303	0.0415	14.8	3000	10

Table 5.21: Option Price for Uncorrelated Stochastic Volatility Model

		$n = 30,\ p = 20000$					
	"True"	Euler			Milstein		
	Value	Std.	CAV	EAV4	Std.	CAV	EAV4
BASKET							
2-Asset	8.2615	8.1502	8.2745	8.2508	8.2010	8.2564	8.2621
Flops $\times 10^7$	304.5001	0.972	1.466	2.69	1.4520	2.426	4.61
7-Asset	0.0622	0.0615	0.0625	0.0621	0.0624	0.0620	0.0622
Flops $\times 10^7$	1345.3	7.592	9.306	13.57	9.272	12.66	20.29
50-Asset	27.7828	27.8771	27.7955	27.7877	27.8213	27.7836	27.7821
Flops $\times 10^7$	2014.1	312.2083	324.4143	354.8223	324.2083	348.4143	428.223
100-Asset	40.351	40.3147	40.3473	40.3483	40.3280	40.3485	40.3512
Flops $\times 10^7$	92816.4	1224.4	1248.8	1309.7	1248.4	1296.8	1405.7
SPREAD							
2-Asset	5.2795	5.3324	5.2426	5.3183	5.2153	5.2682	5.2775
Flops $\times 10^7$	304.2601	0.966	1.454	2.666	1.446	2.406	4.586
STRIKE							
2-Asset	17.6891	17.5314	17.6145	17.6938	17.6924	17.7078	17.6847
Flops $\times 10^7$	304.2601	0.966	1.454	2.666	1.446	2.406	4.586

Table 5.22: Numerical Results for Various Multi-Asset Options.

Bibliography

[1] Ball, C.A., and A. Roma, 1994. "Stochastic Volatility Option Pricing", Journal of Financial and Quantitative Analysis, 29, pp. 589-607.

[2] Barlow, R.E, and F. Proschan, 1975. Statistical Theory of Reliability and Life Testing, Holt, Rinehart and Winston, Inc., New York

[3] Barraquand, J., 1995. "Numerical Valuation of High Dimensional Multivariate European Securities", Management Science, 41:12 (December), pp. 1882-91.

[4] Bartoszyński, R., and M. Niewiadomska-Bugaj, 1996. Probability and Statistical Inference, Wiley Series in Probability and Statistics, Wiley-Interscience, New York.

[5] Bhansali, V., 1998. Pricing and Managing Exotic and Hybrid Options, McGraw-Hill, New York.

[6] Bingham, N. H., and R. Kiesel, 1998. Risk Neutral Valuation, Springer, New York.

[7] Black, F., and M. Scholes, 1973. "The Pricing of Options and Corporate Liabilities", Journal of Finance, 31, 899-908.

[8] Boyle, P. P., 1977. "Options: A Monte Carlo Approach", Journal of Financial Economics, 4, pp. 323-38.

[9] Boyle, P.P., M. Broadie, and P. Glasserman, 1997. "Monte Carlo Methods for Security Pricing", Journal of Economic Dynamics and Control, 21 (8/9), pp. 1267-321.

[10] Boyle, P.P., J. Evnine, and S. Gibbs, 1989. "Numerical Valuation of Multivariate Contingent Claims", Review of Financial Studies, 2:2, pp. 241-50.

[11] Bratley, P., B.L. Fox and L. Schrage, 1987. A Guide to Simulation, 2nd edition, Springer-Verlag, New York.

[12] Broadie, M., and P. Glasserman, 1996. "Estimating Security Price Derivatives by Simulation", Management Science, 42, pp. 269-85.

[13] Clarke, N., and K. Parrott, 1999. "Multigrid for American Option Pricing with Stochastic Volatility", Preprint, 17 pages.

[14] Clewlow, L., and A. Carverhill, 1994. "On the Simulation of Contingent Claims", Journal of Derivatives, 2, pp. 66-74.

[15] Clewlow, L., and C. Strickland, 1999. Implementing Derivatives Models, John Wiley and Sons, New York.

[16] Conte, S.D, and C. de Boor, 1972. Elementary Numerical Analysis, McGraw-Hill, Tokyo.

[17] Cox, J. C., and Ross, S. A., 1976. "A Survey of Some New Results in Financial Option Pricing Theory", The Journal of Finance, 31:2 (May), pp. 383-402.

[18] Glynn, P. W., and W. Whitt, 1992. "Indirect Estimation via $L = \lambda W$", Operations Research, 37, pp. 82-103.

[19] Hammersley, J.M., and D.C. Handscomb, 1964. Monte Carlo Methods, Chapman & Hall, London.

[20] Harrison, J. M., and S. Pliska, 1981. "Martingales and Stochastic Integrals in the Theory of Continuous Trading", Stochastic Processes and Their Applications, 11, pp. 215-260.

[21] Hua He, 1990. "Convergence from Discrete-to Continuous Time Contingent Claims Prices", Review of Financial Studies, 3:4, pp. 523-46.

[22] Hull, J. and A. White, 1988. "The Use of the Control Variate Technique in Option Pricing", Journal of Financial and Quantitative Analysis, 23:3, September, pp. 237-251.

[23] Jarrow, R. A., 1994. "Derivative Security Markets, Market Manipulation, and Option Pricing Theory", Journal of Financial and Quantitative Analysis, June, pp. 241-261.

[24] Jarrow, R. A., and S. Turnbull, 1996. Derivative Securities, International Thompson Publishing, Cincinnati.

[25] Joy, C. P. Boyle, and K. S. Tan, 1996. Quasi-Monte Carlo Methods in Numerical Finance, Management Science, 42 (June), pp. 926-938.

[26] Kalos, M. H., and P. A. Whitlock, 1986. Monte Carlo Methods, Volume 1, Wiley-Interscience, New York.

[27] Kloeden, P.E., and E. Platen, 1999. Numerical Solution of Stochastic Differential Equations, 3rd edition., Springer-Verlag, New York.

[28] Kolb, R. W., 1997. Future, Options and Swaps, 2nd Edition, Blackwell Publishers, Oxford.

[29] Lamberton, D., and B. Lapeyre, 1997. Introduction to Stochastic Calculus Applied to Finance, 2nd edition, Chapman and Hall, New York.

[30] Lari-Lavassani, A., A. Sadeghi, and H. Wong, 2000. " Monte Carlo Simulations and Option Pricing", Preprint, 30 pages.

[31] Lehmann, E.L., and G. Casella, 1998. Theory of Point Estimation, 2nd Edition, Springer, New York.

[32] Lemieux, C. and P. L'Ecuyer, "Using Lattice Rules as a Variance Reduction Technique in Simulation", to appear in the Proceedings of the 2000 Winter Simulation Conference, December.

[33] Milevsky, M.A., and S.E. Posner, 1998. "A Close-form Approximation for Valuing Basket Options", Journal of Derivatives, 5:4, pp. 54-61.

[34] Milstein, G. N., 1974. "Approximate Integration of Stochastic Differential Equations", Theory Probab. Appl., 19, pp. 557-62.

[35] Morgan, B. J. T., 1984. Elements of Simulation, Chapman and Hall, New York.

[36] Morokoff, W. J., and R. E. Caflisch, 1998. "Quasi-Monte Carlo Simulation of Random Walks in Finance", Monte Carlo and Quasi-Monte Carlo Methods 1996, Lecture Notes in Statistics 127, Springer-Verlag, New York, pp. 340-352.

[37] Ninomiya, S. and S. Tezuka, 1996. "Toward Real Time Pricing of Complex Financial Derivatives", Applied Mathematical Finance, 3, pp. 1-20.

[38] Øksendal, 1995. Stochastic Differential Equations, 4th Edtion, Springer-Verlag, New York.

[39] Papageorgiou, A., and J. F. Traub, 1996. "Beating Monte Carlo", RISK, 9 (June), pp. 63-65.

[40] Paskov, S. H., 1997. "New Methodologies for Valuing Derivatives", Mathematics of Derivative Securities, edited by Dempster, M. A. H, and S. Pliska, Isaac Newton Institute, Cambridge University Press, Cambridge, pp. 545-82.

[41] Paskov, S. H., and J. F. Traub, 1995. "Faster Valuation of Financial Derivatives", Journal of Portfolio Management, 22:1 (Fall), pp. 113-120.

[42] Rubinstein, M., 1991. "European Rainbow Options", Exotic Options (7/11), Manuscript, 14 pages.

[43] Schuss, Z., 1980. Theory and Applications of Stochastic Differential Equations,Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, New York.

[44] Taleb, N., 1997. Dynamic Hedging, John Wiley and Sons, New York.

[45] Turnbull, S. M., and L. M. Wakeman, 1991. "A Quick Algorithm for Pricing European Average Options", Journal of Financial and Quantitative Analysis, September, pp. 377-89.

[46] Various authors, 1998. Monte Carlo: Methodologies and Applications for Pricing and Risk Management, RISK Books, Risk Publications, London.

[47] Ware, A., and A. Lari-Lavassani, 2000. "Algorithms for Portfolio Options", Preprint, 15 pages.

[48] Wilmott, P., 1998. Derivatives: The Theory and Practice of Financial Engineering. John Wiley and Sons, New York.