KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

OPTION PRICING : A PARTICLE FILTERING APPROACH

By

Henry Nii Ayitey-Adjin

(Bsc. Mathematics)

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS, KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY IN PARTIAL FUFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF MSc. INDUSTRIAL MATHEMATICS

October 15, 2015

Declaration

I hereby declare that this submission is my own work towards the award of the MSc. degree and that, to the best of my knowledge, it contains no material previously published by another person nor material which had been accepted for the award of any other degree of the university, except where due acknowledgement had been made in the text.

Henry Nii Ayitey-Adjin	....	....
Student	Signature	Date

Certified by:

E. Owusu-Ansah	....	....
Supervisor	Signature	Date

Certified by:

Prof. S.K. Amponsah	....	....
Head of Department	Signature	Date

Dedication

With love and thankful heart to God, family. I am becoming a better person each day.

Abstract

Option pricing is a critical issue in the financial market. An investigation into the use of Sampling Importance Resampling (SIR) filter for financial option pricing in the Black-Schole model is performed. The impact of process noise, measurement noise, and the number of particles on the accuracy and performance of SIR filter is examined. The Black-Schole model is solved by the finite difference scheme. The SIR filter is implemented by the use of the GARCH model and the Black-Schole model with synthetic data. The effect of different process noise, measurement noise, and number of particles on the SIR filter was examined. It was found that the SIR filter performed well at lower process noise and high measurement noise when considering profitability of a call option. Also, as the number of particle decrease the SIR filter performed very well.

Acknowlegdment

Thanks be to God, who gives us the victory through our LORD JESUS CHRIST.

Firstly, I thank God for all the wisdom to come this far. I am very grateful to Mr. Owusu Ansah of the department of Mathematics, KNUST, my supervisor, for the faith in me, the guidance he provided and putting me to this challenge. God richly bless you. To all the lecturers at the department of mathematics, I am very grateful. To my parent and family, thank you all for the support and encouragement. God richly bless you all.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In the financial world, option is basically a contract, which does not oblige but give the right to an investor to either buy or sell a financial asset often called underlying asset. Options have undoubtedly become a major part of financial market, with its ability to cover the risk to certain extent coupled with it high degree of complexity. The thesis focused on American-styled options that permit an exercise at any time from the inception date to the expiration date. The BlackScholes Model(1973) describing a mathematical framework on option pricing is adapted in this thesis.

An investigation into particle filtering as a technique to pricing American-styled options is explored. Obviously over the years, a number of different ways have been used in the pricing of financial instrument.

1.1 Background of the study

Data assimilation schemes are designed to utilize measured observations in conjunction with the dynamic system, with estimates of the uncertainty in the estimated states. The two types of data assimilation schemes are the sequential and variational assimilation. Variational data assimilation use all the observation available over a given period of time to give improved estimates for all the states in that time period. It is based on optimal control theory. Sequential data assimilation uses a probabilistic framework and given estimates of the whole system state sequentially by propagating information only forward in time. Therefore avoid deriving an inverse model and make sequential method easier to adopt for all models according to Bertino et al. (2003)

Data assimilation basically quantify error found in both the model predictions and observations. These errors may be caused by a number of factors. For example, violation in the assumption of the model, the use of incorrect parameter values that are not optimal can result in model error. The continuous dynamic systems are solved numerically and so are transformed into discrete dynamical systems. Computations resulting from this often bring round-off errors in the model predictions. During reading of data and inaccurate instrument often introduce human error into the observation. Finite difference schemes is used in providing a numerical solution to the underlying models of the systems. Data assimilation does not only quantify errors but also reduce errors and to provide a more accurate predictions of both states and parameters.

1.2 Statement of the Problem

Sampling Importance Resampling (SIR) filter is a Monte Carlo (MC) method for implementing a recursive Bayesian filter by representing the required posterior density function by a set of random samples with associated weights and to compute estimates based on these samples and weights. according to Dablemont et al. (2009)

This study investigates the performance of sampling importance resampling filter first proposed by Gordon et al. (1993) in estimating Black-Schole model when the stock price and time to maturity are varied whiles keeping other parameters constant, effect of different measurement and process noise and finally the effect of different number of particles. The performance of the SIR is evaluated by the use of synthetic data.

1.3 Objectives of the Study

In this study, the objective is to examine the performance of the Sampling Importance Resampling (SIR) filter in pricing options. SIR filter's performance is evaluated through experimentation to comparing the effect of different: • stock price and time to maturity on mean volatility

- process noise on the filter.
- measurement noise on the filter.
- number of particles of SIR filter.

1.4 Methodology

The performance and accuracy of the SIR filter is examined based on the impact of the process noise, measurement noise and the number of particles. The statespace estimation problem was investigated by the use of the SIR.
Finite difference scheme was used to solve the Black-Scholes model. A detailed algorithm of the SIR filter was developed and implemented using the GARCH model as state equation and the Black-Scholes model as observation equation. The algorithms used in the study were implemented in Matlab and used in performing a number of experiments with synthetic data.

1.5 Significance of the study

Options provide investors ie individuals or institutions with great leveraging power, lesser risk, higher potential return and limit losses. What is the best price to buy or sell options ie pricing of financial options and how to determine this price is a relevant question to answer. Surely, a number of data assimilation schemes have been applied in determining the price of options, for example Jasra and Del Moral (2010) and Lindstrom and Guo (2013). What has not been done is the use of the GARCH model with the SIR filter. This thesis investigate the application SIR filter with the GARCH model.

1.6 Organization of the study

The study is outlined in five chapters. In Chapter 1, an introduction to research is presented. Chapter 2 contains literature review and a general framework of the study. The methodology employed in this study is discussed in Chapter 3. In Chapter 4, there is the discussion of result and findings from estimation problem of pricing financial instruments. Finally Chapter 5 presents the summary of findings, conclusion and recommendations.

Chapter 2

Literature Review

2.1 Introduction

An option is a financial product usually called instrument whose value is obtained from the value of another asset hence it is called a ’derivative’. The method for pricing such instruments was more or less based on guesswork until 1973. In 1973, Black and Scholes (1973) and Merton (1973) published their work on options. Since then, option pricing has been transformed into science by the Black-Scholes equation. This chapter deals with review of literature on pricing of options and the use of data assimilation methods in option pricing.

2.2 Options

In the financial world, an option is basically a contract, which does not oblige but give the right to an investor to either buy or sell a financial asset often called underlying asset S_0. Options are bought at specific price known as strike price K and can be exercised or acted on, before or on expiration date (T). The sellers of these options also known as the writer, incur the obligation to buy or sell the underlying if the investor choose to exercise his right. Investors pay a premium for this right to writer. Options can either be a call option or put option. Call option give the investor the right to buy an underlying at a specific exercise price. A put option gives the investor the right to sell an underlying at a specific strike price.

Options that are exercised at any time up to its expiration date are known as American Options whiles options that can only be exercised at its expiration date is referred to as European options, Hull (2006).

For American options to be profitable, the underlying asset's price should be greater than the strike price in the case of a call option and in the case of put option the strike price should be greater than the underlying asset's price.

If we suppose the price of the underlying asset at time t is a random variable $S_t := S(t)$ with a strike price K,then the payoff from a call option at time of maturity is

$$V_c(S,T) = max\{S - K,0\}. \tag{2.1}$$

Also for a put option at time of maturity is

$$V_p(S,T) = max\{K - S,0\}. \tag{2.2}$$

2.3 Valuation of Financial Options

Financial options are widely trade assets in the financial market thus there is the need for a structured method for determining it price. A simple way to determine the value of an option is whether or not it will likely be in-the-money or out-themoney at expiration date. For a call option, it is in-the-money if $S > K$ and out-the-money, if $S < K$. A put option is in-the-money, if $S < K$ and out-themoney, if $S > K$. The value of an option is known as the premium. The premium of an option is price paid by the buyer and amount received by the seller. The value of an option (Premium) can be broken down into two simple parts: - Intrinsic value: This is the difference between the price of the underlying asset and strike price.

$$Intrinsic \quad value = S - K \quad (Call \quad Option) \tag{2.3}$$

$$= K - S \quad (Put \quad Option) \tag{2.4}$$

- Time value: This is the price paid for an option greater than its intrinsic value with a belief that before the expiration date the value of the option will increase due to favourable changes in the underlying price. For a greater time value, the option must spend longer time in the market.

$$Time \quad value = Option \quad Premium - Intrinsic \quad V alue \tag{2.5}$$

A number of factor influence the value of an option. In Hull (2006), six major factors affect options: the price of underlying asset (S_0), the strike price K, the time of expiration T, the volatility of the price (σ), the risk free interest rate r, and the dividends expected during the life of the option.

Table 2.1: Determinants of Option value

Factors	Call Value	Put Value
Increase in stock price	Increase	Decrease
Increase in strike price	Decrease	Increase
Increase in expiration time	Increase	Increase
Increase in volatility	Increase	Increase
Increase in interest rate	Increase	Decrease
Increase in dividends	Decrease	Increase

For greater accuracy and consistency, mathematical models are employed. The most famous of these models is the Black-Schole's model (Black and Scholes (1973)) used for the pricing of European put and call option.

A theoretical pricing formula for pricing option was derived by Black and Scholes (1973). The underlying principles of the formula are: if options are correctly priced in the market, it should not be possible to make sure profit by creating portfolios of long and short positions in options and their underlying assets. This model is useful to corporate liabilities such as traded stocks, bonds, commodities and index.

Rubinstein (1983) worked on a option pricing formula that places the main source of risk on the risk of individual underlying assets. In relation to the Black-Scholes equation, the displaced diffusion formula has several desirable features. The equation encompasses differential riskiness of the assets, their relative weights in price determination of the firm, the effect of firm debt and the effect of a dividend payment policy with constant and random components.

In 1992, Gallant et al. (1992), worked on the joint dynamics of price changes and volume on the stock market making use of daily data on S&P composite index and total NYSE trading volume from 1928 to 1987. Nonparametric technique was used in achieving the set goals. Gallant et al. (1992) discovered that there was a positive and nonlinear relationship between daily trading volume and the magnitude of the daily price change and that price change leads to volume movements.

Heston (1993) developed a new technique, based on the Black-Scholes equation, to derive a closed-form solution to the pricing of an European call option on an asset with stochastic volatility. The model allows arbitrary correlation between volatility and spot assets returns. With the introduction of stochastic interest rate, Heston (1993) demonstrated how the model can be applied to bond options and foreign currency options. Result from Heston (1993) showed that correlation between volatility and the spot asset price is important in tell the story of return skewness and strike price biases in the Black-Scholes model (Black and Scholes (1973)). Hull and White (1987), Stein and Stein (1991) have also contributed to the literature on stochastic volatility option pricing.

Pastorello et al. (2000) worked on the estimation of continuous-time stochastic volatility models for pricing options.They developed a Monte Carlo experiment which compared two strategies based on different information sets.Their basic assumptions were:An Ornstein-Uhlenbeck process for log of the volatility, a zerovolatility risk premium and no leverage effect. In their work, they kept to the framework with no over-identifying restrictions, which led to showing that estimation based on option prices were far more precise in samples of typical size for a given option pricing model.

In Harrison and Pliska (1981), the value of American-styled option, is found, guided by the fundamental theorem of no arbitrage pricing. Also the solution to the European option pricing problem in a non-arbitrage, constant volatility framework is provide in the work of 4. A number of references on American option pricing include Brennan and Schwartz (1977), Broadie and Glasserman (1997), Detemple and Tain (2002), Geske and Johnson (1984).

2.4 Data Assimilation

The application of data assimilation which is typically referred to the estimation of the state of a physical system given a model and observation, is specifically applied to option pricing. In Lahoz et al. (2010), the aim of a data assimilation scheme is to use measured observations in combination with a dynamical system model in order to derive accurate estimates of the current and future states of the system, together with estimates of the uncertainty in the estimated states. Data assimilation is interested in the flow and prediction of the state of processes where in most cases are time dependent. A model is useful and necessary to express the temporal changes of the process. Models that are time-dependent of this kind are called dynamical system or dynamical model. In Weisstein (2002), " a means of describing how one state develops into another state over the course of time" is what defines a dynamical system. Thus a mathematical formulation of one or more factors assumed to influence the dynamics of a process is what we call a dynamical system. There are two kinds of dynamical systems: deterministic system and stochastic system. In a deterministic system, given an initial condition, the evolution of the system is completely expressed as a rule relating one state to the future state. Most deterministic system make a number of assumptions about a process which make them incomplete. To account for these assumptions, a stochastic term often referred to as system noise or model error is added to the deterministic system. Thus making such systems stochastic in nature. The focus of this thesis is stochastic systems. The combination of dynamic model and observation to obtain improved estimates is called data assimilation. Most data assimilation schemes are developed for more accurate estimate of the current and future state of dynamic system by use of measured observation and dynamic models. (Kalman (1960); Evensen (2003); Ott et al. (2004)). An analytical technique where observed data is accumulated into the model state by taking advantage of consistency constraints with laws of time evolution and physical properties defines data assimilation by (Bouttier and Courtier (2009)). Errors in models often come from inaccurate parameters in the dynamic model. Data assimilation is often used to estimate the parameters. There are two approaches to data assimilation. These are variational and sequential data assimilation. Variational data assimilation makes use of observation from the future in instances of reanalysis and observation are processed in small batches,Bouttier and Courtier (2009). The focus of this thesis was sequential data assimilation. More specifically the thesis focused on sequential data assimilation to stochastic system.

2.5 Sequential Data Assimilation

In sequential data assimilation, observation are fed back into the model at each time these are available and a best estimate is produced and used to predict future states. To describe the sequential data assimilation technique, we assume a perfect dynamic system modeled by the equation

$$x_k = f_k(x_{k-1}, u_{k-1}, v_k; w) \tag{2.6}$$

with the observational equation is given by:

$$y_k = h_k(x_k, n_k; w) \tag{2.7}$$

where:

x_k:the state vector at the time k, y_k:the measurement vector,

u_k:an external input of the system, assumed known, v_k:the process noise that drives the dynamic system, n_k:the measurement noise corrupting the observation of the state, f_k:a time-variant, linear or non-linear function, h_k:a time-variant, linear or non-linear function, w: the parameters vector.

Assuming that at time t_k, prior background estimate x_k^c for the various states are known. The differences between the observations of the true states and the observations predicted by the background states at this time ($y_k - h_k(x^c{}_k)$), are then used to make a correction to the background state vector in order to obtain improved estimates $x^a{}_k$ referred to as the analysis states. The model is then evolved forward from analysis states to the next time t_{k+1} where observations are available. The evolved states of the system at time t_{k+1} become the background states and are denoted by $x^c{}_{k+1}$. The background is then corrected to obtain an analysis at this time and the process is repeated.

According to Barillec (2009), if the model and the observation operator are linear, and if all distributions are Gaussian, then the Kalman filter (Kalman (1960)) provides an optimal (variance minimising) solution to the filtering problem. If the operators are non-linear, sub-optimal methods can be derived. The Extended Kalman Filter (Jazwinski (1970); Maybeck (1979)) and the Ensemble Kalman Filter (Evensen (1994)) provide respectively a linearised and a Monte Carlo approximations to the Kalman Filter. Another Monte-Carlo approach, the Particle Filter (Doucet et al. (2001)), allows the Gaussian assumption to be relaxed. The Monte Carlo methods is a kind to stochastic sampling approach aiming to tackle the complex systems which are analytically intractable. This Monte Carlo methods are so powerful that they are able to attack the difficult numerical integration problems. Examples of these Monte Carlo methods include Bayesian Bootstrap,

Hybrid Monte Carlo, Quasi Monte Carlo.

2.6 Particle Filtering

Particle filtering as a sequential Monte Carlo method is explored. The sequential Monte Carlo methods have become attractive due to the fact that they allow online estimation by combining the powerful Monte Carlo sampling methods with Bayesian inference, at an expense of reasonable computational cost. Particularly, the sequential Monte Carlo methods has been used in parameter estimation and state estimation where the latter is referred to as particle filter. The basic idea of particle filter is to use a number of independent random variables called particles, sampled directly from the state space, to represent the posterior probability, and update the posterior by involving the new observations; the particle system is properly located, weighted and propagated recursively according to the Bayesian rule. The earliest idea of Monte Carlo method used in statistical inference is found in Handschin (1970) and Akashi and Kumamoto (1975), but the formal establishment of particle filter seems fair to be due to Gordon et al. (1993), who introduced novel resampling technique to the formulation. A number of statisticians also independently rediscovered and developed the sampling-importanceresampling (SIR) idea (Kong et al. (1994), Smith and Gelfand (1992)), which was originally proposed by Rubin (1987) in a non-dynamic framework. The rediscovery of particle filters in the mid-1990s after a long dominant period, partially thanks to the ever increasing computing power. Recently, a lot of work has been done to improve the performance of particle filters which include Musso et al. (2001), Norton and Verse (2002), Torma and Szepesvari (2001). Most of these works are based on the work of Doucet et al. (2001).

Chapter 3

Methodology

3.1 INTRODUCTION

The Black-Schole's equation which is used in the pricing of options is considered. In this chapter, analytical and numerical solutions to the Black-Schole's equation is discussed. Data assimilation is considered next. Particle filtering is discussed as a data assimilation scheme. An investigation into the performance of numerical solution and Particle filter is conducted.

3.1.1 Black-Scholes

In the build up of Black-Schole's equation,it is necessary to understand the fundamental role of **stochastic differential equation (SDE)**.

Stochastic process is a parametrized collection of random variables $\{X_t\}_{t\epsilon T} = x_t(\omega), \omega\epsilon\Omega$, defined on a probability space (ω,f,P) and take values in R^n. The price of underlying are often times very erratic and uncertain in nature. The value of a stock follows directly from Brownian motion, a form of stochastic process.

Brownian motion,W_t, is a stochastic process, with three main properties:

- $W_t = 0$
- {W(t), t≥ 0} has stationary and independent increments.
- W_t has independent increments with $W_t - W_s \sim N(0,t-s)$ for $0 \leq s < t$.

Ito calculus is employed to solve the dilemma where Brownian motion is continuous everywhere and differentiable nowhere in tradition calculus. From the

Taylor's theorem, Ito calculus makes the following assumptions, called as Ito's Multiplication Table in Table 3.1;

Table 3.1: Ito's Multiplication Table

	dW_t	dt
dW_t	dt	0
dt	0	0

Ito's lemma for Brownian motion, which use Table 3.1 above, states the following:

Assuming $X(t)= X_t$ satisfies the following stochastic differential equation

$$dX_t = \mu(x,t)dt + \sigma(x,t)dW_t \tag{3.1}$$

and $f(t,x)$ is any differentiable twice function of x,t.

For

$$df(t,x) = \left(\frac{\partial f}{\partial t} + \mu\frac{\partial f}{\partial x} + \frac{\sigma^2\partial^2 f}{2\partial x^2}\right)dt + \sigma\frac{\partial f}{\partial x}dW_t \tag{3.2}$$

A stochastic process, X_t is called a geometric Brownian motion (GBM) with parameters μ and σ^2 if its logarithm forms a Brownian motion with mean μ and variance rate σ^2.

The price of a stock follows a GBM process with μ and σ constant. Furthermore, the GBM satisfies the following stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{3.3}$$

Where

S_t is the underlying asset price at time t, μ is the rate of return on risk-less asset (or drift), σ captures the volatility of the stock, W_t represent a Brownian motion.

3.1.2 Black -Scholes Model

Black-Scholes model is a parabolic partial differential equation with a closed-form solution obtained by changing the equation by use of a change of variable into a heat equation. Then the Black-Schole equation has become a simple parabolic PDE whose solution is known since that the solution of heat equation is also known.

ASSUMPTION

The Black-Scholes model provides answers to the problem of option pricing by constructing a portfolio made up of cash, options, and underlying asset. By the use of the following assumption, analysis of Black and Scholes (1973) can be done:

- The underlying asset follows a lognormal random walk with μ and σ as constant.
- The short selling of the underlying asset is permitted.
- The risk-free interest rate, r is constant.
- There are no transaction cost or taxes.
- All asset are perfectly divisible.
- There are no dividends on asset.
- Trading in underlying asset can be done continuously.
- There are no risk-less opportunities for arbitrage.

3.1.3 The Black-Scholes Equation

With the assumption that the underlying asset price, S_t follows the GBM,

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{3.3}$$

with μ and σ constant and W_t being a Brownian motion. The value of the option is denoted by $V = V(S,t)$.

With the help of Ito's lemma, (3.1) becomes

$$dV(S,t) = \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial t}dt + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}dSdS + \frac{1}{2}\frac{\partial^2 V}{\partial t^2}dtdt + \frac{\partial^2 V}{\partial S \partial t}dSdt \tag{3.4a}$$

$$= \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{\partial^2 S^2}{2}\frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S\frac{\partial V}{\partial S}dW_t \tag{3.4b}$$

3.1.4 Portfolio

To derive the Black-Scholes equation, we begin with a basic portfolio. Assuming the portfolio comprises of a call option of value V , a function of the stochastic variable S and a deterministic variable t, with Δ units of underlying asset with price S. Assuming Π is the value of a portfolio, then $\Pi(0) = \Pi_0$. The value of the portfolio at time t is

$$\Pi = -V + \Delta S \tag{3.5}$$

And an infinitesimal change in time period dt with Δ remain constant, leads to a change in the value of the portfolio which is given by

$$d\Pi = -V + \Delta S \tag{3.6a}$$
$$= \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S\frac{\partial V}{\partial S}dW_t + \partial dS \tag{3.6b}$$
$$= \left(\frac{\partial V}{\partial t} + \mu S\left(\Delta - \frac{\partial V}{\partial S}\right) + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S\left(\Delta - \frac{\partial V}{\partial S}\right)dW_t \tag{3.6c}$$

Let

$$\Delta = \frac{\partial V}{\partial S}$$

gives a complete deterministic equation because of the dt term and the removal of terms associated with μ. We then have

$$d\Pi = -\left(\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2}\right)dt \tag{3.7}$$

Letting $\Delta = \frac{\partial V}{\partial S}$ removes all uncertainty and makes the portfolio riskless since there is no stochastic variable, dW_t.

If the risk free interest rate is r, then the portfolio becomes

$$d\Pi = r\Pi dt \tag{3.8a}$$
$$= r\left(-V + \Delta S\right)dt$$
$$= \left(-rV + rS\frac{\partial V}{\partial S}\right)dt.$$
(3.8b) (3.8c)

An investor exercise an option only when there is profit to be made. Due to the no-arbitrage principle, riskless portfolio cannot have value greater than the option's portfolio, thus gives

$$-rV + rS\frac{\partial V}{\partial S} \leq -\left(\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2}\right) \quad (3.9)$$

Thus the famous Black-Scholes partial differential equation is

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \quad (3.10)$$

3.1.5 Transformation of Black-Scholes into the Diffusion equation

The solution of the Heat equation is already known. Therefore transforming the Black-Scholes to the second order heat equation makes it easy to solve.

With Black-Scholes equation in the form,

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \quad (3.10)$$

Let $S = Ke^x, \quad t = T - \frac{\tau}{\frac{\sigma^2}{2}}$ be the change of the independent variable and let $v(x,\tau) = \frac{1}{K}V(S,t) = \frac{1}{K}V(Ke^x, T - \frac{\tau}{\frac{\sigma^2}{2}})$ be the change in the dependent variable

$$\begin{aligned} \frac{\partial V}{\partial t} &= -\frac{\sigma^2}{2}K\frac{\partial v}{\partial \tau} \\ \frac{\partial V}{\partial S} &= \frac{K}{S}\frac{\partial v}{\partial x} \end{aligned},$$

Putting this back into (1) we have

$$-\frac{\sigma^2}{2}K\frac{\partial v}{\partial \tau} + \frac{\sigma^2}{2}K\left(\frac{\partial^2 v}{\partial x^2} - \frac{\partial v}{\partial x}\right) + rK\frac{\partial v}{\partial x} - rKv = 0 \quad (3.11)$$

Finally, the Black-Scholes equation can be reduced to a simple heat equation with these transformations:

Constant transformations:

$$\kappa = \frac{r}{\frac{\sigma^2}{2}}, \qquad \gamma = \frac{1}{2}(\kappa - 1), \qquad \beta = \frac{1}{2}(\kappa + 1)$$

Variable transformations:

$$S = Ke^x, \qquad t = T - \frac{\tau}{\frac{\sigma^2}{2}}.$$

Therefore:

$$V(S,t) = Kv(x,\tau) \tag{3.12a}$$

$$v(x,\tau) = e^{-\gamma x-\beta^2\tau}u(x,\tau). \tag{3.12b}$$

$$V(S,t) = Ke^{-\gamma x-\beta^2\tau}u(x,\tau). \tag{3.12c}$$

3.1.6 Pricing Call Option

For a call option,the solution to Black-Scholes equation after transforming it into a heat equation gives the following results:

$$u(x,0) = v(x,0)e^{\gamma x}. \tag{3.13a}$$

$$v(x,0) = max\{e^x - 1,0\} \tag{3.13b}$$

Therefore

$$u(x.0) = max\{e^{(\gamma+1)x} - e^{\gamma x}, 0\} \tag{3.14}$$

Finally, the price of a Call Option is given by

$$V_c(S,t) = Kv(x,\tau) \tag{3.15a}$$

$$= S\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2) \tag{3.15b}$$

With

$$x = ln(\frac{S}{K})$$

$$t = T - \frac{\tau}{\frac{\sigma^2}{2}} \Rightarrow \quad \tau = (T-t)(\frac{\sigma^2}{2})$$

$$d_1 = \frac{In(\frac{S}{K}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{(T-t)}}$$

$$d_2 = \frac{In(\frac{S}{K}) + ((r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{(T-t)}}$$

$S = Ke^x \Rightarrow$

Where

S is the stock price at time t, K is the strike price of the option, r is the risk-free interest rate,

T is the maturity time,

$\Phi(d)$ is the CDF of the standard normal distribution of d.

3.1.7 Pricing Put Option

The value of Put Option is very similar to that of a Call Option but the Put Option is the negative of the Call Option. So we obtain the following.

$$u(x,0) = v(x,0)e^{\gamma x}. \tag{3.16a}$$

$$v(x,0) = max\{1 - e^x, 0\} \tag{3.16b}$$

Giving

$$u(x.0) = max\{e^{\gamma x} - e^{(\gamma+1)x}, 0\} \tag{3.17}$$

Then finally, for a Put Option

$$V_p(S,t) = Kv(x,\tau) \tag{3.18a}$$

$$= Ke^{-r(T-t)}\Phi(-d_2) - S\Phi(-d_1) \tag{3.18b}$$

3.2 NUMERICAL SOLUTION

Three different finite difference methods for the European style option will be used in solving the Black-Scholes equation.

3.2.1 Finite Difference Methods

The finite difference methods are used by approximating the continuous-time differential equation which shows how an option changes over time by a set of discrete-time difference equation.This discrete-time difference equation is then solved iteratively to find a price of the option

The finite difference methods to be used include;

- Implicit method
- Explicit method
- Crank-Nicolson method

Implicit Finite difference method is considered in this thesis. The implicit method is more stable compared with the explicit method. But the Crank-Nicolson method is more stable than the implicit method.

There is a direct relation between the truncation error and the rate of convergence. The explicit and implicit methods both converge at the rate of $O(\delta t)$ and $O(\delta S^2)$. The Crank-Nicolson method converges at the rate of $O(\delta t)$ and $O(\delta S^2)$ which is obviously faster compared to the explicit and implicit method.

Hence the Crank-Nicolson method converges at the rates of $O(\delta t)$ and $O(\delta S^2)$. This is a faster rate of convergence than either the explicit method, or the implicit method. Also the explicit method converges at the rates of $O(\delta t)$ and $O(\delta S^2)$. This is the same convergence rate as the implicit method, but slower than the Crank-Nicolson method. Finally the implicit method converges at the rates of $O(\delta t)$ and $O(\delta S^2)$. This is the same convergence rate as the explicit method, but slower than the Crank-Nicolson method.

3.2.2 Numerical Scheme

The Implicit Finite Difference Method

The Black-Scholes equation,

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \qquad (3.19)$$

the implicit finite difference method discretizes it by use of the following formulae

- the forward approximation for $\frac{\partial V}{\partial t}$

$$\frac{\partial V}{\partial t} = \frac{V_{i,j+1} - V_{i,j}}{\Delta t}$$

- the central approximation for $\frac{\partial V}{\partial S}$

$$\frac{\vartheta V}{\vartheta S} = \frac{V_{i+1,j} - V_{i-1,j}}{2\Delta S}$$

- the standard approximation for $\frac{\partial^2 V}{\partial S^2}$

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

rewriting Black-Scholes equation,it becomes

$$\frac{V_{i,j+1} - V_{i,j}}{\Delta t} + \frac{1}{2}\sigma^2[j\Delta S]^2[\frac{V_{i+1,j} + V_{i-1,j} - 2V_{i,j}}{\Delta S^2}] + rj\Delta S[\frac{V_{i+1,j} - V_{i-1,j}}{2\Delta S}] - rV_{i,j} = 0$$

(3.20)

Multiplying through by $\Delta t \Delta S^2$ and

let

$$a = \frac{1}{2}\sigma^2[j\Delta S]^2, \qquad b = \frac{1}{2}rj\Delta S$$

$$[V_{i,j+1} - V_{i,j}]\Delta S^2 + a[V_{i+1,j} + V_{i-1,j} - 2V_{i,j}]\Delta t + b[V_{i+1,j} - V_{i-1,j}]\Delta t\Delta S - rV_{i,j}\Delta t\Delta S^2 = 0$$

(3.21)

Let $c = \frac{\Delta t}{\Delta S^2}, \quad d = \frac{\Delta t}{\Delta S}$ and simplify making $V_{i,j+1}$ the subject

$$V_{i,j+1} = V_{i,j} + acV_{i+1,j} + acV_{i-1,j} - 2acV_{i,j} + bdV_{i+1,j} - bdV_{i-1,j} - r\Delta tV_{i,j}$$

(3.22a)

$$= [1 - 2ac - r\Delta t]V_{i,j} + [ac + bd]V_{i+1,j} + [ac - bd]V_{i-1,j} \qquad (3.22b)$$

$$= \gamma V_{i,j} + \beta V_{i+1,j} + \alpha V_{i-1,j} \tag{3.22c}$$

where

$$\begin{aligned}
\gamma = 1 - 2ac - r\Delta t &= 1 - [\sigma^2 j^2 + r]\Delta t \\
\beta = ac + bd &= \frac{1}{2}[\sigma^2 j^2 + r]\Delta t \\
\alpha = ac - bd &= \frac{1}{2}[\sigma^2 j^2 - r]\Delta t
\end{aligned}$$

3.3 DATA ASSIMILATION

Data assimilation is primarily interested with the use of observational data into mathematical models. The Bayesian view of data assimilation where prior information about a system is combined with data to give a posterior distribution. Data assimilation involves observing a physical process by use of a model and observed data. There are basically two component of Data Assimilation: a model, which give to the best possible degree the process of interest, and observations, which help in estimating the model parameters (the state) in space and in time if applicable.

There are two major ways to discuss data assimilation. The deterministic approach where a single optimal estimate of the true process is sought after and the stochastic approach where the uncertainty associated with estimate is considered. The Bayesian framework for the formulation of the stochastic data assimilation is considered, where the probability distribution of the estimate is followed instead of the estimate only.

3.3.1 Dynamical System

According to Weisstein (2002), a dynamical system can be defined as a means of describing how one state develops into another state over the course of time. The mathematical formulation that provides factors assumed to be responsible for the dynamic process is a dynamic system. The main design of any dynamic system is to understand and reproduce the evolutionary phenomena observed in the real world.

The major interest of data assimilation is to track and predict the state of process and in most application, these processes are time dependent. The required model to express such temporal change of the process are often referred to as dynamical systems or dynamical models. These dynamical system can further be classified into deterministic systems or stochastic systems.

3.3.2 Deterministic System

The deterministic character of a dynamic system is, for a given initial condition x_0, the future transformation of the system is completely determined. This transformation of a deterministic system may be expressed as a rule relating the state of the system at a given instant to its state at a later time. A differential equation is generally used in expressing this rule.

$$\frac{dx}{dt} = f(x,t) \tag{3.23}$$

where f is the system operator responsible for propagating the state forward in time.

3.3.3 Stochastic System

Physical phenomenon is a result of many causes but only the most important is identified and use in deterministic system. Most phenomenon are approximated by an incomplete system. To account for the approximation between the true process and the system, a stochastic term is added to the deterministic system.

$$\frac{dx}{dt} = f(x,t) + h(x,t) \tag{3.24}$$

h(x,t) represent effect of all the unrepresented factors in the model and is often referred to as system noise or model error.

3.3.4 Stochastic Approach

The stochastic approach of data assimilation in the dynamic framework is considered. The state of the model and the state of the observation are related as follows:

$$x_t = m_t(x_{t-1}) + \eta_t \tag{3.25a}$$

$$y_t = h_t(x_t) + \varepsilon_t \tag{3.25b}$$

One concern lies in evaluating the joint probability density function of the state specified for all observations up to which includes a specified time t: $p(X_t|Y_t)$.

For the model and observation equations that are linear with Gaussian distribution, the Kalman filter (Kalman (1960)) gives an optimal solution which minimises the variance to the filtering problem. The Extended Kalman Filter (Jazwinski (1970); Maybeck (1979)) gives a linear approximation to the Kalman filter whiles the Ensemble Kalman Filter (Evensen (2003)) provides a Monte Carlo approximation as well. The Particle Filter (Doucet et al. (2001)) is another Monte-Carlo approximation that helps in dealing with non-Gaussian distributions. Particle Filtering is discussed in the remaining of section.

3.3.5 Bayesian Framework

Suppose data $x = (x_1,...,x_n)$ with distribution $p(x|\theta)$ where θ is the unknown parameter we want to estimate. The basic idea of the Bayesian approach is to treat the parameter θ as a random variable and to use a priori knowledge of the distribution $\pi(\theta)$ of θ and then to estimate θ by calculating the a posteriori distribution $\pi(\theta|x)$ of θ.

The one-dimension case.

In the one-dimensional case the a posteriori distribution $\pi(\theta|x)$ of θ is calculated by the so called *Bayes*^{0}*sformula* using the priori distribution $\pi(\theta|x)$ as follows

$$\pi(\theta|x) = \frac{p(x|\theta)\pi(\theta}{\int p(\theta)\pi(\theta)d\theta} \tag{3.26}$$

where the denominator is a proportionality making the total a posteriori probability equal to one. Now by using the posterior distribution $\pi(\theta|x)$ the parameter θ can be estimated by the mean $\hat{\theta} = E[\pi(\theta|x)]$.

Multi-dimension case.

In the multi-dimensional case $\theta = (\theta_1,...,\theta_k)$, the a posteriori distribution of θ can be calculated by the Bayes formula as follow

$$\pi(\theta|x) = \frac{p(x|\theta)\pi(\theta)}{\int \ldots \int p(\theta|x)\pi(\theta)d\theta_1 \ldots \theta_k}. \tag{3.27}$$

By using the marginal distribution $\pi(\theta_i|x)$ of the joint posterior distribution $\pi(\theta|x)$

$$\pi(\theta_i|x) = \int \ldots \int \pi(\theta|x)d\theta_1 \ldots d\theta_{i-1}d\theta_{i+1} \ldots d\theta_k \tag{3.28}$$

θ is estimated by the ways described in the one - dimensional case. Usually problems arise in calculating the integrals in equation which require approximation techniques as *Markov Chain Monte Carlo Methods (MCMC)*.

3.3.6 Framework

The hidden system state x_k, with initial probability density $p(x_0)$, over time changes to be a partially observed first order Markov process according to the conditional probability density $p(x_k|x_{k-1})$. The observations y_k are conditionally independent given the state and are generated according to the conditional probability density $p(y_k|x_k)$. The evolution of the state is produced by the Transition equation:

$$x_k = f_k(x_{k-1}, u_{k-1}, v_k; w), \tag{3.29}$$

and the Measurement equation is given by:

$$y_k = h_k(x_k, n_k; w), \tag{3.30}$$

where:

x_k:the state vector at the time k, y_k:the measurement vector, u_k:an external input of the system, assumed known, v_k:the process noise that drives the dynamic system, n_k:the measurement noise corrupting the observation of the state, f_k:a time-variant, linear or non-linear function, h_k:a time-variant, linear or non-linear function, w: the parameters vector.

The state transition density $p(x_k|x_{k-1})$ is fully specified by f_k and the process noise distribution $p(v_k)$, whereas h_k and the observation noise distribution $p(n_k)$ fully specify the observation likelihood $p(y_k|x_k)$.

3.4 PARTICLE FILTERING

The Particle Filter is a filter that is based on the Monte Carlo methods. The Monte Carlo methods gives an approximation to continuous distribution by use of a discrete set of samples $\{x_i\}_{i=1:N}$ called "particles". This filter does not require the state's probability density function to be Gaussian.

Let the system be:

$$x_k = f_k(x_{k-1}, u_{k-1}, v_k; w),$$

$$y_k = h_k(x_k, n_k; w), \tag{3.31}$$

The algorithm of Particle filter is made up of the following steps:*Initialization, Prediction, Updating* and *Resampling*.

According to Dablemont et al. (2009), during the initialization, sampling is taken N times from the initial distribution η_0. By sampling x^i from a distribution μ, for i = 1,...,N, a simulation of N independent random samples, named particles, according to μ. Hence, the N random variables x^i for $i = 1,...,N$ are independent and identical distributed (i.i.d.) according to η_0. Afterwards, the values of the particles are predicted for the next time step according to the dynamics of the state Markov process. During the ”Updating” step, each predicted particle is weighted by the likelihood function $g_k(y_k - h_k(.))$, which is determined by the observation process. The ”Resampling” step can be view as a special case of a ”Selection” step. The particles are selected in accordance with the weighting function g_k. This step gives birth to some particles at the expense of light particles which die.

3.4.1 Sequential Importance Sampling

The *Sequential Importance Sampling* (SIS) has always been a *Monte Carlo MC* method at the foundation of most sequential Monte Carlo filters developed for a number of years now. A recursive Bayesian filter by Monte Carlo simulation is implemented in SIS. The working technique of particle filters is as follows: The state-space is partitioned into many parts, where the particles fill in based on a probability measure. With higher probability comes a concentration of denser particles. The particle system evolves with state equation along with time. The evolving pdf is represented by a number of particles provided by the random sampling of the state space. This ramdom sampling of the state space is an approximation of the initial pdf. There is always a difficult in sampling posterior density model, since it is unknown. An alternative distribution is chosen for the purposes of efficient sampling.

For the detailed algorithm, let $\{x^i_{0:k}, w^i_k\}$ for $i = 1,...,N_s$ a Random measure that characterizes the posterior pdf $p(x_{0:k}|y_{1:k})$, where $\{x^i_{0:k}\}$ for $i = 1,...,N_s$ is a set of support points with associated weight $\{w^i_k\}$ for $i = 1,...,N_s$ and $x_{0:k} = \{x_j, j = 0,...,k\}$ is the posterior density at k can be approximated as

$$p(x_{0:k}|y_{1:k}) \approx \sum_{i=1}^{N_s} w_i \delta(x_{0:k} - x_{i0:k}) \equiv \hat{p}(x_{0:k}|y_{1:k}) \tag{3.32}$$

where $\{x^i_{0:k}\}$ are assumed to be i.i.d. drawn from $p(x_{0:k}|y_{1:k})$. For sufficiently large N_s, $\hat{p}(x_{0:k}|y_{1:k})$ give an approximation of the true posterior $p(x_{0:k}|y_{1:k})$. The mean of a nonlinear function can be estimated by this approximation.

$$E[f(x_{0:k})] = \int f(x_{0:k})\hat{p}(x_{0:k}|y_{1:k})dx_{0:k} \tag{3.33a}$$

$$= \frac{1}{N_s}\sum_{i=1}^{N_s}\int f(x_{0:k})\delta(x_{0:k} - x^i_{0:k})dx_{0:k} \tag{3.33b}$$

$$= \frac{1}{N_s}\sum_{i=1}^{N_s} f(x^i_{0:k}) \equiv \hat{f}_{N_s}(x). \tag{3.33c}$$

The proposal distribution which is an easy-to-implement distribution from where the sampling is done. This is denoted by $q(x_{0:k}|y_{1:k})$, therefore

$$E[f(x_0 : k)] = \int f(x_{0:k})\frac{p(x_{0:k}|y_{1:k})}{q(x_{0:k}|y_{1:k})}q(x_{0:k}|y_{1:k})dx_{0:k} \tag{3.34a}$$

$$= \int f(x_{0:k})\frac{p(y_{1:k}|x_{0:k})p(x_{0:k})}{p(y_{1:k})q(x_{0:k}|y_{1:k})}q(x_{0:k}|y_{1:k})dx_{0:k} \quad (3.34b)\ (3.34c)$$

$$= \int f(x_{0:k})\frac{w_k(x_{0:k})}{p(y_{1:k})}q(x_{0:k}|y_{1:k})dx_{0:k}$$

where $w_k(x_{0:k})$ are the unnormalized importance weights, and are given by

$$w_k(x_{0:k}) = \frac{p(y_{1:k}|x_{0:k})p(x_{0:k})}{q(x_{0:k}|y_{1:k})} \tag{3.35}$$

Then

$$E[f(x_{0:k})] = \frac{1}{p(y_{1:k})}\int f(x_{0:k})w_k(x_{0:k})q(x_{0:k}|y_{1:k})dx_{0:k} \tag{3.36a}$$

$$= \frac{E_q[f(x_{0:k})w(x_{0:k})]}{E_q[w_k(x_{0:k})]} \tag{3.36b}$$

where $E_q[.]$ denotes the expectations taken over the proposal distribution $q(x_{0:k}|y_{1:k})$. The independent identical distribution (iid) sample $\{x^{(i)}_{0:k}\}$ is drawn from thee proposal distribution $q(x_{0:k}|y_{1:k})$, gives an approximation of the expectation by:

$$E[f(x_{0:k})] \approx \tilde{E}[f(x_{0:k})] = \frac{\frac{1}{N_x}\sum_{i=1}^{N_x} f(x)^{(i)}_{0:k} w_k(x)^{(i)}_{0:k}}{\frac{1}{N_x}\sum_{i=1}^{N_x} w_k(x^{(i)}_{0:k})} \tag{3.37}$$

$$= \sum_{i=1}^{N_x} \tilde{w}_k^{(i)} f(x_{0:k}^{(i)}), \tag{3.37b}$$

where $\tilde{w}_k^{(i)}$ denotes the normalized importance weights are provided by:

$$\tilde{w}_k^{(i)} = \frac{w_k(x_{0:k}^{(i)})}{\sum_{i=1}^{N_x} w_k(x_{0:k}^{(i)})} \tag{3.38}$$

If the proposal distribution has the form:

$$q(x_{0:k}|y_{1:k}) = q(x_{0:k-1}|y_{1:k-1})q(x_k|x_{0:k-1}, y_{1:k}) \tag{3.39a}$$

$$= q(x_0)\prod_{j=1}^{k} q(x_j|x_{0:j-1}, y_{1:j}) \tag{3.39b}$$

With the assumptions that the states corresponds to a first order Markov process and the observations are conditionally independent given the states, we obtain:

$$p(x_{0:k}) = p(x_0)\prod_{j=1}^{k} p(x_j|x_{j-1}) \tag{3.40a}$$

$$p(y_{1:k}|x_{0:k}) = \prod_{j=1}^{k} p(y_j|x_j) \tag{3.40b}$$

factorizing the posterior distribution to obtain:

$$p(x_{0:k}|y_{1:k}) = \frac{y_k|x_{0:k}, y_{1:k-1}p(x_{0:k}|y_{1:k-1})}{p(y_k|y_{1:k-1})} \tag{3.41a}$$

$$= \frac{p(y_k|x_k)p(x_k|x_{k-1})}{p(y_k|y_{1:k-1})} p(x_{0:k-1}|y_{1:k-1}) \tag{3.41b}$$

$$\propto p(y_k|x_k)p(x_k|x_{k-1})p(x_{0:k-1}|y_{1:k-1}) \tag{3.41c}$$

Also factorizing a recursive estimate for the importance weights to obtain:

$$w_k^{(i)} = \frac{p(x_{0:k}^{(i)}|y_{1:k})}{q(x_{0:k}^{(1)}|y_{1:k})} \tag{3.42a}$$

$$= w_{k-1}^{(i)} \frac{p(y_k|x_k^{(i)})p(x_k^{(i)}|x_{k-1}^{(i)})}{q(x_k^{(i)} x_{0:k-1}^{(i)}, y_{1:k})} \tag{3.42b}$$

Also, if $q(x_k^{(i)} x_{0:k-1}^{(i)}, y_{1:k}) = q(x_k^{(i)} x_{0:k-1}^{(i)}, y_k)$, then the importance density depends on only x_{k-1} and y_k. The modified weight is:

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(y_k|x_k^i)p(x_k^{(i)}|x_{k-1}^{(i)})}{q(x_k^{(i)}|x_{0:k-1}^{(i)}, y_k)}, \quad (3.43)$$

and the posterior filtered density $p(x_k|y_{1:k})$ can be approximated as:

$$p(x_k|y_{1:k}) \approx \sum_{i=1}^{N_x} w_k^{(i)} \delta(x_k - x_k^{(i)}), \quad (3.44)$$

where the weights are defined in (48). It can be shown that as $N_x \to \infty$ the approximation (49) approaches the true posterior density $p(x_k|y_{1:k})$.

The SIS algorithm consists of recursive propagation of the weights and points as each measurement is received sequentially. A pseudo-code description of this algorithm is provided by algorithm 1.

> $[\{x_k^i, w_k^i\}_{i=1}^{N_s}] = \text{SIS } [\{x_k^i, w_k^i\}_{i=1}^{N_s}, y_k]$
> **for** $i = 1; N_s$ **do**
> Draw $x_k^i \sim q(x_k^{(i)}|x_{0:k-1}^{(i)}, y_k)$
> Evaluate the importance weights up to a normalizing constant
> $w_k^{(i)} = w_{k-1}^{(i)} \frac{p(y_k|x_k^i)p(x_k^{(i)}|x_{k-1}^{(i)})}{q(x_k^{(i)}|x_{0:k-1}^{(i)}, y_k)}$,
> **end for**

Algorithm1: SISParticleFiltering

3.4.2 Degeneracy Problem

A major setback with the SIS particle filter is the degeneracy phenomenon, in that after a few iterations, all but one particle will have negligible weight. Doucet (1998) has shown that the variance of the importance weights can only increase over time and so it is impossible to avoid the degeneracy phenomenon. A large computational effort is put into updating particles whose contribution to the approximation to $p(x_k|y_{1:k})$ is almost zero. One approach to dealing with this situation, is the use of a very large number of particles N_s. But this approach is often very impractical. A good choice of Importance density and the use of Resampling are two alternative method to the solve the degeneracy problem.

3.4.3 Resampling

The elimination of particles which have small weights and the concentration on particles with very large weights is the underline idea of resampling. This method involves generating a new set $x_k^{(t)*}$ for $i = 1,...,N_s$ by resampling with replacement N_s times from an approximate discrete representation of $p(x_k|y_{1:k})$ given by

$$p(x_k|y_{1:k}) \approx \sum_{t=1}^{N_s} w_k^i \delta(x_k - x_k^i) \tag{3.45}$$

so that $Pr(x_k^{(t)*} = x_k^j) = w_k^j$. The resulting sample is in fact an i.i.d sample from the discrete density (6), and the weights are not reset to $w_k^i = \frac{1}{N_s}$. The systematic resampling Kitagawa (1996) is implemented in algorithm 2 since it is simple to implement, takes $O(N_s)$ and minimises the MC variation. For each resampled particle $x_k^{j^*}$, this resampling algorithm also stores the index of its parent, denoted by i^j.

A generic particle filter is described by algorithm 3.

$[\{x_k^{j^*}, w_k^j, i^j\}_{j=1}^{N_s}]$ = RESAMPLE $[\{x_k^i, w_k^i\}_{i=1}^{N_s}]$
Initialize the CDF: $c_1 = w_k^1$
for $i = 2 : N_s$ **do**
  Construct CDF: $c_i = c_{i-1} + w_k^i$
end for
Start at the bottom of the CDF: $i = 1$ Draw
a starting point: $u_j \sim \mathbf{U}(0, N_s^{-1})$ **for** $j = 1 : N_s$
do
  Move along the CDF: $u_j = u_1 + N_s^{-1}(j-1)$
  while $u_j > c_i$ **do**
    $i = i + 1$ **end while**
  Assign sample: $x_k^{j^*} = x_k^i$
  Assign weight: $w_k^j = N_s^{-}$

Assign parent: $i^j = i$ **end for**

Algorithm 2: Resampling Particle Filtering

$[\{x_k^i, w_k^i\}_{i=1}^{N_s}] = \text{PF } [\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^{N_s}, y_k]$
for $i = 1 : N_s$ **do**
Draw $x_k^i \sim q(x_k^{(i)}|x_{0:k-1}^{(i)}, y_k)$
Assign the particle a weight,w_k^i, according to:
$w_k^{(i)} = w_{k-1}^{(i)} \frac{p(y_k|x_k^i)p(x_k^{(i)}|x_{k-1}^{(i)})}{q(x_k^{(i)}|x_{0:k-1}^{(i)}, y_k)}$,
end for
Calculate total weight: $t = SUM[\{w_k^i\}_{i=1}^{N_s}]$
for $i = 1 : N_s$ **do**
Normalize: $w_k^i = t^{-1} w_k^i$
end for
Calculate $\hat{N}_{eff}$ using
$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N_s}(w_k^i)^2}$
if $\hat{N}_{eff} < N_T$ **then**
Resample using algorithm 2:
$[\{x_k^{j^*}, w_k^j, i^j\}_{j=1}^{N_s}] = \text{RESAMPLE } [\{x_k^i, w_k^i\}_{i=1}^{N_s}]$
end if

Algorithm 3: Generic Particle Filtering

3.4.4 Sampling Importance Resampling Filter

The Sampling Importance Resampling (SIR) filter proposed in Gordon et al. (1993) is a Monte Carlo method that can be applied to recursive Bayesian filtering problems. The assumptions required to use the SIR filter are very weak.

- The state dynamics and measurement function, $f_k(.,.)$ and $h_k(.,.)$ in (3.29) and (3.30)respectively, needs to be known.

- It is required to be able to sample realizations from the process noise distribution of v_{k-1} and from the prior.

- The likelihood function $p(y_k|x_k)$ needs to be available for pointwise evaluation (at least up to proportionality).

It is easy to derive the SIR algorithm from the SIS algorithm by an appropriate choice of:

- The importance density: $q(x_k|x^i_{k-1},y_{1:k})$ is chosen to be the prior density

$p(x_k|x^i_{k-1})$,

- Resampling step: to be applied at every time index.

A sample $x^i_k \sim p(x_k|x^i_{k-1})$ can be generated by first generating a process noise sample $v^i_{k-1} \sim p_v(v_{k-1})$ and setting $x^i_k = f_k(x^i_{k-1}, v^i_{k-1})$, where $p_v(.)$ is the pdf of v_{k-}. For this particular choice of importance density, it is evident that the weights are given by

$$w^i_k \propto w^i_{k-} p(y_k|x^i_k). \tag{3.46}$$

However, noting that resampling is applied at every time index, given $w^i_{k-} = \frac{1}{N} \forall i$ and so

$$w^i_k \propto p(y_k|x^i_k) \tag{3.47}$$

The weights given by the proportionality is (7) are normalized before the resampling stage. The iteration of the algorithm is described by algorithm 4.

$[\{x^i_k, w^i_k\}^{N_s}_{i=1}]$ = SIR $[\{x^i_{k-1}, w^i_{k-1}\}^{N_s}_{i=1}, y_k]$
for $i = 1 : N_s$ **do**
  Draw $x^i_k \sim p(x_k|x^i_{k-1})$
  Calculate $w^i_k = p(y_k|x^i_k)$
end for
Calculate total weight; t = SUM $[\{w^i_k\}^{N_s}_{i=1}]$
for $i = 1 : N_s$ **do**
  Normalize: $w^i_k = t^{-1} w^i_k$
end for
Resample using algorithm 2:
$[\{x^{j*}_k, w^j_k, i^j\}^{N_s}_{j=1}]$ = RESAMPLE $[\{x^i_k, w^i_k\}^{N_s}_{i=1}]$

Algorithm 4: SIR (Bootstrap) Particle Filtering

3.5 Application of Particle Filtering

This section looks at the implementation of the particle filtering method discussed in this chapter to the Black-Scholes European call option and the GARCH model. The Generalized Autoregressive Conditional Heteroskedasticity GARCH model was first introduced by Bollerslev (1986) in the parameterization form. The model is a weighted average of past squared residuals, with declining weights that never goes to zero completely. The standard GARCH(1,1) is of the form:

$$dV_t = x_n = \omega + \alpha u_{n-1} + \beta v_{n-1} + w_n \qquad (3.48)$$

The experimental setup used seek to investigate the performance of the SIR method. The effect of the process noise, measurement noise and the number of particle on the performance of the SIR method is investigated.

In applying the particle filtering, the GARCH model estimates the variance from the underlying stock returns. The state therefore becomes the variance x with control signal being stock returns u and the observations are option price c. The state or dynamic equation therefore is;

$$x_n = \omega + \alpha u_{n-1} + \beta v_{n-1} + w_n \qquad (3.49)$$

where w_n is the normally distributed noise term and ω is the long run average volatility. The state equation as $\alpha + \beta < 1$ as a constraint. The non-linear observation equation thus is the standard Black-Scholes equation for an European call option:

$$C = S_0 N(d_1) - Ke^{-rT} N(d_2) \qquad (3.50)$$

where

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

$$d_2 = d_1 - x\sqrt{T}$$

The experiment is setup to test the performance of SIR in the estimation of volatility as the process noise is varied at different underlying price S_0 for a call option. This was achieved by setup of different experiment for varying the process noise, measurement noise and number of particles to evaluate the performance of the SIR.

In the first experiment, the process noise is varied P_0 = [0.5,1,5] in order to estimate volatility with time t = 1,2,...,100. Figures 3.1 displays volatility with time when the process noise is at P_0 = [0.5,1,5] for the underlying stock price S_0 of \$60 respectively.

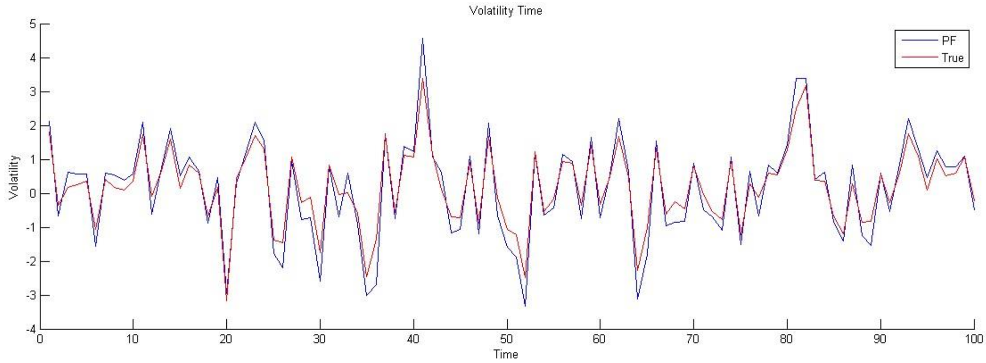

Figure 3.1: The estimated volatility over time of 100 days with a process noise of P_0 of 1 at an underlying price S_0 of \$60 and a strike price K_0 of \$50.

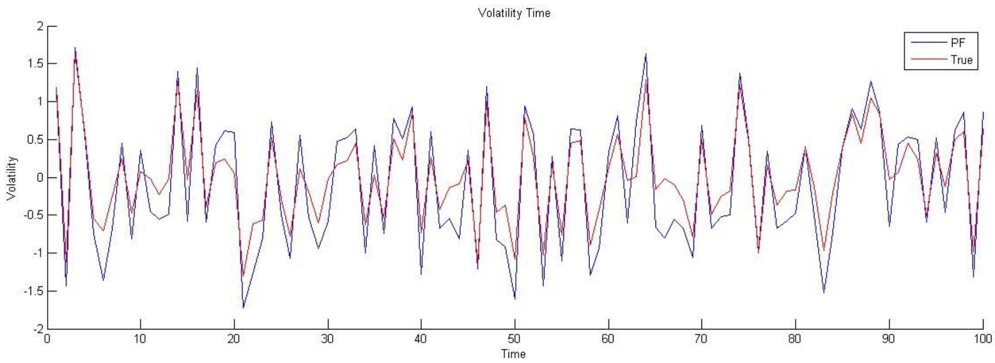

Figure 3.2: The estimated volatility over time of 100 days with a process noise of P_0 of 0.5 at an underlying price S_0 of \$60 and a strike price K_0 of \$50.

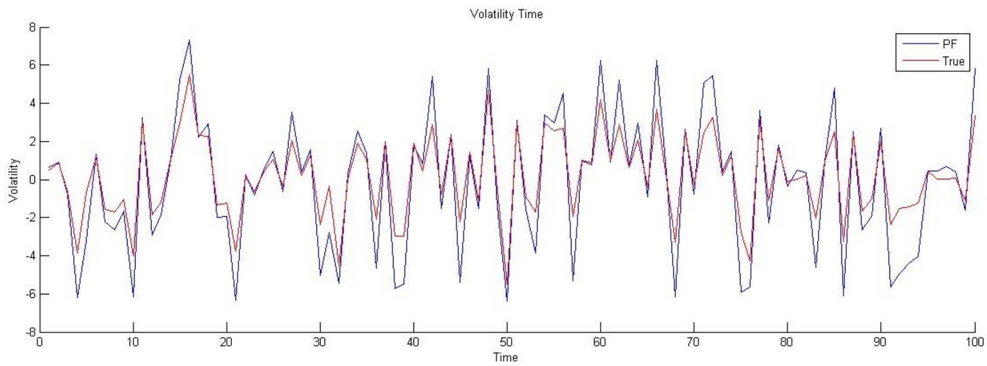

Figure 3.3: The estimated volatility over time of 100 days with a process noise of P_0 of 5 at an underlying price S_0 of \$60 and a strike price K_0 of \$50.

Table 3.2: RMSE of Particle Filter for various process noise

RMSE OF Particle Filter for P_0			
Underlying Price	0.5	1	5
40	0.2220	0.2145	0.8882
50	0.1212	0.1974	1.1781
60	0.3424	0.4282	NaN

Table 3.2 shows the root means square error of the SIR when the process noise is at P_0 = [0.5,1,5]. It is observed that the SIR performs better when the process noise P_0 of 0.5 for underlying asset price S_0 of \$50 and \$60. This is so because of the low RMSE it produce compared to P_0 of 1 and P_0 of 5. Plot of volatility at P_0 of 0.5 displayed in Figure 3.2 show the better estimation of SIR compared to the volatility at P_0 of 1, and at P_0 of 5.

The next experiment vary the measurement noise while other parameters are kept fixed. Table 3.3 summarizes the performance of the SIR under varying measurement noise. Figure 3.2 also display the volatility with time of 100 days at an underlying asset price of \$60.

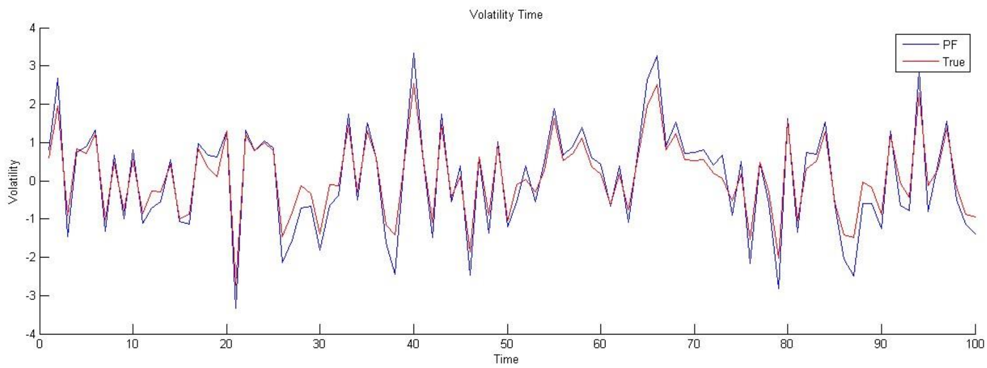

Figure 3.4: The estimated volatility over time of 100 days with a measurement noise of M_0 of 1 at an underlying price S_0 of \$60 and a strike price K_0 of \$50.

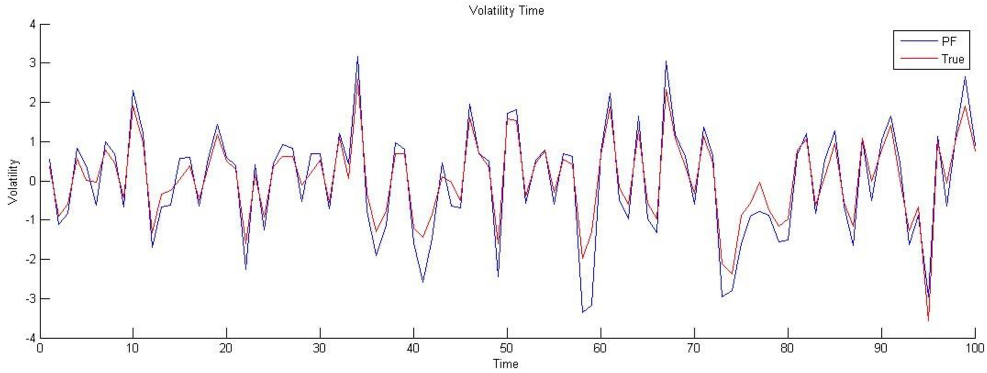

Figure 3.5: The estimated volatility over time of 100 days with a measurement noise of M_0 of 0.5 at an underlying price S_0 of \$60 and a strike price K_0 of \$50.

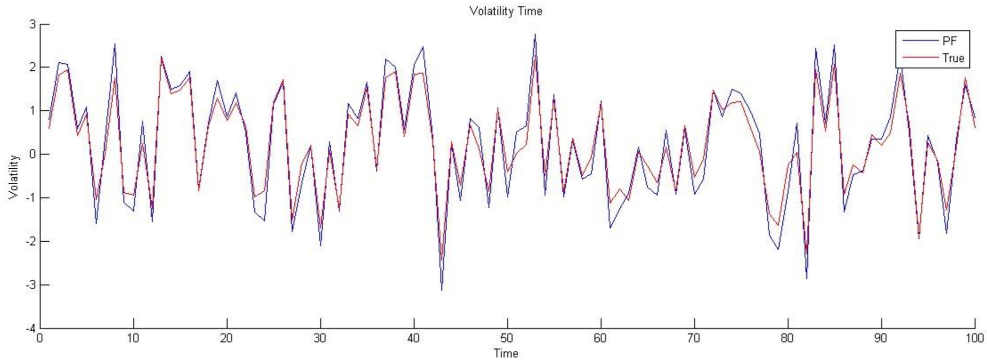

Figure 3.6: The estimated volatility over time of 100 days with a measurement noise of M_0 of 5 at an underlying price S_0 of \$60 and a strike price K_0 of \$50.

Table 3.3: RMSE of Particle Filter for various measurement noise

RMSE OF Particle Filter for M_0			
Underlying Price	0.5	1	5
40	0.2022	0.2131	0.2889
50	0.1907	0.1910	0.2540
60	0.4426	0.4328	0.4054

Table 3.3 reveals that the estimate best when the underlying asset price is equal to the strike price ie S_0 = \$50 = K_0. The SIR performs poorly at S_0 of \$60 given the highest RMSE for all measurement noise tested. In order to be profitable, the call option should have underlying asset S_0 ¿ K_0 the strike price. With interest in the performance of SIR at S_0 of \$60, the SIR performance best at high measurement noise M_0 as suggested by measurement noise M_0 of 5 with RMSE of 0.4054 and poorly at low measurement noise in the case of M_0 of 0.5 with RMSE of 0.4426.

Finally, an experiment to investigate the performance of SIR when the number of particle (nPF) in SIR is varied is performed. The process noise and measurement noise is kept at 1 respectively. Figures 3.7,3.8,3.9 also shows volatility over time for 100 days. A summary of result of the performance of SIR in terms of RMSE is displayed in Table 3.4.

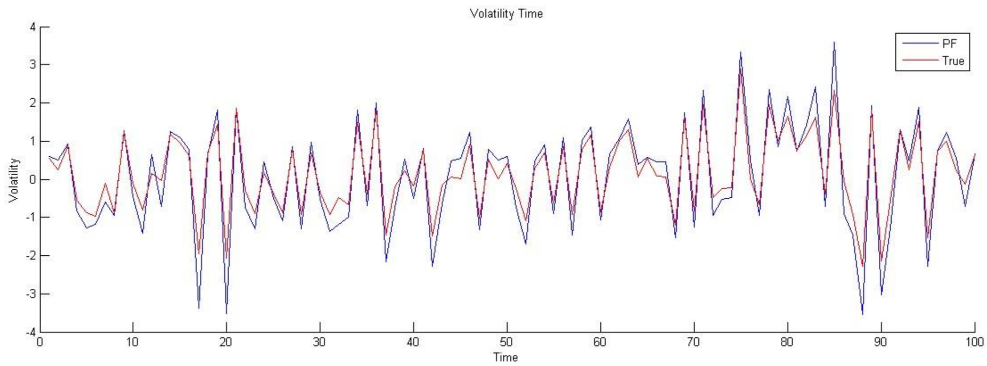

Figure 3.7: The estimated volatility over time of 100 days with 1000 particles

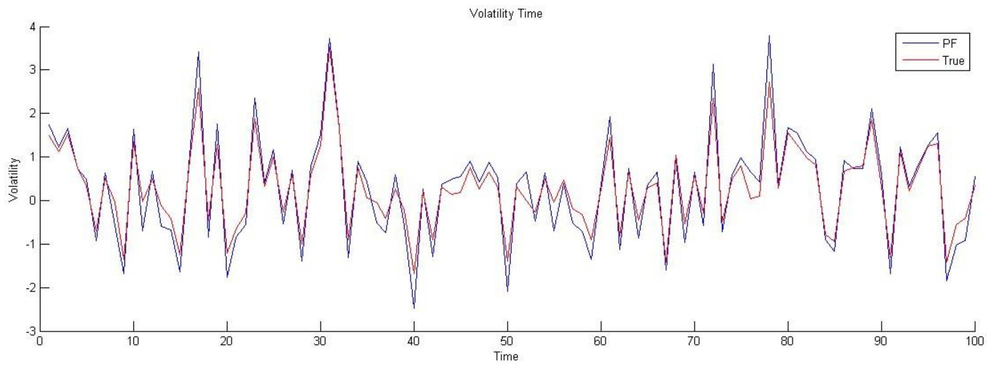

Figure 3.8: The estimated volatility over time of 100 days with 10000 particles

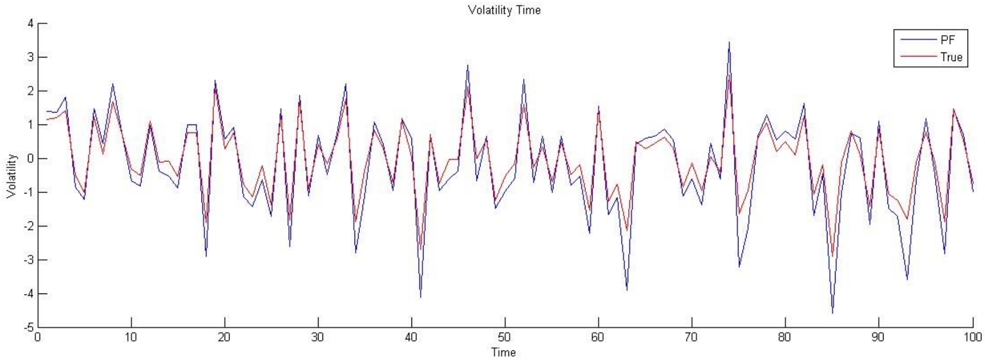

Figure 3.9: The estimated volatility over time of 100 days with 100000 particles
Table 3.4: RMSE of Particle Filter for different number of particle

RMSE OF Particle Filter for nPF			
Underlying Price	1000	10000	100000
40	0.2158	0.2161	0.2211
50	0.1957	0.2139	0.2263
60	0.4352	0.4603	0.4862

The experimental result in Table 3.4 indicates that the RMSE of SIR increases as the number of particles nPF increases. This indicate that the SIR performs best with a small number of particles.

Chapter 4

Analysis

4.1 Introduction

In this chapter,a number of experiments are performed by the use of the particle filtering method. The results of the predictions of volatility of financial options in the Black-Scholes models by the use of the particle filter specifically the bootstrap method is outlined. The estimates were obtained for 100 samples and experiment was repeated 100 times given a mean value, with associated variances similar to zero. Synthetic data is generated and used in evaluation of the pricing algorithm. At the foundation of the option pricing is the famous Black-Scholes model. The partial differential equation describes the Black-Scholes model best.

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{4.1}$$

where $V(S,t)$ is a European put or call option with an underlying asset price $S$and at time t, $r(t)$ is the risk free interest rate with volatility of underlying asset σ. The Black-Scholes model describes the European call option $V_c(S,t)$ and put option $V_p(S,t)$ with no-dividend payment on stocks as:

$$V_c(S,t) = S\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2) \tag{4.2}$$

$$V_p(S,t) = Ke^{-r(T-t)}\Phi(-d_2) - S\Phi(-d_1) \tag{4.3}$$

where K is the strike price, $T - t$ is the time until expiration. $\Phi(.)$ is the cumulative normal distribution function and d_1 and d_2 are:

$$
\begin{aligned}
d_1 &= \frac{In(\frac{S}{K}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{(T-t)}} \\
d_2 &= \frac{In(\frac{S}{K}) + ((r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{(T-t)}}
\end{aligned}
$$

The hidden states is the volatility of underlying states whereas the call and put options are considered as the output observation. The input observations are the current value of underlying asset price and the time to maturity. Thus the model setup represents a parameter estimation problem with the observation equation given by Equation 4.2 and Equation 4.3 which allows us to compute for the daily probability distributions for the volatility whiles keeping the risk free rate constant.

4.2 Result and Discussion

4.2.1 Call Option

Figure 4.1 and figure 4.2 present the estimated volatility and the estimated price with respect to time at underlying price S_0= \$50 and \$60 respectively, with strike price K=\$50, for a call option.

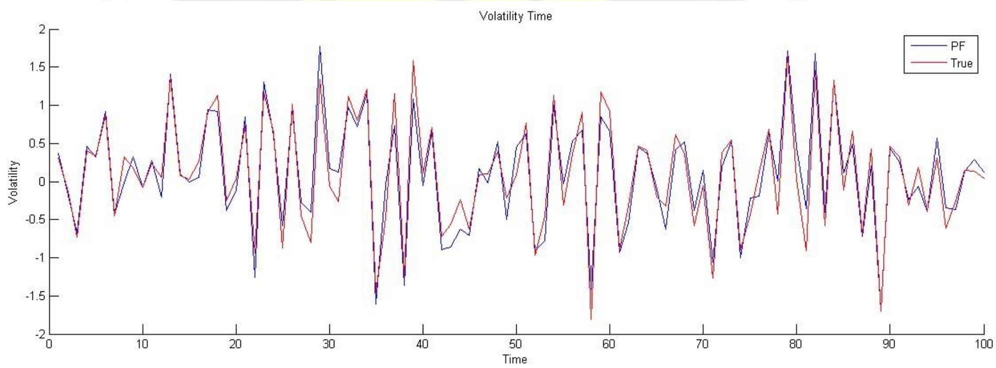

Figure 4.1: The estimated volatility at an underlying price of \$50 over time of 100 days

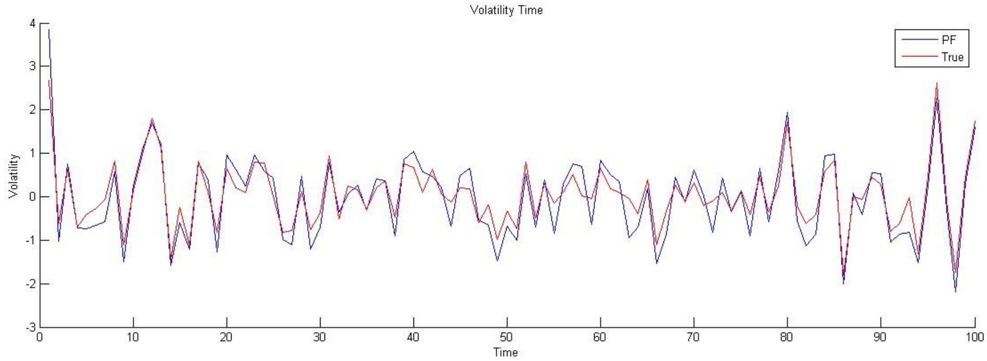

Figure 4.2: The estimated volatility at an underlying price of $60 over time of 100 days

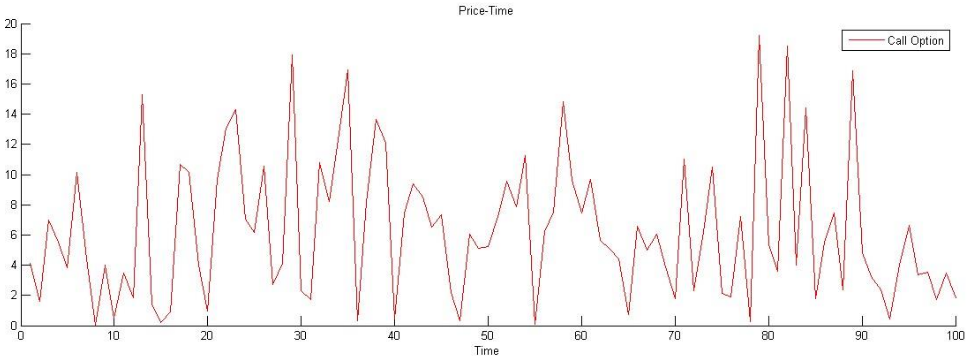

Figure 4.3: The estimated price at an underlying price of $50 over time of 100 days

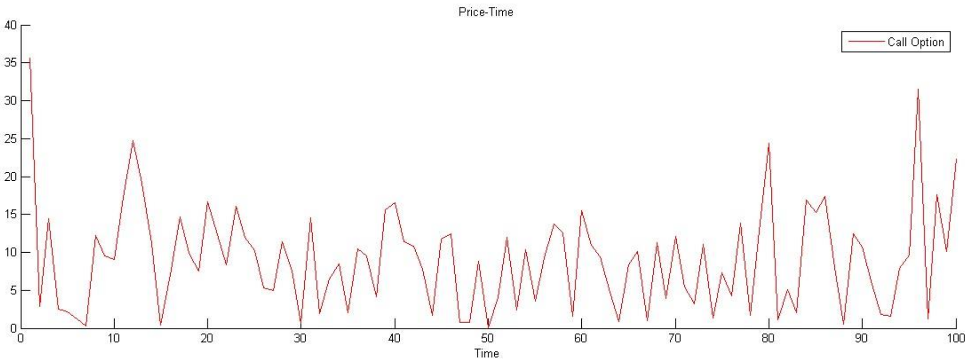

Figure 4.4: The estimated price at an underlying price of $60 over time of 100 days

Table 4.1: Estimated mean volatility and RMSE for a Call Option

Call Option		
Underlying Price	Mean PF	RMSE PF
50	0.0304	0.2157
60	0.0063	0.3440

Table 4.1 provides the root mean square error RSME showing the performance of the particle filter.

4.2.2 Put Option

The estimated volatility and price over time for a put option is displayed in figure 4.3 and 4.4 respectively. These estimates where obtained for an underlying stock price of $40 and $50 respectively at a strike price of $50 over a 100 days time period.

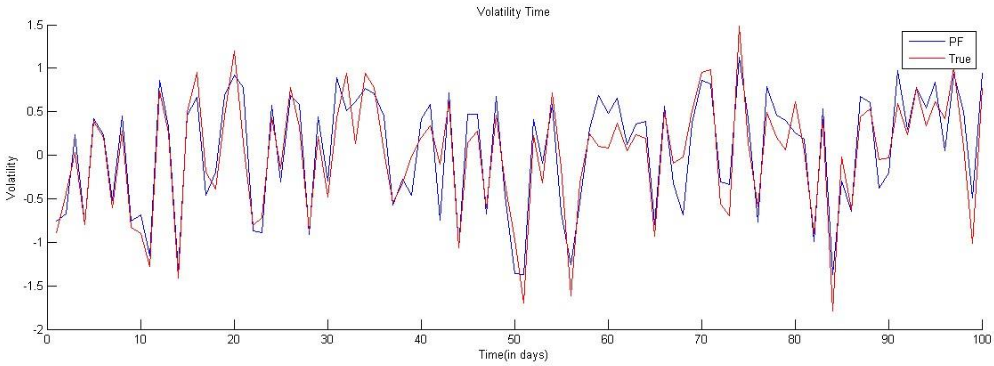

Figure 4.5: The estimated volatility at an underlying price of $40 over time of 100 days

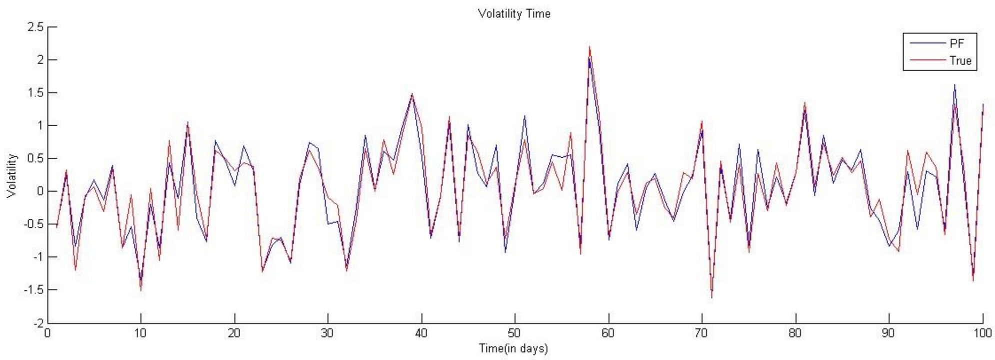

Figure 4.6: The estimated volatility at an underlying price of $50 over time of 100 days

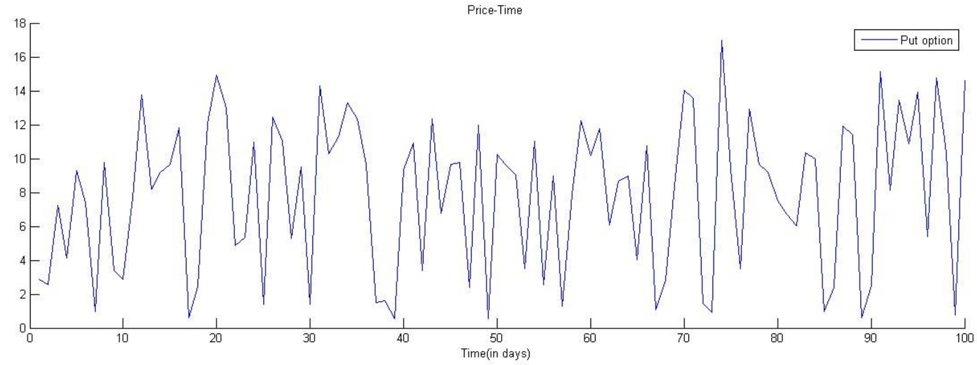

Figure 4.7: The estimated price at an underlying price of $40 over time of 100 days

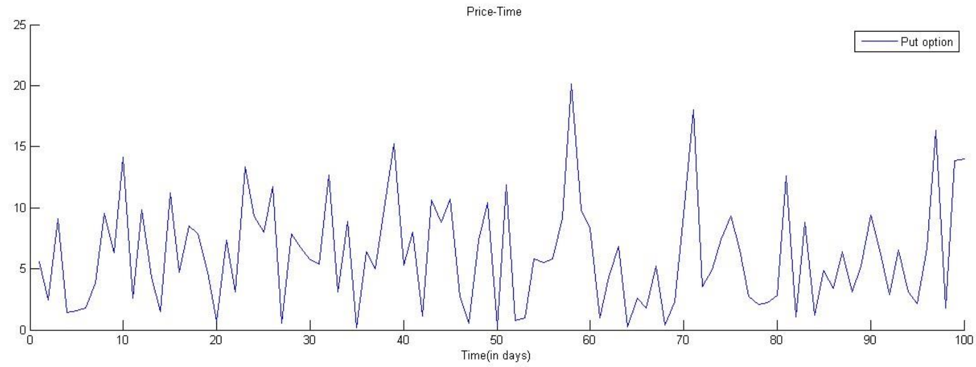

Figure 4.8: The estimated price at an underlying price of $50 over time of 100 days

Table 4.2: Estimated mean volatility and RMSE for a Put Option

Put Option		
Underlying Price	Mean PF	RMSE PF
40	0.0218	0.2823
50	0.0339	0.2156

Table 4.2 provides the estimated mean volatility and the root mean square error RSME showing the performance of the particle filter. Mean volatility decrease as the underlying price S decrease from the position where underlying price is equal to strike price K of $50. Whiles RMSE increase as underlying price decrease from $50 to $40.

Chapter 5

Conclusion

5.1 Introduction

This study has investigated the performance of the Sampling Importance Resampling(SIR) also known as the bootstrap filter by taking into consideration the effect of different process noise, different measurement noise, different number of particles. The bootstrap filter was implemented by use of synthetic data. A comparison test for when the options are profitable (in the money) was performed for both put and call options. This Chapter present conclusion and a number of recommendations.

5.2 Conclusion

Gorden et al. (1993) proposed the SIR filter. SIR filter was implemented under the following of scenarios:

- effect of different underlying price

- effect of different process noise on filter

- effect of different measurement noise of filter

- effect of different number of particles of SIR filter.

  At different process noise, SIR filter was found to perform better when $P_0 < 1$. In term being in-the-money for a call option, SIR filter still performed better when process noise $P_0 < 1$. At-the-money ($S = K$), SIR filter performed well when process noise was $P_0 < 1$ whiles performing poorly in-the-money ($S > K$) Generally, the SIR filter performed well at low measurement noise $M_0 < 1$. In-the-money, a measurement noise $M_0 > 1$ is required for a call option for a good performance of SIR filter. Also the performance of SIR filter is better at-the-money when compared with inthe-money for measurement noise. The performance of SIR filter improved as number of particles decreased. This is shown in the fact that the RMSE declined as the number of particles also declined. Once again, the SIR filter performs better at-the-money than in-the-money.

5.3 Recommendation

On the basis of our findings, it is recommended that the process noise, measurement noise and number of particles should be minimal. A performance comparison of Auxiliary SIR particle filter and Regularized particle filter in option pricing.

REFERENCES

Akashi, H. and Kumamoto, H. (1975). Construction of discrete-time nonlinear filter by monte carlo methods with variance-reducing techniques. *Journal of Systems and Control*, 19:211–221.

Barillec, R. (2009). Bayesian data assimilation.

Bertino, L., Evensen, G., and Wackernagel, H. (2003). Sequential data assimilation techniques in oceanography.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, Volume 81.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307 – 327.

Bouttier, F. and Courtier, P. (March 2009). *Data Assimilation concepts and methods*.

Brennan, M. and Schwartz, E. (May1977). The valuation of american put options. *The Journal of Finance*, 32:449–462.

Broadie, M. and Glasserman, P. (1997). Pricing american-style securities using simulation. *Journal of Economic Dynamics and Control*.

Dablemont, S., Bellegem, V., and Verleysen, M. (2009). Forcasting high and low of financial time series by paticle systems and kalman filters.

Detemple, J. and Tain, W. (July 2002). The valuation of american options for a class of diffusion processes. *Management Science*, 48:917–937.

Doucet, A. (1998). On sequential simulation-based methods for bayesian filtering. *technical report Cambridge University*.

Doucet, A., de Freitas, N., and Gordon, N. (2001). *Sequential Monte Carlo Methods in Practice*. Springer-Verlag.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasigeostrophic model us monte carlo methods to forcast error statistics. *Journal of Geophsical Research, Vol. 99*.

Evensen, G. (2003). *Data Assimilation: The Ensemble Kalman Filter: Theoretical Formulation and practical Implementation*. Springer-Verlag.

Gallant, A., Rossi, P., and Tauchen, G. (1992). Stock prices and volume. *Reveiw of Financial Studies*, 5:199–242.

Geske, R. and Johnson, H. (1984). The american put option valued analytically. *The Journal of Finance*, 39:1511–1524.

Gordon, N., Salmond, D., and Smith, A. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceedings F*, 140:107– 113.

Handschin, J. (1970). Monte carlo techniques for prediction and filtering of nonlinear stochastic processes. *Journal Automatica (Journal of IFAC)*, 6:555–563.

Harrison, J. and Pliska, S. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11:215– 260.

Heston, S., L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of Financial Studies*, 6:327–343.

Hull, John, C. (2006). *Options, Futures and Other Derivatives*. Pearson Prentice Hall.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42:281–300.

Jasra, A. and Del Moral, P. (26th May 2010). Sequential monte carlo methods for option pricing.

Jazwinski, A. (1970). *Stochastic Processes and Filtering Theory*. Academic Press.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problem. *Transactions of the ASME - - J ournal of Basic Engineering*, 82:35–45.

Kitagawa, G. (1996). Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 15:1– 25.

Kong, A., Liu, J., and Wong, W. (March 1994). Sequential imputations and bayesian missing data problems. *Journal of American Statistical Association*, 89:278–288.

Lahoz, W., Boris, K., and Menard, R. (2010). *Data Assimilation: Making sense of Observations*. Springer-Verlag Berlin Heidelberg.

Lindstrom, E. and Guo, J. (2013). Simultaneous calibration and quadratic hedging of options.

Maybeck, P. (1979). *Stochastic Models, Estimation, and Control*. Mathematics in Science and Engineering.

Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science*, pages 141–183.

Musso, C., Oudjane, N., and LeGland, F. (2001). *Improving Regularised Particle Filters*. Springer.

Norton, J. and Verse, G. (July 2002). Improvement of the particle filter by better choice of the predicted sample set. *IFAC World Congress b'02 Barcelona*.

Ott, E., Hunt, B., Szunyogh, I., Zimin, A., Kostelich, E., Corazza, M., Kalnay, E., Patil, D., and Yorke, J. (2004). A local ensemble kalman filter for atmospheric data assimilation.

Pastorello, S., Renault, E., and Touzi, N. (June 2000). Statistical inference for random-variance option pricing. *Journal of Business and Economic Statistics*, 18:358–367.

Rubin, D. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley and Sons.

Rubinstein, M. (1983). Displaced diffusion options pricing. *The Journal of Finance*, 34:213–217.

Smith, A. and Gelfand, A. (May 1992). Bayesian statistics without tears: A sampling -resampling perspective. *The American Statistician*, 46:84–88.

Stein, J. and Stein, E. (1991). Stock price distributions with stochastic volatility: An analytic approach. *Review of Financial Studies*, 4:727–752.

Torma, P. and Szepesvari, C. (2001). Ls-n-ips: An improvement of particle filters by means of local search. *IFAC Symposium on Nonlinear Control System*, pages 715–719.

Weisstein, E. (2002). *CRC Concise Encyclopedia of Mathematics 2nd ed*.