Path-dependent volatility and the preservation of PDEs

by

Michael Light

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae

in the Department of Mathematics and Applied Mathematics
in the Faculty of Natural and Agricultural Sciences

University of Pretoria
Pretoria

August 2016

Declaration

I, Michael Light declare that the dissertation, which I hereby submit for the degree *Magister Scientiae in Mathematics of Finance* at the University of Pretoria, is my own work and has not previously been submitted by me or any other person for a degree at this or any other tertiary institution.

Michael Light

Date

Abstract

The classical theory of risk neutral derivative pricing relies on the underlying market model being Markovian and complete. We present the theory of stochastic differential equations relevant to risk neutral pricing, with a particular focus on the Markov property and its links to partial differential equations. We demonstrate when this classical theory can still be applied to derivative pricing in models with path dependent volatility.

A link between these models and the local volatility framework is derived via the representation of local volatility as the conditional expectation of some, more complicated, process. Julien Guyon used this link as a tool in fitting a large class of models to the market. We will propose a fitted, complete and Markovian market model, which incorporates past asset levels in future volatility levels. The numerical implementation of such a model is addressed through a Monte Carlo scheme incorporating Guyon's particle method, as well as a finite difference scheme.

Acknowledgement

I would like to thank my supervisor Dr Van Zyl, whose guidance and encouragement has helped me grow as an aspiring mathematician. Thank you for answering my long technical emails with equally lengthy replies when direct communication was not possible. I would also like to acknowledge Prof Swart who was always willing to listen and comment on any ideas I had whenever I knocked on his door.

Lastly I would like to thank my parents. They have afforded me the honour and privilege of an education and have always supported me in pursuing my passion, mathematics.

Contents

1 Introduction

A paper on option pricing almost surely begins by mentioning Myron Scholes and Fischer Black's work in deriving the Black-Scholes option pricing framework [4], and since this dissertation is in this space, it is no exception.

Although the 1973 paper was a hallmark moment for modern finance, the Black-Scholes framework is not without criticism, least of all in its treatment of the volatility of the underlying, which is assumed constant.

The existence of so called "volatility smiles" is a clear violation of the assumption of constant volatility. Some of the more well studied alternatives to constant volatility include local and stochastic volatility models. In the local volatility framework much of the classical theory is retained, but the volatility dynamics are limited and remain unrealistic. At the opposite end, most stochastic volatility models give richer dynamics, but markets lose completeness and much of the classical theory collapses.

There is another class of models that, until recently, has received comparatively less attention to the two just mentioned. Path dependent volatility models have begun to attract more interest, and this is well deserved. They offer much of the desirable qualities of their counterparts, while sacrificing less. We will see that it is possible to have a complete model which incorporates past information into future volatility levels, while still retaining the Markov property, somewhat of a contradiction.

In this dissertation we will present the most important literature so far in this field, and look to unify some of the work into a single framework. The eventual goal will be to arrive at a fitted, complete and Markovian market model, for which we can derive a pricing PDE.

We will begin by discussing the underlying mathematical theory, before reviewing the relevant work done in the field of local volatility with Dupire's work in this field being of particular importance [14]. The link between local and path-dependent volatility, and the key to fitting our model to the market, is a classical theory by Gyöngy [22] which we will present and discuss.

We will then launch into the current path-dependent volatility theory before proposing our final model. Finally, we discuss numerical techniques for the application of our chosen model. In particular we consider a finite difference method for a pricing PDE, as well as a Monte Carlo scheme that makes use of the particle method of Julien Guyon [20].

2 Measure theory, stochastic calculus, and martingale pricing

In this section we lay the foundations for the rest of the dissertation. In an attempt to make the dissertation self contained, we will give a brief description of the mathematical framework under which we work, as well as present some key results that will be used in the work that follows.

2.1 The basics

A natural starting point is to discuss the space in which we work. All definitions and results can be viewed under the context of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the whole space, $\mathcal{F}$ is a σ-field made up of subsets of Ω and $\mathbb{P}$ is a probability measure, meaning it is a function $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ such that all the usual axioms for a measure are satisfied and in addition $\mathbb{P}(\Omega) = 1$.The σ-field can be interpreted as the set of events in the space Ω, and the function $\mathbb{P}$ is said to assign a probability to each of these events. We say that a function $f : X \to \mathbb{R}$ on the measure space $(\Omega, \mathcal{F}, \mathbb{P})$ is $\mathcal{F}$-measurable (or simply measurable if there is no possible ambiguity) if the set $\{f \in B\}$ is in $\mathcal{F}$ for every $B \in \mathcal{B}(\mathbb{R})$. Here we have used the shorthand $\{f \epsilon B\} := \{x \in X : f(x) \epsilon B\}$.

The stage is now set for us to define random variables. A random variable is simply a measurable function f mapping Ω to $\mathbb{R}$. Every random variable $f : X \to \mathbb{R}$ gives rise to a measure

$$\mu_f(B) := \mathbb{P}\big(\{f \epsilon B\}\big) \tag{2.1}$$

on $\mathbb{R}$ defined on the σ-field of Borel sets $B \epsilon \mathcal{B}(\mathbb{R})$. We call μ_f the distribution of f, and we define the function $F(x) := \mathbb{P}(f \leq x)$ as the cumulative distribution of f.

We say that a sequence of random variables $\{f_n\}$ converges to f *in measure* if for every $\epsilon > 0$ we have that

$$\lim_{n \to \infty} \mathbb{P}\Big(\omega : |f_n(\omega) - f(\omega)| > \epsilon\Big) = 0.$$

In probability theory the concepts of expected value and variance of a random variable are of particular importance. In order to define these two crucial quantities we must discuss integration with respect to a probability measure $\mathbb{P}$.

Consider a measurable simple function ϕ on Ω. This is a function that can be written in the form $\phi = \sum_{i=1}^{n} a_i \chi_{A_i}$, where n is some finite whole number, a_i are constants, A_i are measurable sets with $A_i = [x : \phi(x) =$

$a_i]$ and χ_{A_i} is the characteristic function of the set A_i. The characteristic function is defined for any set A by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}.$$

The integral of this type of function over Ω with respect to $\mathbb{P}$ is defined as

$$\int_\Omega \phi d\mathbb{P} = \sum_{i=1}^{n} a_i \mathbb{P}(A_i).$$

Using this definition we can define the integral of a non-negative random variable f as

$$\int_\Omega f d\mathbb{P} = \sup \left[\int_\Omega \phi d\mathbb{P} : \phi \leq f, \phi \text{ is a measure simple function} \right].$$

The integral of any measurable random variable, not necessarily non-negative, is then defined by subtracting the integral of the absolute value of the negative part of the function from the integral of the non-negative part. Over a set E, the integral is simply equal to the integral of the function $f\chi_E$.

A function is said to be integrable if the integral of its absolute value is finite. The usual properties, such as linearity and that the absolute value of the integral is less than the integral of the absolute value, hold in this setting.

We can now define the expected value of a random variable f, with respect to a probability measure $\mathbb{P}$, simply as

$$\mathbb{E}^{\mathbb{P}}(f) := \int_\Omega f d\mathbb{P}$$

and the variance as

$$\begin{aligned} var(f) &:= \int_\Omega \left(f - \mathbb{E}^{\mathbb{P}}(f)\right)^2 d\mathbb{P} \\ &= \mathbb{E}^{\mathbb{P}}(f^2) - (\mathbb{E}^{\mathbb{P}}(f))^2. \end{aligned}$$

Next we consider the key concept of conditional expectation. We will use the most general definition which involves conditioning on a σ-field, however this easily translates to conditioning on a set or a random variable by simply considering the σ-field generated by that set or random variable respectively. With that in mind the definition is as follows [8]:

Definition Let f be an integrable function on our usual probability space and let $\mathcal{G}$ be a σ-field contained in $\mathcal{F}$. Then the **conditional expectation** of f given $\mathcal{G}$ is defined as the random variable $\mathbb{E}^{\mathbb{P}}(f|\mathcal{G})$ such that

1. $\mathbb{E}^{\mathbb{P}}(f|\mathcal{G})$ is $\mathcal{G}$-measurable.

2. For any $A \in \mathcal{G}$
$$\int_A \mathbb{E}^{\mathbb{P}}(f|\mathcal{G})d\mathbb{P} = \int_A f d\mathbb{P}.$$

In addition to the expected properties, such as linearity etc, some less obvious properties that will be useful are

1. If a r.v f is $\mathcal{F}$-measurable then, for any integrable r.v g, $\mathbb{E}(fg|\mathcal{F}) = g\mathbb{E}(g|\mathcal{F})$.
2. If a r.v f is independent of $\mathcal{F}$ then $\mathbb{E}(f|\mathcal{F}) = \mathbb{E}(f)$.
3. If we have sigma fields $\mathcal{F}, \mathcal{F}_1, \mathcal{F}_2$ such that $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}$ then
$$\mathbb{E}\big(\mathbb{E}(f|\mathcal{F}_2)\big|\mathcal{F}_1\big) = \mathbb{E}(f|\mathcal{F}_1).$$

Now that most of the tools from measure theory have been defined we will state some key results from the subject that will be used in this dissertation. The first is a famous lemma from one of the main contributors to the subject.

Theorem 2.1 (Fatou's Lemma) *Let $\{f_n\}$ be a sequence of measurable functions, $f_n : \Omega \to \mathbb{R}^+$, and let $f = \liminf_{n\to\infty} f_n$. Then*
$$\mathbb{E}^{\mathbb{P}}(f) \leq \liminf_{n\to\infty} \mathbb{E}^{\mathbb{P}}(f_n).$$

Proof See Theorem 3 page 57 in [9].

This can be extended to conditional expectation in the natural way so that under the same conditions and for any σ-field $\mathcal{G}$ we have that
$$\mathbb{E}^{\mathbb{P}}(f|\mathcal{G}) \leq \liminf \mathbb{E}^{\mathbb{P}}(f_n|\mathcal{G}).$$

The next result sets the scene for a change of measure. This is important in the context of finance since we will often move from the risky to a risk neutral world, which are linked by a change of probability measure. First we must define the concept of absolute continuity. Suppose we have two measures $\mathbb{P}$ and $\mathbb{Q}$, we say that $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$, written $\mathbb{Q} << \mathbb{P}$, if for every $A \in \mathcal{F}$ such that $\mathbb{P}(A) = 0$ we have that $\mathbb{Q}(A) = 0$.We can now state the Radon-Nikodym Theorem.

Theorem 2.2 (Radon-Nikodym Theorem) *Let $(\Omega, \mathcal{F})$ be a measurable space, and $\mathbb{P}$ and $\mathbb{Q}$ finite measures on $(\Omega, \mathcal{F})$. If $\mathbb{Q} << \mathbb{P}$ then there exists a measurable function $f : \Omega \to \mathbb{R}^+$ such that for all subsets $A \in \mathcal{F}$*
$$\mathbb{Q}(A) = \int_A f d\mathbb{P}.$$

The function f is called the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ and is denoted by $\frac{d\mathbb{Q}}{d\mathbb{P}}$.

Proof See Theorem 5 on page 139 of [9].

If we have that $\mathbb{Q} << \mathbb{P}$ and $\mathbb{P} << \mathbb{Q}$, we say that the measures $\mathbb{P}$ and $\mathbb{Q}$ are equivalent. A trivial, but important, fact for probability measures is that

$$\mathbb{Q} << \mathbb{P} \Rightarrow \mathbb{Q}(A) = 1 \text{ where } A \in \mathcal{F} \text{ is such that } \mathbb{P}(A) = 1$$

this can be easily shown by considering A^c, which has measure zero, and the fact that $\mathbb{P}(\Omega) = \mathbb{Q}(\Omega) = 1$.

We have now defined most the key concepts and results we will need to dive into some important definitions and results concerning stochastic processes.

2.2 Stochastic processes

A stochastic process can be defined as a family of random variables $\big(X(t)\big)$ parametrized by $t \in T \subset \mathbb{R}$. We will often use the notation X to denote a stochastic process where possible. If $T = \{1, 2, 3 \dots\}$ then X is said to be a discrete-time random process, but when T is some interval in $\mathbb{R}$ (typically $[0, \infty)$) then X is said to be a continuous-time stochastic process.

We will consider processes over a time period $[0, T]$, and so we need to introduce a concept that captures the information flow over this interval. A filtration, denoted by $(\mathcal{F}_t)_{t\geq 0}$, is a family of σ-fields satisfying

$$\mathcal{F}_0 \subseteq \mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}_T, \text{ for every } s, t \in [0, T] \text{ such that } s < t.$$

For each $t \in [0, T]$ the σ-field $\mathcal{F}_t$ can simply be interpreted as the information available at time t.

We need to extend the concept of measurability to stochastic processes. A process X is said to be adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$, if $X(t)$ is $\mathcal{F}_t$-measurable for every $t \in T$.

We call a process X a *martingale* with respect to $\mathcal{F}_t$ and a measure $\mathbb{P}$ if

1. $\mathbb{E}^{\mathbb{P}}\big(|X_t|\big) < \infty$ for all $t \geq 0$
2. $\mathbb{E}^{\mathbb{P}}\big(X_t|\mathcal{F}_s\big) = X_s$ for all $t \geq s \geq 0$

If we replace equality with $\leq$ $(\geq)$ then we say that X is a *supermartingale* (*submartingale*). An important result from the study of martingales was due to Joseph Doob.

Theorem 2.3 (Doob's Martingale inequality) *Let X be a continuous submartingale taking non-negative real values, either in discrete or continuous time. Then, for any constant $C > 0$ and $p \geq 1$,*

$$\mathbb{P}\left(\sup_{u \in [0,T]} X_u \geq C \right) \leq \frac{\mathbb{E}(X_T^p)}{C^p}.$$

Proof See Proposition 4.1 on page 68 of [8].

Obviously this inequality also holds when X is a martingale.

Another important concept will be that of the σ-field generated by a random variable f. This is the smallest σ-field containing all sets of the form $\{f \in B\}$ where $B \in \mathcal{B}(\mathbb{R})$. This can be naturally extended to the concept of a filtration generated by a random process. A deep and useful result relating to this concept is presented next

Theorem 2.4 (Doob-Dynkin) *Let f be a random variable. Then each $\sigma(f)$-measurable random variable g can be written as*

$$g = \phi(f)$$

for some Borel function function $\phi : \mathbb{R} \to \mathbb{R}$.

Proof see proposition 3 on page 8 of [36].

A *stopping time* with respect to a filtration $\mathcal{F}_t$ and an index set T, is a random variable τ taking values in T such that for each $t \in T$ the set

$$\{\tau \leq t\} \in \mathcal{F}_t.$$

We will often be interested in a process up to some unknown future time, for this reason for every process X and stopping time τ we define a new process by $X_\tau(t) := X(t \wedge \tau)$, referred to as X stopped at τ.

There is a class of processes, called local martingales, which will be of particular importance for our purposes. They can be defined as follows.

Definition Let X be an $\mathcal{F}_t$ adapted stochastic process. Then X is called an $\mathcal{F}_t$-**local martingale** if there exists a sequence of $\mathcal{F}_t$-stopping times (s_n) such that

1. the sequence is almost surely increasing
2. $\mathbb{P}(s_k \to \infty \text{ as } k \to \infty) = 1$
3. the stopped process
$$X_t^k := X(s_k \wedge t)$$
is an $(\mathcal{F}_t)$-martingale for every k.

Clearly all martingales are local martingales, but the converse is not necessarily true. A typical example of a local martingale which is not a martingale, is the reciprocal of the norm of Brownian Motion in $\mathbb{R}^3$ (see example 2 on page 37 of [35]). We will often define a new measure via a Radon-Nikodym derivative, which is a local martingale. However we will need the process to be a true martingale in order for our new measure to have the desired properties. Therefore, it is important for us to understand when a local martingale is a true martingale.

We see that the definition only requires the existence of a sequence of stopping times, and does not say anything about its form. We will be particularly interested in the following sequence of random times

$$\tau_n = \inf\{t \geq 0, |X_t| > n\}$$

for any process X. That τ_n is a stopping time follows by noticing that, for any $t \geq 0$,

$$\begin{aligned}\{\tau_n \leq t\} &= \{\omega \in \Omega : X_s(\omega) \geq n \text{ for some } s \in [0,t]\} \\ &= \bigcup_{s\in[0,t]} \{X_s \geq n\} \in \mathcal{F}_t.\end{aligned}$$

The following result shows that a local martingale stopped using the above sequence of stopping times, is a martingale. We will need the following lemma (see, for example, [3]).

Lemma 2.5 *Let X_t be a martingale on $(\Omega, \mathcal{F}_n, \mathbb{P})$ and τ be a stopping time. Then X_t^τ is a true martingale, where $X_t^\tau := X(t \wedge \tau)$ is the stopped version of X_t.*

Proof We only prove the lemma for a discrete time local martingale X_n. Define the discrete time process h by $h_n := \chi_{\{n\leq\tau\}}$ $n = 1, 2, \ldots$. Clearly

$$\{n \leq \tau\} = \{n < \tau\}^c = \{n - 1 \leq \tau\}^c \in \mathcal{F}_{n-1}.$$

Which shows that h_n is $\mathcal{F}_{n-1}$ measurable or *predictable*. Clearly we also have that

$$X_n^\tau = \sum_{k=0}^{n} h_k(X_k - X_{k-1})$$

which yields the following

$$
\begin{aligned}
\mathbb{E}(X_n^\tau|\mathcal{F}_{n-1}) &= \mathbb{E}\big(\sum_{k=0}^{n} h_k(X_k - X_{k-1})\big) \\
&= \mathbb{E}\big(h_n(X_n - X_{n-1}) + \sum_{k=0}^{n-1} h_k(X_k - X_{k-1})|\mathcal{F}_{n-1}\big) \\
&= \mathbb{E}(h_n X_n|\mathcal{F}_{n-1}) - \mathbb{E}(h_n X_{n-1}|\mathcal{F}_{n-1}) + \sum_{k=0}^{n-1} \mathbb{E}(h_k(X_k - X_{k-1})|\mathcal{F}_{n-1}) \\
&= h_n X_{n-1} - h_n X_{n-1} + \sum_{k=0}^{n-1} h_k(X_k - X_{k-1}) = X_{n-1}^\tau
\end{aligned}
$$

which shows that X_n^τ is a martingale. ∎

Proposition 2.6 *Let X_t be a continuous local martingale on $(\Omega, (\mathcal{F}_t), \mathcal{F}, \mathbb{P})$ such that $X_0 = 1$. Let*

$$\tau_n = \inf\{t \geq 0, |X_t| > n\}.$$

Then, for $n \in \mathbb{N}$ the process $X_t^n := X(\tau_n \wedge t)$ is a bounded martingale.

Proof Let (s_k) be the sequence of stopping times described in the definition of a local martingale, so that $X_t^k := X(t \wedge s_k)$ is a true martingale. Then for any $s, t \in [0, T]$, such that $s \leq t$, and for any $k, n \geq 0$ we have that

$$\mathbb{E}(X(t \wedge s_k \wedge \tau_n)|\mathcal{F}_s) = X(s \wedge s_k \wedge \tau_n).$$

Since X_t^k is a martingale, and τ is a stopping time, so that by our lemma, the process stopped by τ_n is a martingale. Taking the limit as $k \to \infty$ on both sides we get that

$$\mathbb{E}(X(t \wedge \tau_n)|\mathcal{F}_s) = X(s \wedge \tau_n).$$

Notice that we were able to take the limit inside the expectation since, by the construction of τ_n, the stopped process $X(t \wedge s_k \wedge \tau_n)$ is bounded by n. Therefore, all the conditions of Lebesgue's dominated convergence Theorem hold (See, for example, Theorem 10 on page 63 of [9]), and we can take the limit inside the integral.This completes the proof. ∎

2.3 Stochastic calculus

The most important process that we will encounter is the Wiener process, or alternatively, Brownian motion. This is a process W that starts at 0 almost surely (with probability 1), which has almost surely continuous sample paths and finally for any finite sequence of times $0 < t_1 < \cdots < t_n$ and Borel sets $B_1, \ldots, B_n \subset \mathbb{R}$ we have that

$$\mathbb{P}\big(\{W(t_1) \in B_1, \ldots, W(t_n) \in B_n\big)$$
$$= \int_{B_1} \cdots \int_{B_n} p(t_1, 0, x_1)p(t_2 - t_1, x_1, x_2) \ldots p(t_n - t_{n-1}, x_{n-1}, x_n)dx_n \ldots dx_1,$$

where $p(t, x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}}$ is called the transition density of W. The Wiener process also happens to be a *martingale.*

We are now ready to define the stochastic integral, which is the essential ingredient in the discussion of stochastic differential equations. We start by defining the stochastic integral for the class of piecewise constant processes, known as step processes. Let g be such a process, the integral over $[a, b]$ is defined as

$$\int_a^b g(s)dW(s) := \sum_{k=0}^{n-1} g(t_k)\big[W(t_{k+1}) - W(t_k)\big]$$

where the set $\{t_k; k = 0, \ldots n\}$ is a partition so that $a = t_0 \leq t_1 \leq \cdots \leq t_k = b$. It is important to notice that in the definition we use the value of g at the lower bound of each interval, unlike in the Riemann case, this choice will effect the properties and value of a stochastic integral.

We can extend this to general random variables using the definition above and the concept of L^2 convergence.

Definition The space $L^2(\Omega, \mathcal{F}, \mathbb{P})$, is the space of equivalence classes of square integrable measurable functions with respect to the measure $\mathbb{Q}$. i.e for any $f \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ we have that

$$\int |f|^2 d\mathbb{P} < \infty.$$

We will write $L^2(\Omega, \mathcal{F}, \mathbb{P}) = L^2(\mathbb{P})$ when no ambiguity is possible. It can be shown that the function $||f - g||_2 = \left(\int |f - g|^2 d\mathbb{P}\right)^{1/2}$, defines a norm so that the pair $(L^2, ||\cdot||_2)$ is a norm space.

In addition, it can be shown that convergence in L^2 implies convergence in measure. For the proof of this see, for example, Theorem 5 on page 123 of [9].

Using this, for any process f we set up a sequence of step processes f_n that converge to f in the sense that

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^\infty |f(t) - f_n(t)|^2 dt\right) = 0.$$

We then define the stochastic integral of f as the random variable,

$$\int_a^b f(s)dW(s) = \lim_{n\to\infty} \int_a^b f_n(s)dW(S)$$

it is possible to prove that the right hand side converges to a random variable in L^2, and it can be shown that the result does not depend on our choice of an approximating sequence of functions.

We are now in a position to discuss stochastic differentials. As is standard, we work with a process X in the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Assume that there exists a real number a, and two adapted processes μ and σ such that we have

$$X(t) = a + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s), \text{ for all } t \geq 0$$

then for ease of notation we say that X has stochastic differential

$$dX(t) = \mu(t)dt + \sigma(t)dW(t)$$

with initial condition

$$X(t) = a.$$

A stochastic process X is called an Itô process if it has almost surely continuous paths and can be represented as

$$X(t) = a + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s), \text{ for all } t \geq 0$$

where $\sigma(s)$, $\mu(s)$ are adapted processes such that

$$\mathbb{E}(\int_0^t |\sigma(s)|^2 ds) < \infty$$

and

$$\int_0^t |\mu(s)|ds < \infty, \ a.s$$

for all $t > 0$. If these properties hold we say that σ and μ are in M_t^2 and $\mathcal{L}_t^1$ respectively.

With this notation we can present the workhorse of stochastic calculus, Itô's formula.

Theorem 2.7 (Itô's formula) *Suppose the process X is an Itô process, and let f be a $C^{1,2}$-function (first and second order derivatives are continuous). Define the process $Z(t) := f(t, X(t))$. Then Z has stochastic differential given by*

$$df(t, X(t)) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX)^2$$

where we use the following formal multiplication table

$$(dt)^2 = 0$$

$$dt \cdot dW = 0$$

$$(dW)^2 = dt.$$

Proof See the proof of Theorem 7.5 on page 196 of [8].

This theorem extends to the case of a multidimensional process in a natural way. Given a process, Itô's formula allows us to calculate its stochastic differential. Often we will be interested in the reverse of this by asking the question: given a stochastic differential, can we find a stochastic process that satisfies it?

2.4 Stochastic Differential Equations

Throughout this subsection we will consider the d-dimensional stochastic differential equation (SDE)

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t) \text{ on } t \in [0, T] \tag{2.2}$$

with initial value $X(0) = x_0$. Here $W(t) = [W_1(t), \ldots, W_m(t)]^T$ is an m-dimensional Wiener process and x_0 is a $\mathcal{F}_0$-measurable, $\mathbb{R}^d$ valued random variable with finite variance. The functions $\mu : \mathbb{R}^d \times [0, T] \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times [0, T] \to \mathbb{R}^{d \times m}$ are both Borel measurable. We define the solution of this problem as follows:

Definition An $\mathbb{R}^d$-valued stochastic process X on $t \in [0, T]$ is called a solution to (2.2) if it has the following properties:

1. X has a.s continuous paths and is $\mathcal{F}_t$ adapted.
2. $\mu(t, X(t)) \in \mathcal{L}_t^1$ and $\sigma(t, X(t)) \in M_t^2$
3. X satisfies (2.2) for every $t \in [0, T]$

A solution X is called unique if

$$\mathbb{P}(X(t) = Y(t) \text{ for all } t \in [0, T]) = 1$$

for any other solution Y. Putting the definition more simply: when finding a solution to an SDE we are looking for an *Itô process* which has the correct stochastic differential.

It will be important to have conditions under which there exists a unique solution to (2.2). The following results and their proofs can be found in, for example, [34].

Theorem 2.8 *Assume that there exists positive constants C_1,C_2,C_3 and C_4 such that*

1. *(Lipschitz condition) for all $x, y \in \mathbb{R}^d$ and $t \in [0, T]$*

$$|\mu(t,x) - \mu(t,y)|^2 \leq C_1|x-y|^2 \tag{2.3}$$

$$|\sigma(t,x) - \sigma(t,y)|^2 \leq C_2|x-y|^2 \tag{2.4}$$

2. *(Linear growth condition) for all $(t, x) \in [0, T] \times \mathbb{R}^d$,*

$$|\mu(t,x)|^2 \leq C_3(1+|x|) \tag{2.5}$$

$$|\sigma(t,x)|^2 \leq C_4(1+|x|). \tag{2.6}$$

The there exists a unique solution process X to equation (2.2).

Proof See Theorem 3.1 on page 69 of [34].

An essential property for our purposes will be the so-called *Markov property* of a stochastic process. Simply put, if a process has the Markov property it has no memory. When considering possible future states of the process the only information that needs to be considered is the current position. More formally we can define the Markov property as follows.

Definition A d-dimensional $\mathcal{F}_t$-adapted process $X(t)$ is called a *Markov Process* if: for all $0 \leq s \leq t < \infty$ and $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\mathbb{P}(X(t) \in A|\mathcal{F}_s) = \mathbb{P}(X(t) \in A|X(s)). \tag{2.7}$$

We follow [34] and define the *transition probability* of a Markov process $X(t)$ as the function $P(s, t, x, A)$ defined on $0 \leq s \leq t < \infty$, $x \in \mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d)$, with the properties:

1. For every $0 \leq s \leq t < \infty$ and $A \in \mathcal{B}(\mathbb{R}^d)$,

$$P(s, t, X(s), A) = \mathbb{P}(X(t) \in A|X(s)).$$

2. For every $0 \leq s \leq t < \infty$ and $x \in \mathbb{R}^d$ $P(s, t, x, \cdot)$ is a probability measure on $\mathcal{B}(\mathbb{R}^d)$.

3. $P(s, t, \cdot, A)$ is Borel measurable for every $0 \leq s \leq t < \infty$ and $A \in \mathcal{B}(\mathbb{R}^d)$

4. For every $0 \leq s \leq t < \infty$, $x \in \mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d)$ it holds that

$$P(s, t, x, A) = \int_{\mathbb{R}^d} P(r, t, y, A)P(s, t, dy, A).$$

This is known as the Chapman-Kolmogorov equation.

Using this notation we can write the Markov property in (2.7) as

$$\mathbb{P}(X(t) \in A|\mathcal{F}_s) = P(s, t, X(s), A).$$

Making use of the above definition and notation we can formulate an elegant result concerning the solution on an SDE.

Theorem 2.9 *Let $X(t)$ be a solution of the equation*

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t) \text{ on } 0 \leq t$$

whose coefficients satisfy the conditions of Theorem 2.8, for some initial condition $X_0 = x \in \mathbb{R}$. Then X is a Markov process with respect to the filtration generated by W, whose transition probability is defined by

$$P(s, t, x, A) = \mathbb{P}(X_{x,s}(t) \in A)$$

where $X_{x,s}(t)$ is the solution of the equation

$$X_{x,s}(t) = x + \int_s^t \mu(u, X_{x,s}(u))du + \int_s^t \sigma(u, X_{x,s}(u))dW(u) \text{ on } t \geq s.$$

Proof For any $t > 0$, by construction we have that

$$\begin{aligned}
X(t) &= x + \int_0^t \mu(t, X_t)dt + \int_0^t \sigma(t, X(t))dW(t) \\
&= x + \int_0^s \mu(t, X_t)dt + \int_0^s \sigma(t, X(t))dW(t) + \int_s^t \mu(t, X_t)dt + \int_s^t \sigma(t, X(t))dW(t) \\
&= X(s) + \int_s^t \mu(t, X_t)dt + \int_s^t \sigma(t, X(t))dW(t) \\
&= X_{X(s),s}(t).
\end{aligned}$$

Since X and $X_{X(s),s}$ satisfy the same SDE, and by assumption we have uniqueness, we must have that $X_{X(s),s}(t) = X(t)$ almost surely. Clearly the process $X_{X(s),s}$ is determined by increments of the form $W(t) - W(s)$ and therefore is independent of $\mathcal{F}(s)$. We also have that $X(s)$ is $\mathcal{F}_s$ measurable.

Using this and the properties of conditional expectation gives

$$\mathbb{E}\big(f(X_t)|\mathcal{F}_s\big) = \mathbb{E}\big(f(X_{X(s),s}(t)|\mathcal{F}_s\big) = \mathbb{E}\big(f(X_{X(s),s}(t))\big|_{x=X(s)}.$$

∎

Next we would like to study when the solution of an SDE has a density function, and if we can derive some useful properties of this density assuming that it exists, and for this we will need the *classical Hörmander's Theorem.*

We present this result as seen in [24]. We will consider an n dimensional stochastic differential equation of the form

$$dX(t) = V_0(X(t))dt + \sum_{i=1}^{m} V_i(X(t)) \circ dW_i, \tag{2.8}$$

where $V_0, \ldots, V_m$ are smooth vector fields on $\mathbb{R}^n$, i.e $V_0, \ldots, V_m$ are vectors who's components are functions taking values in $\mathbb{R}$. $W_1, \ldots, W_m$ are all independent standard n dimensional Wiener processes and $\circ$ is the usual dot product. We assume throughout that all theses vectors fields satisfy the conditions necessary to guarantee that a solution of (2.8) exists, and that derivatives of all orders exist.

The Lie bracket between two vector fields U and V on $\mathbb{R}^n$ denoted by $[U, V]$ is the vector field defined by

$$[U, V](x) = DV(x)U(x) - DU(x)V(x),$$

where D is the derivative operator such that $(DU)_{i,j} = \frac{\partial U_i}{\partial x_j}$. For any SDE of the form of (2.8) we can then define a collection of vector fields $\mathcal{V}_k$ by the recursive relationship

$$\mathcal{V}_0 = \{V_i : i > 0\}, \ \mathcal{V}_{k+1} = \mathcal{V}_k \cup \{[U, V_j] : U \in \mathcal{V}_k \text{ and } j \geq 0\}$$

and we define the vector spaces $\mathcal{V}_k(x)$ by

$$\mathcal{V}_k(x) = span\{V(x) : V \in \mathcal{V}_k\}.$$

Using this notation we can state the seminal result of Hörmander.

Theorem 2.10 (Hörmander's Theorem) *If a stochastic differential equation of the form of (2.8) satisfies the parabolic Hörmander condition, i.e we have that*

$$\cup_{k\geq 1}\mathcal{V}_k(x) = \mathbb{R}^n \text{ for every } x \in \mathbb{R}^n,$$

then its solutions admits a smooth density with respect to the Lebesgue measure.

Proof See [24].

This Theorem tells us that, if the coefficients of a SDE are such that the parabolic Hörmander condition is satisfied, the solution to the SDE has a density function and that density function is smooth.

We will also be interested in whether or not the density, given it exists, is strictly positive. In other words, under what conditions can we say that a solution to an SDE can achieve any value with positive probability. The following proposition will be useful for the special cases in mind.

Proposition 2.11 *If the process X is of the form*

$$X_t = \int_0^t \sigma(\xi_u) dW_u$$

where ξ is a nonexplosive stochastic process with a transition density function f, and $\sigma(\cdot) > \epsilon$ for some $\epsilon > 0$, then X_t has a strictly positive density on $\mathbb{R}$.

Proof For any $t > 0$, let $\{t_j^n\}$ be a partition of the interval $[0, t]$ such that $0 = t_1^n \leq t_2^n \leq \cdots \leq t_n^n = t$, and $t_{i+1}^n - t_i^n = \frac{t}{n}$ for every i. Consider the random variable

$$X_t^n := \sum_{i=1}^n \sigma_i \Delta W_i$$

where $\sigma_i := \sigma(\xi(t_i^n))$ and $\Delta W_i = W(t_{i+1}^n) - W(t_i^n)$ for every $i = 1, \ldots, n$. Now, let $b > 0$ be an arbitrary real value then

$$\begin{aligned}
\mathbb{P}(|X_t^n| > b) &= \mathbb{P}\Big(\Big|\sum_{i=1}^n \sigma_i \Delta W_i\Big| > b\Big) \\
&= \mathbb{E}\Big(\mathbb{P}\Big(\big|\sum_{i=1}^n \sigma_i \Delta W_i\big| > b | \sigma_i = x_i, \text{ for } i = 1, \ldots, n\Big)\Big) \\
&= \int_\epsilon^\infty \ldots \int_\epsilon^\infty \mathbb{P}\Big(\big|\sum_{i=1}^n \sigma_i \Delta W_i\big| > b | \sigma_i = x_i, \text{ for } i = 1, \ldots, n\Big) f_\sigma(x_1, \ldots, x_n) dx_1 \ldots dx_n \\
&= \int_\epsilon^\infty \ldots \int_\epsilon^\infty \mathbb{P}\Big(\big|\sum_{i=1}^n x_i \Delta W_i\big| > b\Big) f_\sigma(x_1, \ldots, x_n) dx_1 \ldots dx_n \\
&= \int_\epsilon^\infty \ldots \int_\epsilon^\infty \mathbb{P}\Big(|Z| > b\Big) f_\sigma(x_1, \ldots, x_n) dx_1 \ldots dx_n, \text{ where } Z \sim N(0, \frac{1}{n}\sum_{i=1}^n x_i^2) \\
&\geq \int_\epsilon^\infty \ldots \int_\epsilon^\infty \mathbb{P}\Big(|Z_\epsilon| > b\Big) f_\sigma(x_1, \ldots, x_n) dx_1 \ldots dx_n, \text{ where } Z \sim N(0, \epsilon^2) \\
&= \mathbb{P}(|Z_\epsilon| > b) \\
&= \mathbb{P}\Big(\big|\frac{Z_\epsilon}{\epsilon}\big| > \frac{b}{\epsilon}\Big) \\
&= 2\Phi\Big(-\frac{b}{\epsilon}\Big) > 0 \qquad (2.9)
\end{aligned}$$

where $\Phi(\cdot)$ is the standard normal distribution. By the construction of X_t^n we know that

$$X_t^n \xrightarrow{L^2} X_t$$

and since convergence in mean implies convergence in measure taking limits on either side of

$$\mathbb{P}(|X_t^n| > b) > 0$$

we have that

$$\mathbb{P}(|X_t| > b) > 0.$$

Since this argument holds for any $b > 0$ we have shown that X achieves all values in $\mathbb{R}$ with positive probability. ∎

Notice that this argument will also hold if X is of the form

$$X_t = x + \int_0^t \sigma(\xi_u) dW_u$$

for any $x \in \mathbb{R}$, i.e

$$dX_t = \sigma(\xi_t) dW_t.$$

We have shown that if X can be written as the stochastic integral of a strictly positive and integrable random variable, then X has a strictly positive density on the whole of $\mathbb{R}$, provided that the density exists.

2.5 Stochastic functional differential equations

We now come to the subject of Stochastic functional differential equations (SFDEs), the theory of which is far less developed then that of ordinary SDEs. Throughout we will be working with a d-dimensional equation of the form:

$$dX(t) = \mu(t, (X(u), t-\tau \leq u \leq t))dt + \sigma(t, (X(u), t-\tau \leq u \leq t))dW(t) \; t \in [0, T], \tag{2.10}$$

where $\tau > 0$ is fixed. Clearly at time t, we have that the drift and volatility functions do not only depend on X at time t as in (2.2), but instead depend on X over the whole interval $[t-\tau, t]$. As in the case of the ordinary SDE we need to ask ourselves how the solution will be defined and under what condition will that solution exist? Again we take definitions and results as found in [34]. For any fixed $\tau > 0$ we define $C\big([-\tau, 0]; \mathbb{R}^d\big)$ as the space of continuous functions $\phi : [-\tau, 0] \to \mathbb{R}^d$ with the norm $||\phi|| := \sup_{\tau \leq u \leq 0} |\phi(u)|$.

We say that the stochastic functional differential equation has initial data $X_0 = \psi(u)$ for $-\tau \leq u \leq 0$ where ψ is an $\mathcal{F}_0$-measurable, $C\big([-\tau, 0]; \mathbb{R}^d\big)$ valued random variable such that $\mathbb{E}(||\psi||^2) < \infty$.

We can now define the solution as follows

Definition An $\mathbb{R}^d$- valued stochastic process X on $-\tau \leq t \leq T$ is called a solution to equation (2.10), with initial data ψ as above, if it has the following properties:

1. It is continuous and X is $\mathcal{F}_t$-adapted

2. $\mu(t, (X(u), t-\tau \leq u \leq t)) \in \mathcal{L}^1([0,T];\mathbb{R}^d)$ and
$\sigma(t, (X(u), t-\tau \leq u \leq t)) \in L^2([0,T];\mathbb{R}^{d\times m})$
3. $X_0 = \psi$ and, for every $0 \leq t \leq T$,

$$X(t) = \psi(0) + \int_0^t \mu(s, (X(s), t-\tau \leq s \leq t))ds + \int_0^t \sigma(s, (X(s), t-\tau \leq s \leq t))dB(s),\ a.s$$

A solution X is called unique if any other solution Y is such that

$$\mathbb{P}(\{X(t) = Y(t) \text{ for all } -\tau \leq t \leq T\}) = 1.$$

As in the case of ordinary stochastic differential equations we then have some results regarding the existence and uniqueness of solutions.

Theorem 2.12 *Assume that there exists positive constants C_1, C_2, C_3 and C_4 such that*

1. *(uniform Lipschitz condition) for all $\phi, \xi \in C([-\tau, 0];\mathbb{R}^d)$ and $t \in [0,T]$*

$$|\mu(t,\phi) - \mu(t,\xi)|^2 \leq C_1||\phi - \xi||^2$$
$$|\sigma(t,\phi) - \sigma(t,\xi)|^2 \leq C_2||\phi - \xi||^2.$$

2. *(linear growth condition)for all $(t,\phi) \in [0,T] \times C([-\tau,0];\mathbb{R}^d)$*

$$|\mu(t,\phi)|^2 \leq C_3(1 + ||\phi||^2)$$
$$|\sigma(t,\phi)|^2 \leq C_4(1 + ||\phi||^2).$$

Then there exists a unique solution X to equation (2.10) with initial data $X_0 = \psi(u)$ for $-\tau \leq u \leq 0$.

Proof See the proof of Theorem 2.2 on page 150 of [34].

As one might expect there is no result for stochastic functional differential equations that is equivalent to the one concerning the Markov property of an ordinary stochastic differential equations. In fact in most cases the solution to equation (2.10) will be non-Markovian.

A typical example, and one we will encounter in later chapters, is that of the stochastic delay equation. Which takes the general form

$$dX(t) = F(t, X(t), X(t-\tau))dt + G(t, X(t), X(t-\tau))dW(t)$$

on $t \in [0,T]$ with initial data ψ as above, where $F, G : \mathbb{R} \times \mathbb{R} \times [0,T] \to \mathbb{R}$. A less trivial example, and one that may depend on the entire path of the r.v, is

$$dX(t) = F(t, X(t))dt + \sup_{s\in[a,t]} |X(s)|G(t, X(t))dW(t).$$

Here a is some nonnegative constant, and $F, G : \mathbb{R} \times [0,T] \to \mathbb{R}$.

We will now present several important results for the application of stochastic calculus to finance.

2.6 Martingale pricing

We start with the martingale representation Theorem. This Theorem guarantees that every martingale adapted to the filtration generated by the Wiener process can be represented by a stochastic integral of another adapted process.

Theorem 2.13 (The Martingale Representation Theorem) *Let W be a Wiener process, and assume that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is defined as*

$$\mathcal{F}_t = \sigma(W_s, s \leq t),\ t \in [0,T],$$

and let M be any $\mathcal{F}_t$-adapted martingale. Then there exists a uniquely determined $\mathcal{F}_t$-adapted process ξ such that M has the representation

$$M(t) = M(0) + \int_0^t \xi(u)dW(u),\ t \in [0,T]. \tag{2.11}$$

If the martingale M is in $\mathcal{L}^2$ then ξ is in $\mathcal{L}^2$.

Proof See, for example, Theorem 11.2 in [3].

This result will be useful in the quest of proving that certain market models are complete. The next Theorem is one that allows us to derive the dynamics of stochastic processes, driven by Wiener processes, under a change of measure.

Theorem 2.14 (Girsanov) *Let $W(t)$, $t \in [0,T]$, be a standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. Let h be any adapted process such that $\int_0^T |h(u)|du < \infty$ a.s, and let*

$$L_t = e^{\int_0^t h(u)dW(u) - \frac{1}{2}\int_0^t |h(u)|^2 du},\ t \in [0,T],$$

and suppose that $\mathbb{E}^{\mathbb{P}}(L_t) = 1$. If we define a probability measure $\mathbb{Q}$ on the measurable space $(\Omega, \mathcal{F})$ by $d\mathbb{Q} = L_t d\mathbb{P}$, then the process

$$\bar{W}(t) := W(t) - \int_0^t h(u)du,\ t \in [0,T], \tag{2.12}$$

is a standard Brownian motion under the measure $\mathbb{Q}$. We call h_t the Girsanov Kernel of the measure transformation.

Proof See, for example, Theorem 11.3 in [3].

Since we will be primarily interest in the risk neutral world, this will be an extremely important tool for us to have control over the change to the risk neutral measure. In the conditions for the Girsanov Theorem we require $\mathbb{E}^{\mathbb{P}}(L_T) = 1$. Therefore, it will be useful to have a condition for L_t, or equivalently $h(t)$, under which this holds. The standard result to use in this regard is Novikov's condition.

Theorem 2.15 (Novikov's Condition) *Assume that the Girsanov Kernel h_t is such that*

$$\mathbb{E}^P\left[e^{\frac{1}{2}\int_0^T ||h_t||^2 dt}\right] < \infty. \tag{2.13}$$

Then L, as defined in the previous Theorem, is a martingale and in particular $\mathbb{E}^P[L_t] = 1$.

However this can be difficult to verify in some cases. When our measures are related through two stochastic differential equations, there is a more useful condition. Before we look at this special case, let us consider the following general setting:

Let $(\Omega, \mathcal{F}_t, \mathcal{F}, \mathbb{P})$ be a probability space, on the compact interval $[0, T]$. Suppose we define a measure $\mathbb{Q}$ by

$$d\mathbb{Q} = L_T d\mathbb{P}, \text{ on } \mathcal{F}_T$$

where L_T is some $\mathcal{F}_T$-measurable, nonnegative and integrable random variable. By definition, L_T is the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ on $\mathcal{F}_T$ which, by the Radon-Nikodym Theorem, implies that $\mathbb{Q} << \mathbb{P}$ on $\mathcal{F}_T$.

Now, for every $t \leq T$ we have that $\mathcal{F}_t \subseteq \mathcal{F}_T$, and as a consequence $\mathbb{Q} << \mathbb{P}$ on $\mathcal{F}_t$. This implies the existence of the Radon-Nikodym derivative for every $t < T$, and so we can define the random process L_t on $[0, T]$ by

$$L_t = \frac{d\mathbb{Q}}{d\mathbb{P}}, \text{ on } \mathcal{F}_t.$$

Let us narrow our focus slightly to the case where L_t, as defined above, is a nonnegative $\mathbb{P}$-Local martingale with $L_0 = 1$. Clearly, in order for $\mathbb{Q}$ to be a probability measure we require that $\mathbb{E}^{\mathbb{P}}(L_T) = 1$. The following proposition, as found in [27], is useful for proving when this is the case.

Proposition 2.16 *Define the stopping time $\tau_n = \inf\{u : L(u) > n\} \wedge T$, and set $L^n(t) = L(t \wedge \tau_n)$. Define the measure $\mathbb{Q}^n$ via $\frac{d\mathbb{Q}^n}{d\mathbb{P}}\big|_{\mathcal{F}_t} = L^n_t$ so that for $A \in \mathcal{F}_t$, $\mathbb{Q}^n(A) = \mathbb{E}^{\mathbb{P}}(L^n_T \chi_A)$. Then the following are equivalent:*

1. *L is a martingale*
2. *$\mathbb{E}^{\mathbb{P}}(L_T) = 1$*
3. *$\mathbb{Q}^n(\tau_n < T) \to 0$ as $n \uparrow \infty$.*

If any of these conditions hold then $\mathbb{Q}$ is well defined and $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$.

Proof Let us begin by noticing that

$$\liminf L^n(t) = \liminf L(t \wedge \tau_n) = L(t)$$

and that $L^n(t)$ is a true martingale by (2.6). As a consequence, for every $s, t \in [0, T]$ with $s < t$, by Fatou's Lemma for conditional expectation we have that

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}(L(t)|\mathcal{F}_s) &\leq \liminf \mathbb{E}^{\mathbb{P}}(L^n(t)|\mathcal{F}_s) \\ &= \liminf L^n(s) \\ &= L(s) \end{aligned} \tag{2.14}$$

which implies that L is a supermartingale. For the implication that (1) $\Rightarrow$ (2) notice that

$$\mathbb{E}^{\mathbb{P}}(L_T) = \mathbb{E}^{\mathbb{P}}\big(\mathbb{E}^{\mathbb{P}}(L_T|\mathcal{F}_0)\big) = \mathbb{E}^{\mathbb{P}}(L_0) = 1.$$

For the reverse implication suppose that (1) does not hold. This implies that there exist an $s \in [0, T]$ such that $\mathbb{E}^{\mathbb{P}}(L_T|\mathcal{F}_s) \neq L(s)$, and since L is a supermartingale we must have

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}(L_T|\mathcal{F}_s) &< L(s) \\ \Rightarrow \mathbb{E}^{\mathbb{P}}\big(\mathbb{E}^{\mathbb{P}}(L_T|\mathcal{F}_s)|\mathcal{F}_0\big) &< \mathbb{E}^{\mathbb{P}}(L(s)|\mathcal{F}_0) \\ &\leq L_0 = 1 \\ \Rightarrow \mathbb{E}^{\mathbb{P}}(L_T|\mathcal{F}_0) &< 1 \\ \Rightarrow \mathbb{E}^{\mathbb{P}}\big(\mathbb{E}^{\mathbb{P}}(L_T|\mathcal{F}_0)\big) &< 1 \\ \mathbb{E}^{\mathbb{P}}(L_T) &< 1 \end{aligned}$$

so that (2) does not hold. For the remaining implications recall that L^n is a true martingale and that

$$\begin{aligned} 1 = \mathbb{E}^{\mathbb{P}}(L^n_T) &= \mathbb{E}^{\mathbb{P}}(L^n(T)\chi_{\{\tau_n = T\}}) + \mathbb{E}^{\mathbb{P}}(L^n(T)\chi_{\{\tau_n < T\}}) \\ &= \mathbb{E}^{\mathbb{P}}(L(T)\chi_{\{\tau_n = T\}}) + \mathbb{Q}^n(\{\tau_n < T\}). \end{aligned}$$

We then take limits on both sides of the above equation. If (3) holds then $\mathbb{Q}^n(\{\tau_n < T\})$ goes to zero so that $\mathbb{E}^{\mathbb{P}}(L(T)\chi_{\{\tau_n = T\}}) \to 1$, and by uniqueness of limits $E^{\mathbb{P}}(L_T) = 1$. Finally, if (2) holds then Doob's martingale inequality for $p = 1$, $C = n$ gives,

$$\mathbb{P}(\tau_n < T) = \mathbb{P}(\sup L_t > n) \leq \frac{1}{n} \to 0$$

which implies that $\mathbb{E}^{\mathbb{P}}(L_t\chi_{\{\tau_n = T\}}) \to 1$, so that $\mathbb{Q}^n(\{\tau_n < T\}) \to 0$ by the equation above.

For any $A \in \mathcal{F}_T$ we can, therefore, safely define $\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}(L_T\chi_A)$, which in turn implies that if $\mathbb{P}(A) = 0$ then $\mathbb{Q}(A) = 0$ and we have that $\mathbb{Q} << \mathbb{P}$. ∎

We are now ready to consider a special type of measure transformation. Consider a general probability space $(\Omega, \mathcal{F}, \mathbb{P}_0)$, and consider the following stochastic differential equations,

$$dX_t = \mu_0(t, X_t)dt + \sigma(t, X_t)dW_t \tag{2.15}$$
$$dX_t = \mu_1(t, X_t)dt + \sigma(t, X_t)dW_t \tag{2.16}$$

and let $\mathbb{P}$, $\mathbb{Q}$ be the law of the solution of (2.15) and (2.16) respectively, for a given initial condition $X_0 = x_0 \in \mathbb{R}$. Also, assume that $\sigma(t, \cdot) > 0$ and σ, μ_i are such that both equations have unique solutions. Let us also assume that the filtration is right-continuous and complete with respect to the measure $\frac{1}{2}(\mathbb{Q} + \mathbb{P})$.

If we define the process B_0 and B_1 as the solutions to the following

$$dB_t^0 = \frac{dX_t - \mu_1(t, X_t)}{\sigma(t, X_t)}$$

$$dB_1^1 = \frac{dX_t - \mu_0(t, X_t)}{\sigma(t, X_t)}.$$

Then B^0 and B^1 are standard Wiener processes under $\mathbb{P}$ and $\mathbb{Q}$ respectively, and by defining

$$\theta_t := \frac{\mu_1(t, X_t) - \mu_0(t, X_t)}{\sigma(t, X_t)}$$

we have that

$$dB_t^0 = dB_t^1 + \theta_t dt.$$

Under the above conditions we have the following proposition.

Proposition 2.17 *If* $\mathbb{Q}\big(\int_0^T \theta_u^2 du < \infty\big)$ *then* $\mathbb{Q} << \mathbb{P}$

Proof Define

$$L_t := e^{\int_0^t \theta_u dB_u^0 - \frac{1}{2}\int_0^t \theta_u^2 du}.$$

By the Itô formula this process solves the SDE,

$$dL_t = \theta_t L_t dB_t^0, \; L_0 = 1$$

and is therefore a $\mathbb{P}$-Local martingale. We also have that

$$L_t^{-1} = e^{-\int_0^t \theta_u dB_u^0 + \frac{1}{2}\int_0^t \theta_u^2 du}$$

$$= e^{-\int_0^t \theta_u dB_u^1 - \frac{1}{2}\int_0^t \theta_u^2 du}$$

so by our assumption we have that $\mathbb{Q}(\inf L_t^{-1} = 0) = 0$.

Now let τ_n and $\mathbb{Q}^n$ be defined as in proposition (2.16). Under $\mathbb{Q}^n$ the process X solves

$$dX_t = \sigma(t, X_t)dW_t + (\mu_1(t, X_t)\chi_{\{t\leq\tau_n\}} + \mu_1(t, X_t)\chi_{\{t>\tau_n\}})dt.$$

Path-wise uniqueness then implies that

$$\mathbb{Q}^n(\tau_n < T) = \mathbb{Q}^n(\sup Y_t > n) = \mathbb{Q}(\inf Y_t^{-1} < \frac{1}{n}) \to 0$$

so that by proposition 2.16 we have that L is a $\mathbb{P}$-martingale and that $\mathbb{Q} << \mathbb{P}$ $\blacksquare$

By interchanging the roles of $\mathbb{P}$ and $\mathbb{Q}$ we get the following corollary

Corollary 2.18 *If $\int_0^T \theta_u^2 du$ is almost surely finite under both $\mathbb{Q}$ and $\mathbb{P}$ then $\mathbb{Q}$ is equivalent to $\mathbb{P}$*

The above results will be extremely useful in our quest to find a martingale measures in the market models that we consider, where a martingale measure is defined as a measure under which the discounted price process is a martingale.

Martingale measures are extremely useful and in particular they allow us to make use of the martingale pricing Theorem

Theorem 2.19 *Assuming the existence of a short rate r_t, the arbitrage free process for the any contingent financial cliam at time T (a T-claim) X is given by*

$$V(t, X) = \mathbb{E}^{\mathbb{Q}}\left(e^{\int_t^T r(s)ds} X | \mathcal{F}_t\right)$$

where $\mathbb{Q}$ is a martingale measure.

Proof See, for example, Theorem 10.9 in [3].

We now have the theoretical background necessary for us to understand and construct certain market models of interest. In particular, our goal will be to find a model which has path dependent volatility, yields a pricing partial differential equation and is fitted to the market. We start our quest for such a model, ironically, by presenting the theory of local volatility.

3 Local volatility

3.1 Dupire's equation

In essence, local volatility models seek to solve the problem of volatility smiles. In his seminal paper on the subject Dupire aimed to solve the problem by formulating the following question: Is there a spot process for the underlying that is compatible with observed smiles, and also retains completeness as in the Black-Scholes framework? [14].

In order to satisfy the first part of the question, we assume that the prices for European options of all strikes K and maturities T are given. This will ensure that the correct prices are built into the model and so will fit the smile. In order to satisfy the requirement for completeness, we must not introduce a new source of randomness into the model for the underlying. We are therefore looking for a process with the following risk neutral dynamics:

$$\frac{dS_t}{S_t} = (r_t - d_t)dt + \sigma(t, S_t)dW_t. \tag{3.1}$$

Here W_t is a Wiener process, r_t is the instantaneous forward rate implied by the yield curve (we will assume this is a known deterministic function of t), d_t is the dividend yield (we will use $\mu_t := r_t - d_t$) and $\sigma(t, S_t)$ is a deterministic function of the current spot and time (the local volatility) [14]. Dupire then goes on to derive a form for the function $\sigma(t, S_t)$ based on the above, and with $\mu_t = 0$. We will prove the result for any μ_t.

We will need the following preliminary results which is a simple application of the martingale pricing Theorem to a European call option.

Theorem 3.1 (European Call option) *The price of a standard European call is equal to the discounted expected payoff conditional on the current underlying price. For an option with maturity T strike K, this can be written as:*

$$C_t = e^{-\int_t^T r(s)ds}\mathbb{E}_{t,x}[(S_T - K)^+]$$
$$C_t = D_{t,T}\mathbb{E}_{t,x}[(S_T - K)^+],$$

where $\mathbb{E}_{t,x}(\cdot) = \mathbb{E}(\cdot|S_t = x)$, $\mathbb{E} = (\cdot)$ is the expectation under the risk neutral measure and $D_{t,T} := e^{-\int_t^T r(s)ds}$.

If there exists a transition probability density $p(t_0, x_0; t, x)$, going from state (t_0, x_0) to (t, x), then (**??**) *can be written as follows:*

$$\begin{aligned} C_t &= D_{t,T}\int_0^\infty (s-K)^+ p(T,s)ds \\ &= D_{t,T}\int_K^\infty (s-K)p(T,s)ds, \end{aligned}$$

where we have written $p(t_0, x_0; t, x) := p(t, x)$ for brevity.

Notice that if we fix the current time t, (**??**) gives us the price of a call as a function of the strike K and maturity T, $C_t(T, K)$.

Theorem 3.2 (Fokker-Plank equation) *Assume that S_t, the solution of (3.1), has transition density $p(s, y; t, x)$. Then p will satisfy the* $\mathbf{Fokker-Plank\ equation}$

$$\frac{\partial}{\partial t}p(s, y; t, x) = -\frac{\partial}{\partial x}[\mu_t x p(s, y; t, x)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[\sigma^2(t, x)x^2 p(s, y; t, x)] \quad (3.2)$$

Proof See page 76 in [3].

We are now in a position to derive the main result from [14].The derivation is as in [10], but we provide additional detail that was omitted. If we let $p(T, x)$ denote the risk neutral density for the underlying at time T (again adopting our short hand $p(t_0, x_0; t, x) := p(t, x)$), then Theorem (3.1) gives us:

$$C_t = D_{t,T}\int_K^\infty (s-K)p(T, s)ds \quad (3.3)$$

From this relationship we can get the following:

1. First derivative with respect to K:

$$\begin{aligned}
\frac{\partial C}{\partial K} &= -D_{t,T}\frac{\partial}{\partial K}\int_\infty^K (s-K)p(T, s)ds \\
&= -D_{t,T}\left[\frac{\partial}{\partial K}\int_\infty^K sp(T, s)ds - \frac{\partial}{\partial K}K\int_\infty^K p(T, s)ds\right] \\
&= -D_{t,T}\left[Kp(T, K) - \int_\infty^K p(T, s)ds - Kp(T, K)\right] \\
\frac{\partial C}{\partial K} &= -D_{t,T}\int_K^\infty p(T, s)ds. \quad (3.4)
\end{aligned}$$

2. Second derivative with respect to K:

$$\begin{aligned}
\frac{\partial^2 C}{\partial K^2} &= -D_{t,T}\frac{\partial}{\partial K}\int_K^\infty p(T, s)ds \\
\frac{1}{D_{t,T}}\frac{\partial^2 C}{\partial K^2} &= p(T, K). \quad (3.5)
\end{aligned}$$

3. finally, the first derivative with respect to the maturity

$$\begin{aligned}
\frac{\partial C}{\partial T} &= \frac{\partial}{\partial T}D_{t,T}\int_K^\infty (s-K)p(T, s) + D_{t,T}\int_K^\infty (s-K)\frac{\partial}{\partial T}p(T, s) \\
\frac{\partial C}{\partial T} &= -r(T)D_{t,T}\int_K^\infty (s-K)p(T, s)ds + D_{t,T}\int_K^\infty (s-K)\frac{\partial}{\partial T}p(T, s)ds \\
\frac{\partial C}{\partial T} &= -r(T)C_t + D_{t,T}\int_K^\infty (s-K)\frac{\partial}{\partial T}p(T, s)ds. \quad (3.6)
\end{aligned}$$

For the next part of the derivation we will need Theorem (3.2). Since $p(T,s)$ is a density function, it satisfies (3.2) i.e:

$$\frac{\partial}{\partial T}p(T,s) = -\frac{\partial}{\partial s}[\mu_T sp(T,s)] + \frac{1}{2}\frac{\partial^2}{\partial s^2}[\sigma^2(T,s)s^2p(T,s)]. \tag{3.7}$$

By substituting (3.7) into (3.6), we get the following expression.

$$\begin{aligned}\frac{\partial C}{\partial T} &= -r_T C_t + D_{t,T}\int_K^\infty (s-K)\Big(-\frac{\partial}{\partial s}[\mu_T sp(T,s)] + \frac{1}{2}\frac{\partial^2}{\partial s^2}[\sigma^2(T,s)s^2p(T,s)]\Big)ds \\ &= -r_T C_t + D_{t,T}\Big(-\mu_T\int_K^\infty (s-K)\frac{\partial}{\partial s}[sp(T,s)]ds \\ &+ \frac{1}{2}\int_K^\infty (s-K)\frac{\partial^2}{\partial s^2}[\sigma^2(T,s)s^2p(T,s)]ds\Big)\end{aligned} \tag{3.8}$$

We derive the main result from equation (3.8) by simplifying the two integrals. In order to do this we first notice that combining equations (3.3) and (3.4) gives:

$$\int_K^\infty sp(T,s)ds = \frac{C}{D_{t,T}} - \frac{K}{D_{t,T}}\frac{\partial C}{\partial K}. \tag{3.9}$$

Now we consider the first integral:

$$\mu_T\int_K^\infty (s-K)\frac{\partial}{\partial s}[sp(T,s)]ds = [\mu_T(s-K)sp(T,s)]_K^\infty - \mu_T\int_K^\infty [sp(T,s)]ds$$

this by integration by parts. We assume that $p(t,s)$ decays exponentially as $s\to\infty$, this means that the term in brackets will converge to 0 at the boundary as $s\to\infty$. Using the assumption and equation (3.9), we have that:

$$\begin{aligned}\mu_T\int_K^\infty (s-K)\frac{\partial}{\partial s}[sp(T,s)]ds &= [0-0] - \mu_T\int_K^\infty [sp(T,s)]ds \\ \mu_T\int_K^\infty (s-K)\frac{\partial}{\partial s}[sp(T,s)]ds &= -\frac{\mu_T C}{D_{t,T}} + \frac{\mu_T K}{D_{t,T}}\frac{\partial C}{\partial K}.\end{aligned} \tag{3.10}$$

We use integration by parts again for the second integral

$$\begin{aligned}&\int_K^\infty (s-K)\frac{\partial^2}{\partial s^2}[\sigma^2(T,s)s^2p(T,s)]ds = \Big[(s-K)\frac{\partial}{\partial s}\sigma^2(T,s)s^2p(T,s)\Big]_K^\infty + \int_K^\infty \frac{\partial}{\partial s}[\sigma^2(T,s)s^2p(T,s)]ds \\ \Rightarrow &\int_K^\infty (s-K)\frac{\partial^2}{\partial s^2}[\sigma^2(T,s)s^2p(T,s)]ds = [0-0] + \big[\sigma^2s^2p(T,s)\big]_K^\infty \\ \Rightarrow &\int_K^\infty (s-K)\frac{\partial^2}{\partial s^2}[\sigma^2(T,s)s^2p(T,s)]ds = \sigma^2(T,K)K^2p(T,K) \\ \Rightarrow &\int_K^\infty (s-K)\frac{\partial^2}{\partial s^2}[\sigma^2(T,s)s^2p(T,s)]ds = \sigma^2(T,K)\frac{K^2}{D_{t,T}}\frac{\partial^2 C}{\partial K^2}.\end{aligned} \tag{3.11}$$

Again we have assumed that the transition density behaves appropriately at the boundary as s gets very large. If we now substitute expressions (3.10) and (3.11) into (3.8). We get:

$$\frac{\partial C}{\partial T} = -r_T C_t + D_{t,T}\left[\frac{\mu_T C}{D_{t,T}} - \frac{\mu_T K}{D_{t,T}}\frac{\partial C}{\partial K} + \frac{1}{2}\sigma^2(T,K)\frac{K^2}{D_{t,T}}\frac{\partial^2 C}{\partial K^2}\right]$$

$$\Rightarrow \frac{\partial C}{\partial T} = \frac{1}{2}\sigma^2(T,K)K^2\frac{\partial^2 C}{\partial K^2} + \mu_T\left(C - \frac{\partial C}{\partial K}\right) - r_T + C \tag{3.12}$$

$$\Rightarrow \sigma^2(T,K) = \frac{\frac{\partial C}{\partial T} + d_T C + (r_T - d_T)K\frac{\partial C}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}} \tag{3.13}$$

$$\Rightarrow \sigma(T,K) = \sqrt{\frac{\frac{\partial C}{\partial T} + d_T C + (r_T - d_T)K\frac{\partial C}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}}}. \tag{3.14}$$

Notice that if we set $r(T) = 0$ we recover the result from [14] exactly. Equation (3.12) is known as Dupire's forward equation. Equation (3.14) describes a mapping from the surface of call prices as a function of (T,K) to a volatility surface $\sigma = \sigma(T,K)$. It is often considered the definition of local volatility and we sometimes denote this σ^2 as σ^2_{LV}. The volatility surface that the equation describes is the unique function that, when used in (3.1), gives the correct prices for European options of all strikes and maturities, i.e it fixes the volatility smile.

There are some noticeable issues with formula (3.14). Firstly we notice the possibility for imaginary local volatility if the expression under the square root is negative, however this possibility is removed by considering arbitrage arguments. In the denominator, K^2 is always positive and $\frac{\partial^2 C}{\partial K^2}$ can also be show to always be positive by considering a infinitesimal butterfly spread [30]. No arbitrage also implies that a calendar spread must have positive value, this ensures the positivity of the numerator [30].

The second issue with (3.14) is one of stability. In practice we do not have a full continuum of option prices and so there is no analytical formula for the option price surface. This means we have to use numerical approximations for the derivatives in the formula. Since there is only one term in the denominator of the equation, and this term includes the second derivative with respect to strike, there are potential stability issues when this derivative is small since numerical errors are large relative to the derivative value at these points resulting in large absolute errors. An option has a very small dual gamma when it is far out-of-the-money, this effect is exaggerated for options of short maturity. The issue of stability can be addressed by considering a clever change of variables.

3.2 Local volatility in terms of implied volatility

The Black-Scholes implied volatility of an option can be considered a measure of its price. In fact, options are often traded using implied volatilities and not price. The goal of this section will be to express equation (3.12) in terms of the implied volatility of the European options and not their prices. We provide additional detail and insight, as well as a slight change, to the method presented in [18].

The goal is to express the derivatives in (3.12) using implied volatilities. Without loss of generality we can assume $t = 0$ and S_0 as our initial state. We start by defining the following dimensionless parameters:

$$y := \ln \frac{K}{F_T} \tag{3.15}$$
$$w(S_0, T, K) := \sigma^2_{imp}(S_0, T, K)T.$$

Here we use the futures price $F_T = S_0 e^{\int_0^T \mu_t dt}$. These parameters are known as the "log strike" and the "Black Scholes implied Total variance" respectively. With these new parameters we have that $C(S_0, T, K, \sigma) = C(S_0, T, y(T, K), w(T, K))$. We also have, by the definition of implied volatility:

$$C(S_0, T, K, \sigma) = C_{BS}(S_0, T, K, \sigma_{imp}(S_0, T, K)). \tag{3.16}$$

Where C_{BS} is the price obtained using the Black-Scholes formula. By the seminal paper on option pricing [4] we know:

$$C_{BS}(S_0, T, K, \sigma) = e^{-\int_0^T r_t dt}[F_T N(d_1) - KN(d_2)]$$
$$d_1 = \frac{1}{\sigma\sqrt{T}}\left[\ln\left(\frac{S_0 e^{\int_0^T \mu_t dt}}{K}\right) + \left(\frac{\sigma^2}{2}\right)(T)\right] \tag{3.17}$$
$$d_2 = d_1 - \sigma\sqrt{T}.$$

This is for a process of the form (3.1). Using the new parameters, defined in (3.15), in equation (3.17) we have the following:

$$C_{BS}(F_T, y, w) = S_0 e^{-\int_0^T d_t dt}[N(d_1) + e^y N(d_2)]$$
$$d_1 = -\frac{y}{\sqrt{w}} + \frac{\sqrt{w}}{2} \tag{3.18}$$
$$d_2 = -\frac{y}{\sqrt{w}} - \frac{\sqrt{w}}{2}.$$

It is now relatively easy to derive the following expressions using total derivatives and relationship (3.16).

1. First Derivative with respect to K

$$\begin{aligned}\frac{\partial C}{\partial K} &= \frac{\partial C}{\partial y}\frac{\partial y}{\partial K} + \frac{\partial C}{\partial w}\frac{\partial w}{\partial K} \\ \Rightarrow \frac{\partial C}{\partial K} &= \frac{\partial C}{\partial y}\frac{1}{K} + \frac{\partial C}{\partial w}\frac{\partial w}{\partial K}. \end{aligned} \tag{3.19}$$

2. Second derivative with respect to K

$$\begin{aligned}\frac{\partial^2 C}{\partial K^2} &= \frac{\partial}{\partial K}\left(\frac{\partial C}{\partial y}\frac{1}{K} + \frac{\partial C}{\partial w}\frac{\partial w}{\partial K}\right) \\ &= -\frac{1}{K^2}\frac{\partial C}{\partial y} + \frac{1}{K}\frac{\partial}{\partial y}\left(\frac{\partial C}{\partial K}\right) + \frac{\partial C}{\partial w}\frac{\partial^2 w}{\partial K^2} + \frac{\partial w}{\partial K}\frac{\partial}{\partial w}\left(\frac{\partial C}{\partial K}\right) \\ &= -\frac{1}{K^2}\frac{\partial C}{\partial y} + \frac{1}{K}\frac{\partial}{\partial y}\left(\frac{\partial C}{\partial y}\frac{1}{K} + \frac{\partial C}{\partial w}\frac{\partial w}{\partial K}\right) + \frac{\partial C}{\partial w}\frac{\partial^2 w}{\partial K^2} + \frac{\partial w}{\partial K}\frac{\partial}{\partial w}\left(\frac{\partial C}{\partial y}\frac{1}{K} + \frac{\partial C}{\partial w}\frac{\partial w}{\partial K}\right) \\ &= -\frac{1}{K^2}\frac{\partial C}{\partial y} + \frac{1}{K}\Big(\frac{\partial^2 C}{\partial y^2}\frac{1}{K} + \frac{\partial^2 C}{\partial w\partial y}\frac{\partial w}{\partial K}\Big) + \frac{\partial C}{\partial w}\frac{\partial^2 w}{\partial K^2} + \frac{\partial w}{\partial K}\Big(\frac{\partial^2 C}{\partial y\partial w}\frac{1}{K} + \frac{\partial^2 C}{\partial w^2}\frac{\partial w}{\partial K}\Big) \\ \Rightarrow \frac{\partial^2 C}{\partial K^2} &= \frac{1}{K^2}\Big(\frac{\partial^2 C}{\partial y^2} - \frac{\partial C}{\partial y}\Big) + \frac{2}{K}\frac{\partial^2 C}{\partial w\partial y}\frac{\partial w}{\partial K} + \frac{\partial C}{\partial w}\frac{\partial^2 w}{\partial K^2} \\ &\quad + \frac{\partial^2 C}{\partial w^2}\Big(\frac{\partial w}{\partial K}\Big)^2. \end{aligned} \tag{3.20}$$

3. First derivative with respect to T

$$\begin{aligned}\frac{\partial C}{\partial T} &= \frac{\partial C}{\partial T}\frac{\partial T}{\partial T} + \frac{\partial C}{\partial y}\frac{\partial y}{\partial T} + \frac{\partial C}{\partial w}\frac{\partial w}{\partial T} \\ &= \frac{\partial C}{\partial T} + \frac{\partial C}{\partial y}(d_T - r_T) + \frac{\partial C}{\partial w}\frac{\partial w}{\partial T} \\ \frac{\partial C}{\partial T} &= -d_T C + \frac{\partial C}{\partial y}(d_T - r_T) + \frac{\partial C}{\partial w}\frac{\partial w}{\partial T}. \end{aligned} \tag{3.21}$$

For the final derivative we used $y = \ln(\frac{K}{S_0}) - \int_0^T (r_t - d_t)dt$ to get $\frac{\partial y}{\partial T} = d_T - r_T$. The fact that (3.16) holds together with formula (3.18) gives us $\frac{\partial C}{\partial T} = -d_T C$. Substituting expressions (3.19),(3.20) and (3.21) into (3.12) gives the modified Dupire equation:

$$\begin{aligned}\sigma^2(T,K) &= \frac{(-d_T C + \frac{\partial C}{\partial y}(d_T - r_T) + \frac{\partial C}{\partial w}\frac{\partial w}{\partial T}) + d_T C + (r_T - d_T)K(\frac{\partial C}{\partial y}\frac{1}{K} + \frac{\partial C}{\partial w}\frac{\partial w}{\partial K})}{\frac{1}{2}K^2(\frac{1}{K^2}(\frac{\partial^2 C}{\partial y^2} - \frac{\partial C}{\partial y}) + \frac{2}{K}\frac{\partial^2 C}{\partial w\partial y}\frac{\partial w}{\partial K} + \frac{\partial C}{\partial w}\frac{\partial^2 w}{\partial K^2} + \frac{\partial^2 C}{\partial w^2}(\frac{\partial w}{\partial K})^2)} \\ \sigma^2(T,K) &= \frac{\frac{\partial C}{\partial w}\frac{\partial w}{\partial T} + K(r_T - d_T)\frac{\partial C}{\partial w}\frac{\partial w}{\partial K}}{\frac{1}{2}\Big(\frac{\partial^2 C}{\partial y^2} - \frac{\partial C}{\partial y} + K\frac{\partial^2 C}{\partial w\partial y}\frac{\partial w}{\partial K} + K^2\frac{\partial C}{\partial w}\frac{\partial^2 w}{\partial K^2} + K^2\frac{\partial^2 C}{\partial w^2}\big(\frac{\partial w}{\partial K}\big)^2\Big)}. \end{aligned} \tag{3.22}$$

In order to simplify further we use some simple identities from [18] derived from formula (3.18). Again we assume that relationship (3.16) holds.

$$\frac{\partial^2 C}{\partial w^2} = \Big(-\frac{1}{8} - \frac{1}{2w} + \frac{y^2}{2w^2}\Big)\frac{\partial C}{\partial w} \tag{3.23}$$

$$\frac{\partial^2 C}{\partial w \partial y} = \Big(\frac{1}{2} - \frac{y}{w}\Big)\frac{\partial C}{\partial w} \tag{3.24}$$

$$\frac{\partial^2 C}{\partial y^2} - \frac{\partial C}{\partial y} = 2\frac{\partial C}{\partial w}. \tag{3.25}$$

Substituting these identities into (3.22) gives us:

$$\sigma^2(T,K) = \frac{\frac{\partial w}{\partial T} + K(r_T - d_T)\frac{\partial w}{\partial K}}{\frac{1}{2}\big[2 + K(\frac{1}{2} - \frac{y}{w})\frac{\partial w}{\partial K} + K^2\frac{\partial^2 w}{\partial K^2} + K^2(-\frac{1}{8} - \frac{1}{2w} + \frac{y^2}{2w^2})\big(\frac{\partial w}{\partial K}\big)^2\big]}. \tag{3.26}$$

This is the equation seen in [18]. Finally we remember the goal of this section was to express equation (3.12) in terms of implied volatility and not Black-Scholes implied total variance. We therefore use the definition of w to get the partial derivatives in equation (3.26) in terms of σ_{imp}.

$$\frac{\partial w}{\partial K} = 2\sigma_{imp}T\frac{\partial \sigma_{imp}}{\partial K} \tag{3.27}$$

$$\frac{\partial^2 w}{\partial K^2} = 2T(\frac{\partial \sigma_{imp}}{\partial K})^2 + 2\sigma_{imp}T\frac{\partial^2 \sigma_{imp}}{\partial K^2} \tag{3.28}$$

$$\frac{\partial w}{\partial T} = \sigma_{imp}^2 + 2T\sigma_{imp}T\frac{\partial \sigma_{imp}}{\partial T}. \tag{3.29}$$

After much work we are finally able to express the formula for local volatility as a function of the implied volatility, as in [30], by substituting equations (3.27),(3.28) and (3.29) into (3.26) and performing some algebraic manipulations:

$$\begin{aligned} v_L &= \sigma^2(T,K) \\ &= \frac{\sigma_{imp}^2 + 2T\sigma_{imp}\frac{\partial \sigma_{imp}}{\partial T} + 2(r_T - d_T)KT\sigma_{imp}\frac{\partial \sigma_{imp}}{\partial K}}{\Big(1 + Kd_1\sqrt{T}\frac{\partial \sigma_{imp}}{\partial K}\Big)^2 + K^2T\sigma_{imp}\Big(\frac{\partial^2 \sigma_{imp}}{\partial K^2} - d_1\sqrt{T}\big(\frac{\partial \sigma_{imp}}{\partial K}\big)^2\Big)}. \end{aligned} \tag{3.30}$$

With d_1 as in equation (3.17) with $\sigma = \sigma_{imp}$. We see that the transformation from call prices to implied volatility results in a denominated with multiple terms, unlike the one term in (3.14). This means small errors resulting from numerical approximations for the derivatives will not be compounded. Therefore (3.26) is a more stable formula.

In the two preceding subsections we have derived formulas for the local volatility function, without giving a proper interpretation of what the local volatility represents. It is possible to gain more insight in to the nature and interpretation of the local volatility function by considering a stochastic volatility model and a result from measure theory that has grown in importance since its application to volatility models was discovered. We discuss this result next.

3.3 Gyöngy's mimicking process

Option traders make trading decisions based on future expectations, it is therefore natural to assume that Black-Scholes implied volatility for an option is a measure of the markets' expectations for volatility over the life of that option. Although this is a tempting line of thought, it is wrong, since implied volatility is in fact a measure of price and can not be considered a statistical measure [37]. This reasoning is, however, useful when applied to local volatility.

Derman et al. [11] were able to show the following.

Proposition 3.3 (Local Volatility as a conditional expectation) *Assume that a process has dynamics,*

$$\begin{aligned} \frac{dS_t}{S_t} &= \mu_t dt + \sigma_t dW_t, \\ S_0 &= x \end{aligned} \tag{3.31}$$

where W_t is a standard Wiener process under some measure, μ_t is a deterministic function and σ_t follows a, as yet, unspecified random process. If, in addition to this, we have risk-neutral prices for all European options then the local variance, σ^2_{LV} as in (3.13), is the conditional risk-neutral expectation of the instantaneous future variance of the underlying given that $S_T = K$. In other words we have that:

$$\sigma^2(T, K) = E(\sigma_T^2 | S_T = K). \tag{3.32}$$

Proof We will prove this result for the case when S_t has dynamics as in (3.31) with $\mu_t = r$ where $r > 0$ is a constant, since we can always move to this case from (3.31) using a change in measure.

Assuming that the transition density for the spot price, $p(t, x)$, exists we recall that the price of a call option at time 0 is

$$\begin{aligned} C &= e^{-rT}\mathbb{E}_{0,x}[(S_T - K)^+] \\ &= \mathbb{E}_{0,x}\big(e^{-rT}(S_T - K)\chi_{\{S_T > K\}}\big), \end{aligned} \tag{3.33}$$

where $\chi_{\{S_T>K\}}$ is the characteristic function of the set $\{S_T > K\}$. We note that the characteristic function has the following properties:

$$\frac{\partial}{\partial s}\chi_{\{s>K\}} = \delta(s-K) \tag{3.34}$$

$$\begin{aligned}\frac{\partial}{\partial K}\chi_{\{s>K\}} &= \frac{\partial}{\partial K}(1-\chi_{\{K>s\}})\\ &= -\delta(s-K),\end{aligned} \tag{3.35}$$

where $\delta(\cdot)$ is the Dirac-delta function. Using this, and assuming that we have the appropriate integrability conditions required for Fubini's Theorem to hold, we have the following useful identities

$$\begin{aligned}\frac{\partial C}{\partial K} &= e^{-rT}\frac{\partial}{\partial K}\mathbb{E}_{0,x}((S_T-K)\chi_{\{S_T>K\}})\\ &= e^{-rT}\big(\mathbb{E}_{0,x}((K-S_T)\delta(S_T-K)) - \mathbb{E}_{0,x}(\chi_{\{S_T>K\}})\\ &= -e^{-rT}\mathbb{E}_{0,x}(\chi_{\{S_T>K\}})\end{aligned} \tag{3.36}$$

$$\frac{\partial^2 C}{\partial K^2} = e^{-rT}\mathbb{E}_{0,x}(\delta(S_T-K)). \tag{3.37}$$

Next we consider the function $g(t,x) = e^{-rt}(x-K)\chi_{\{x>K\}}$. Clearly this gives

$$\begin{aligned}g_1 &= -rg\\ g_2 &= e^{-rt}(x-K)\delta(x-K) + e^{-rt}\chi_{\{x>K\}} = e^{-rt}\chi_{\{x>K\}}\\ g_{22} &= e^{-rt}\delta(x-K).\end{aligned}$$

Applying the Itô formula to the process $g(T,S_T)$ then gives that

$$\begin{aligned}d(e^{-rT}(S_T-K)^+) &= -re^{-rT}(S_T-K)\chi_{\{s_T>K\}} + e^{-rT}\chi_{\{S_T>K\}}dS_T + \frac{1}{2}e^{-rt}\delta(S_T-K)d\langle S\rangle_T\\ &= -re^{-rT}(S_T-K)\chi_{\{s_T>K\}}dT + e^{-rT}\chi_{\{S_T>K\}}(S_T r dT + S_T\sigma_T dW_T) + \frac{1}{2}e^{-rt}\delta(S_T-K)(S_T^2\sigma_T^2 dT)\\ &= e^{-rT}\big(Kr\chi_{\{s_T>K\}} + \frac{1}{2}\delta(S_T-K)S_T^2\sigma_T^2\big)dT + e^{-rT}\chi_{\{S_T>K\}}S_T\sigma_T dW_T.\end{aligned}$$

This, together with (3.33), implies that

$$dC = e^{-rT}\mathbb{E}_{0,x}\Big(\big(Kr\chi_{\{s_T>K\}} + \frac{1}{2}\delta(S_T-K)S_T^2\sigma_T^2\big)dT + \chi_{\{S_T>K\}}S_T\sigma_T dW_T\Big)$$

since the expected value of a stochastic integral is always 0 we have that

$$\begin{aligned}\frac{\partial C}{\partial T} &= e^{-rT}\mathbb{E}_{0,x}\Big(Kr\chi_{\{s_T>K\}} + \frac{1}{2}\delta(S_T-K)S_T^2\sigma_T^2\Big)\\ &= e^{-rT}Kr\mathbb{E}_{0,x}(\chi_{\{s_T>K\}}) + e^{-rT}\frac{1}{2}K^2\mathbb{E}_{0,x}(\delta(S_T-K)\sigma_T^2).\end{aligned}$$

We can write the second term as

$$e^{-rT}\frac{1}{2}K^2\mathbb{E}_{0,x}(\delta(S_T-K)\sigma_T^2) = e^{-rT}\frac{1}{2}K^2\mathbb{E}_{0,x}(\sigma_T^2|S_T=K)\mathbb{E}_{0,x}(\delta(S_T-K)),$$

which gives

$$\frac{\partial C}{\partial T} = e^{-rT}Kr\mathbb{E}_{0,x}(\chi_{\{s_T>K\}}) + e^{-rT}\frac{1}{2}K^2\mathbb{E}_{0,x}(\sigma_T^2|S_T = K)\mathbb{E}_{0,x}(\delta(S_T - K)),$$

using equations (3.36) and (3.37) gives,

$$\frac{\partial C}{\partial T} = -Kr\frac{\partial C}{\partial K} + \frac{1}{2}K^2\mathbb{E}_{0,x}(\sigma_T^2|S_T = K)\frac{\partial^2 C}{\partial K^2}$$

so that, finally we have

$$\mathbb{E}_{0,x}(\sigma_T^2|S_T = K) = \frac{\frac{\partial C}{\partial T} + Kr\frac{\partial C}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}}.$$

Comparing this equation with (3.13) (in the case where $d_T = 0$) it is clear that

$$\mathbb{E}_{0,x}(\sigma_T^2|S_T = K) = \sigma_{LV}^2.$$

∎

Dupire also proved a similar result in [15], however, neither were first in arriving to this conclusion. The work done by Dupire and Derman et al. can be considered an application of a more general result proposed by Krylov [32] and proved by Gyöngy [22]. The Theorem is as follows:

Theorem 3.4 (Gyöngy's Theorem) *Suppose that ξ_t is a real-valued one-dimensional Itô process starting at 0 with dynamics*

$$d\xi_t = \alpha(t,\omega)dt + \beta(t,\omega)dW_t, \tag{3.38}$$

where W_t is a k-dimensional Wiener process on the probability space $(\Omega, P, \mathscr{F}_t)$, $\alpha(t,\omega)$ and $\beta(t,\omega)$ are bounded $\mathscr{F}_t$ non-anticipative processes such that $\beta\beta^T$ is uniformly positive definite. Then there exists another, one-dimensional, stochastic process $\tilde{\xi}_t$ which has the same marginal probability distribution as ξ_t for all t, and $\tilde{\xi}_t$ is a solution of the SDE

$$d\tilde{\xi}_t = a(t,\tilde{\xi}_t)dt + b(t,\tilde{\xi}_t)d\tilde{W}_t \tag{3.39}$$

with non-random coefficients a and b on some space $(\Omega, \tilde{P}, \tilde{\mathscr{F}}_t)$. These coefficients have the simple interpretation

$$a(t,x) = E[\alpha(t,\omega)|\xi_t = x]$$

$$b(t,x) = \sqrt{E[\beta\beta^T(t,\omega)|\xi_t = x]}. \tag{3.40}$$

The result was then extended by Brunick and Shreve [7]. They were able to relax the regularity conditions on $\beta(t, \omega)$, requiring only integrability. This allowed the result to be applied to the popular stochastic volatility model created by Heston [26]. Since we have already proved a special case of this result (the most relevant case for our purposes) and considering that the proof of the Theorem requires advanced measure theory and is beyond the scope of this dissertation, we will not prove it here. What is important for our purposes is the insight it gives.

Gyöngy's result tells us that if we have a process with random coefficients, subject to some conditions, there is a process which solves an SDE with non-random coefficients and has the same marginal distribution as the original process. The new process is said to be a simpler mimicking process of the original [1]. At this point we remember that the value of a European call is only dependent on the risk neutral density of the underlying at expiry conditional on the initial state. The Theorem, therefore, tells us that if we have an underlying that is governed by a process with stochastic volatility, such as the one in [26], then there is a new process with non-random volatility that generates the same European option prices. If we compare equation (3.32) with (3.40), we see that this new process has a volatility coefficient of the exact same form as the one shown by Dupire in [14]. This essentially shows that the Dupire local volatility function is the best non-random approximation to a random volatility process that generates the current market prices for European options. This is a nice justification for the use of local volatility models.

3.4 Construction of the implied volatility surface

Up to this point we have proved general results that hold for all local volatility models given certain assumptions, these results are widely accepted and used. Throughout we have assumed that a full continuum of option prices is known, which in turn means that the we have a full, continuous and sufficiently smooth implied volatility function that we can use in formula (3.30) in order to generate a Local volatility surface. This is, of course, a naive assumption since in reality we only have option prices for a finite number of maturities and strikes.

This leads to the question of how to construct the implied volatility surface, and it is in this area where debate is most fierce. The various methods of construction can be grouped into three broad groups, theoretical construction, representations based on interpolation and smoothing techniques, and finally, parametric representations.

3.4.1 Theoretical construction

This type of construction normally involves a stochastic volatility model. From the stochastic volatility model it may be possible to derive a analytical expression for the implied volatility surface, and then obtain the appropriate derivatives that are used in (3.30). This technique can be employed in the frameworks such as that of the Heston volatility model [26] and the stochastic-$\alpha\beta\rho$,or SABR, model [23].

As an example we will look at the SABR model. In this framework we have the following dynamics:

$$\begin{aligned} dF_t &= \alpha_t F_t^{\beta} dW_t^1 \\ d\alpha_t &= \nu \alpha_t dW_t^2 \\ dW_t^1 \cdot dW_t^2 &= \rho dt, \end{aligned} \tag{3.41}$$

where F_t is the forward price with $F_0 = S_0 e^{(r_0-d_0)T}$, α_t the volatility function starting at $\alpha_0 = \alpha$, ν is the volatility of volatility and ρ the correlation between the two Wiener processes. Using this, the following approximation for the implied volatility surface was derived in [23]:

$$\begin{aligned} \sigma_{imp}(F_0, K) \approx & \alpha (F_0 K)^{\frac{\beta-1}{2}} \left[1 + \frac{(1-\beta)^2}{24} \ln\left(\frac{F_0}{K}\right) + \frac{(1-\beta)^4}{1920} \ln^4\left(\frac{F_0}{K}\right)\right]^{-1} \left(\frac{z}{x(z)}\right) \\ & \left[1 + \left(\frac{\alpha^2(1-\beta)^2}{24(F_0K)^{1-\beta}} + \frac{\rho\beta\nu\alpha}{4(F_0K)^{\frac{1-\beta}{2}}} + \frac{2-3\rho^2}{24}\nu^2\right)T\right], \end{aligned} \tag{3.42}$$

where,

$$z = \frac{\nu}{\alpha}(F_0K)^{\frac{1-\beta}{2}} \ln\left(\frac{F_0}{K}\right) \tag{3.43}$$

$$x(z) = ln\left(\frac{\sqrt{1-2\rho z + z^2} + z - \rho}{1-\rho}\right). \tag{3.44}$$

This formula can then be used to get analytic expressions for the derivatives in formula (3.30), however it should be clear by looking at formula (3.42), that the resulting expression will be too large to be written into a computer programme if we want timely results. This highlights the main issue with these kind of theoretical construction: they are very computationally expensive. We can, of course, generate a volatility surface on a discrete grid using (3.42), and then numerically approximate the derivatives, but this will introduce more errors.

The second issue is one concerning the fit of the model to the data. We see that in the construction of the model we have some parameters, namely α, β and ρ. These can be selected to minimize the error between the market data and the implied volatility the model generates at these specific strikes and maturities, however their is no guarantee that all the market data will fit onto the generated surface exactly.

Despite these technical issues, it has been shown that deriving the volatility surface with more technical approximations does lead to some predictive power [19], and the volatility surface generated has realistic dynamics.

Another subclass of such constructions are those based on Levy processes. These processes are useful when we have steep short term skews. They allow for jumps in the diffusion process. Some popular models of this form are the model by Kou [31], and the variance gamma model by Carr et al. [33].

3.4.2 Interpolation and smoothing

We can also construct a implied volatility surface by interpolating market data, and then smooth the resulting surface so that the derivatives in formula (3.30) are stable.

Kahale [28] proposes a method that involves uses piecewise convex polynomials that mimic the Black-Scholes pricing formula. This results in an arbitrage free price surface, and so is any resulting volatility smile.

Another scheme, that was proposed as a candidate for this problem in [16] and [6], is the interpolation of the volatility surface using Thin Plate Splines. The surface is then smoothed out, the resulting surface may not go through each data point exactly, but the smoothing can be controlled so that we lie within the bid-ask spread at these points.

These kinds of schemes have a computational advantage over those in the preceding sub section and they also fit market data more closely. They do not, however, allow us to derive analytical derivatives of the volatility surface and they do not have the same predictive power either.

3.4.3 Parametric construction

These methods are extremely popular due to their simplicity and usefulness. In [13] it was proposed that the volatility surface be approximated as a quadratic function of the moneyness $\mathscr{M} := \ln\left(\frac{F}{K}\right)\sqrt{T}$, i.e:

$$\sigma_{imp}(\mathscr{M}, T) = b_1 + b_2\mathscr{M} + b_3\mathscr{M}^2 + b_4 T + b_5\mathscr{M}T. \tag{3.45}$$

When applied to oil markets, it was found in [5] that the method gives a very rough average of the volatility surface due to the assumption that the quadratic function is the same across all maturities. It was proposed that each maturity be treated separately. A quadratic function of moneyness was fitted across all strikes for each maturity, and then data was interpolated between each maturity. Safex use a method of this form for liquid ALSI option (which is an option on a future) [30]. For this reason we will discuss their method in detail.

Safex fits a three-parameter quadratic polynomial to the market data for each traded expiry [30], i.e a function of the form:

$$\sigma_{imp}(M) = \beta_0^k + \beta_1^k M + \beta_2^k M^2, \tag{3.46}$$

where $k = 1, 2, ..., n$ are the n listed expiries, $M := K/F$ is the moneyness. We remember that the ALSI options are options on the ALSI futures contract, and so our spot price is the futures level F. β_i^k for $i = 1, 2, 3$ are parameters that we can choose for each listed expiry. β_1^k is known as the correlation term, no-spread-arbitrage requires that $-1 < \beta_1^k < 0,$ for all k. This also ensures that the function gives the desirable property that volatility is negatively correlated with the spot price. β_2^k is known as the volatility of volatility (vol-of-vol) parameter, and no-calendar-spread arbitrage convexity condition implies that $\beta_2^k > 0,$ for all k [30].

Equation (3.46) describes n parabolas that can be fitted to the market, giving n skews. The goal, however, is to get a three-dimensional surface. To do this Safex gives a functional form for the At-the-Money (ATM) volatility term structure. It is the following exponential:

$$\sigma_{ATM}(t) = \frac{\theta}{t^{\lambda}}. \tag{3.47}$$

Here t is the time to expiry, λ controls the slope (which may be positive or negative), and θ controls the short term ATM curvature. In order to link the two models in (3.46) and (3.47), and to ensure our resulting surface is continuous we assume that:

$$\beta_i(t) = \frac{\theta_i}{t^{\lambda_i}}, \text{ for } i = 1, 2, 3. \tag{3.48}$$

Notice we have dropped the subscript k, and the β's are now a function of t instead. This is because we have moved away from the discrete number of listed expiries to a continuous setting. We now have

$$\sigma_{imp}(t, M) = \frac{\theta_0}{t^{\lambda_0}} + \frac{\theta_1}{t^{\lambda_1}} M + \frac{\theta_2}{t^{\lambda_2}} M^2. \tag{3.49}$$

This fully describes a 3D volatility surface and we can select the 3 parameters to fit the market data using optimization techniques. Switching back to the strike price and futures level we have that

$$\sigma_{imp}(t, M) = \frac{\theta_0}{t^{\lambda_0}} + \frac{\theta_1}{t^{\lambda_1}}\frac{K}{F} + \frac{\theta_2}{t^{\lambda_2}}\left(\frac{K}{F}\right)^2. \tag{3.50}$$

It is now a straight forward exercise to obtain the derivatives:

$$\begin{aligned}
\frac{\partial \sigma_{imp}}{\partial K} &= \frac{\theta_1}{Ft^{\lambda_1}} + \frac{2\theta_2 K}{F^2 t^{\lambda_2}} \\
\frac{\partial^2 \sigma_{imp}}{\partial K^2} &= \frac{2\theta_2}{F^2 t^{\lambda_2}} \\
\frac{\partial \sigma_{imp}}{\partial t} &= -\lambda_0 \frac{\theta_0}{t^{(\lambda_0+1)}} - \lambda_1 \frac{K\theta_1}{Ft^{(\lambda_1+1)}} - \lambda_2 \frac{K^2\theta_2}{F^2 t^{(\lambda_2+1)}}.
\end{aligned} \tag{3.51}$$

These are the derivatives that can be used in formula (3.30). The parameters θ_i and λ_i for $i = 0, 1, 2$ are published every two weeks by Safex.

4 Path dependent volatility

4.1 A general model

We now begin the hallmark chapter of this dissertation by introducing path dependent models in their most natural form, as done in [25]. The construction of the model contains more heuristics than rigour. However, it is included only to build an intuition for the following subsections. With this in mind, suppose that we have a financial market with a risk-free asset and a single risky asset over the time period $[0, T]$. We assume that the risk-free asset, defined as the process B, has dynamics governed by the equation

$$dB_t = rB_t dt,\ t \in [0, T], \tag{4.1}$$

for some constant $r > 0$. In addition, we assume that the risky asset has dynamics of the form

$$dS_t = \mu S_t dt + \Sigma_t dW_t,\ t \in [0, T], \tag{4.2}$$

where W is a Brownian motion under some probability measure $\mathbb{P}$, and μ is a positive constant representing the drift under the measure $\mathbb{P}$. Σ is assumed to be a stochastic process, such that the filtration generated by S is the same as the filtration generated by W, meaning that, knowing W up to time t is equivalent to knowing S up to time t. This filtration will be denoted by $(\mathcal{F}_t)_{t\geq 0} := (\sigma(S_u, u \leq t))$. In order to force our model to have some desired properties, we make the following assumptions about the process Σ,

1. We assume that Σ is adapted to the filtration $(\mathcal{F}_t)$. This will ensure that future values of Σ are only influenced by past and the present values of S_t and not future values. i.e that the process in not anticipative.

2. Σ is such that the risky asset prices are positive. i.e that $S_t > 0$, for all $t \in [0, T]$.

The first assumption together with the Doob-Dynkin Theorem suggests that Σ should be of the form

$$\Sigma_t = f(t, (S_u, u \leq t)),$$

where $f(t, \cdot)$ is some Borel measurable function. The second assumption will force Σ to be of the form [25]

$$\Sigma_t = S_t \sigma(t, (S_u, u \leq t)), \tag{4.3}$$

where again $\sigma(t, \cdot)$ is a Borel measurable function. We then have, with these two assumptions, a financial market model with dynamics

$$dB_t = rB_t dt, \tag{4.4}$$
$$dS_t = \mu S_t dt + S_t \sigma(t, (S_u, u \leq t))dW_t, \; t \in [0, T]. \tag{4.5}$$

Clearly we have a geometric model where the volatility term $\sigma(t, \cdot)$ is a function of the entire path of S from time zero to time t. For this reason, we will call any model of the form of (4.4)-(4.5) a path-dependent volatility model.

Essential to the success of any diffusion model is the existence of a martingale measure $\mathbb{Q}$. This measure can be constructed by setting $h_t = \frac{-\mu+r}{\sigma(t,(S_u,u\leq t))}$ in Theorem 2.14. Whether or not h_t satisfies the conditions for the Theorem to hold will depend on the properties of the function $\sigma(t, \cdot)$. For now, if we assume that $\sigma(t, \cdot)$ has the necessary properties and that the martingale measure does exist (we will address this question more rigorously when considering specific models), then the $\mathbb{Q}$ dynamics of the model are

$$dB_t = rB_t dt, \tag{4.6}$$
$$dS_t = rS_t dt + S_t \sigma(t, (S_u, u \leq t))dW_t, \; t \in [0, T]. \tag{4.7}$$

It should be clear at this point that the $\mathbb{Q}$ Brownian motion W is the only source of randomness in the above model. The fact that we have an equal number of risky assets and Brownian motions suggests that this market is complete. This can be shown, as in [25], using the martingale representation Theorem. We first construct a self financing portfolio consisting of the risk-free and risky asset defined as

$$V_t = \alpha_t B_t + \Delta_t S_t. \tag{4.8}$$

The self financing condition implies that

$$dV_t = \alpha_t dB_t + \Delta_t dS_t.$$

This, together with 4.7, implies that

$$dV_t = r(\Delta_t S_t) + \alpha_t B_t)dt + \Delta_t \sigma(t, (S_u, u \leq t))S_t dW_t. \tag{4.9}$$

We can then use the martingale representation Theorem to prove that this portfolio can be used to replicate any financial claim at time T. We refer to any such claim as a "T-claim". This is formulated in the following proposition.

Proposition 4.1 (Completeness) *For every T-claim X with finite variance, there exists a unique self-financing portfolio defined by (4.8) such that $V_T = X$ a.s.*

Proof Define the random variable $M_t := \mathbb{E}_Q\left[e^{-rT}X|\mathcal{F}_t\right]$, clearly M is a $\mathbb{Q}$ martingale since, for every $s \in [0,T]$ such that $s < t$,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[M_t|\mathcal{F}_s] &= \mathbb{E}_{\mathbb{Q}}\Big[\mathbb{E}_{\mathbb{Q}}\left[e^{-rT}X|\mathcal{F}_t\right]|\mathcal{F}_s\Big]\\ &= \mathbb{E}_{\mathbb{Q}}\left[e^{-rT}X|\mathcal{F}_s\right]\\ &= M_s.\end{aligned}$$

Therefore by the martingale representation Theorem there exists a process ξ adapted to $(\mathcal{F}_t)_{t\geq 0}$ such that

$$\begin{aligned}M_t &= M_0 + \int_0^t \xi_u dW_u,\ t \in [0,T]\\ &= \mathbb{E}_{\mathbb{Q}}\left[e^{-rT}X\right] + \int_0^t \xi_u dW_u. \end{aligned} \tag{4.10}$$

Therefore, in order to have that $V(T) = X$ we must have that

1. $V_0 = \mathbb{E}_{\mathbb{Q}}[e^{-rT}X]$
2. $dV_t = \xi_t dW_t$.

The first condition is straight forward. By (4.9), the second condition requires us to set

$$\Delta_t = \frac{\xi_t}{S_t\sigma(t,(S_u,u\leq t))} \tag{4.11}$$

$$\text{and, } \alpha_t = V_t - \Delta_t S_t. \tag{4.12}$$

∎

The above does not give us a way of calculating the hedging strategy exactly. We have only shown that the process ξ exists, not what it is explicitly. Regardless, we have shown that a hedging strategy exists for any claim with finite variance, and so the market is complete. This highlights one of the major advantages of path-dependent volatility models: they can incorporate past data, making room for richer volatility dynamics, while still retaining completeness.

There are, however, some serious difficulties that arise in the pricing of contingent claims under such a model. To illustrate this, we look at the simplest example of an equation of the form of (4.5).

4.2 A Delayed model

4.2.1 The model and its solution

We will now consider the following market model, suggested in [2], as a special case of equations (4.4) and (4.5). One advantage of this model is that we get a closed form solution to the price of a European call in a certain domain of $t \in [0, T]$. In this section we will replicate the results found in [2], varying the method of derivation only slightly, while providing more detail and discussion of the results and their derivations.

Suppose we have a risk-free asset, B, with dynamics as in (4.4), furthermore suppose we have a risky asset with $\mathbb{P}$ dynamics governed by the following SDE

$$dS_t = \mu S_t dt + \sigma(S_{(t-b)}) S_t dW_t, \ t \in [0, T] \tag{4.13}$$

$$S_t = f(t), \ t \in [-b, 0]. \tag{4.14}$$

Here μ and b are positive constants and the function $\sigma : \mathbb{R} \to \mathbb{R}$ is assumed to be strictly positive and Lipschitz continuous. The filtration which we consider is as it is in section 4.1. The initial process $f : \Omega \to C([-b, 0], \mathbb{R})$ is $\mathcal{F}_0$-measurable with respect to the Borel σ-algebra of $C([-b, 0], \mathbb{R})$. We recall, by our brief study of stochastic functional equations, that this is enough information to guarantee a unique solution to (4.13) with initial data (4.14). The solution to this equation was found in [2], we will verify their result by the following proposition.

Proposition 4.2 *For a given $\mathcal{F}_0$-measurable function, f, the stochastic process S defined by $S_t = f(0)e^{\mu t + \int_0^t \sigma(f(u-b))dW_u - \frac{1}{2}\int_0^t \sigma^2(f(u-b))du}$ solves (4.13)-(4.14) when $t \in [0, b]$.*

Proof 4.2 We start by defining the following process and its quadratic variation as

$$N_t = \mu t + \int_0^t \sigma(f(u-b))dW_u$$
$$\langle N, N \rangle_t = \int_0^t \sigma^2(f(u-b))du$$

taking differentials we then have

$$dN_t = \mu dt + \sigma(f(t-b))dW_t$$
$$d\langle N, N \rangle_t = \sigma^2(f(t-b))dt.$$

Next we consider the function $g : [0, b] \times \mathbb{R}^2 \to \mathbb{R}$ defined by $g(t, x, y) = f(0)e^{x-\frac{1}{2}y}$, remembering that $f(0)$ is known since $t \in [0, b]$ and f is $\mathcal{F}_0$-measurable. By the Itô formula

$$
\begin{aligned}
dS_t &= dg(t, N_t, \langle N, N\rangle_t) \\
&= g(t, N_t, \langle N, N\rangle_t)dN_t - \frac{1}{2}g(t, N_t, \langle N, N\rangle_t)d\langle N, N\rangle_t + \frac{1}{2}g(t, N_t, \langle N, N\rangle_t)\big(dN_t\big)^2 \\
&= \big(\mu S_t + \frac{1}{2}S_t\sigma^2(f(t-b) - \frac{1}{2}S_t\sigma^2(f(t-b)\big) + S_t\sigma(f(t-b))dW_t \\
&= \mu S_t dt + \sigma(S_{t-b})S_t dW_t \qquad (4.15)
\end{aligned}
$$

since $t \in [0, b]$ implies that $t - b \in [-b, 0]$ so that $f(t - b) = S_{t-b}$. Clearly $S(0) = f(0)$ and so we have shown that the proposed solution satisfies equations (4.13) and (4.14) when $t \in [0, b]$.
∎

Using this solution and following the same argument we can show that

$$S_t = S_b e^{\mu(t-b)+\int_b^t \sigma(S_{u-b})dW_u - \frac{1}{2}\int_b^t \sigma^2(S_{u-b})du}$$

solves the stochastic functional differential equation on the interval $t \in [b, 2b]$. Then by induction we can construct a solution over the whole interval $[0, T]$. This has the form

$$S_t = f(0)e^{\mu t+\int_0^t \sigma(S_{u-b})dW_u - \frac{1}{2}\int_0^t \sigma^2(S_{u-b})du} \qquad (4.16)$$

or alternatively, for any $s \in [0, T]$ such that $s < t$,

$$S_t = S_s e^{\mu(t-s)+\int_s^t \sigma(S_{u-b})dW_u - \frac{1}{2}\int_s^t \sigma^2(S_{u-b})du}. \qquad (4.17)$$

Clearly, when $f(t) > 0$ for all $t \in [-b, 0]$, we will have that $S_t > 0$ for all $t \in [0, T]$, and so the solution has the desirable property of being strictly positive at all times. We can now proceed with the question of pricing a contingent claim.

4.2.2 Option pricing

Consider a European call option with maturity time T and strike price K on our risky asset S. This will be done, as in [2], using the martingale pricing technique. The first step is to find the appropriate measure, $\mathbb{Q}$, under which the discounted share price is a martingale, this is equivalent to the stock price having a drift of $\mu = r$ in (4.13). In order to find the new measure we can use Girsanov's Theorem. By Girsanov's Theorem, assuming that h satisfies certain conditions, the process

$$d\bar{W}_t := dW_t - h_t dt, \; t \in [0, T], \qquad (4.18)$$

is a $\mathbb{Q}$-Brownian motion. Substituting the above into (4.13) in order to get the $\mathbb{Q}$ dynamics of S gives

$$dS_t = (\mu + \sigma(S_{t-b})h_t)S_t dt + \sigma(S_{t-b})S_t d\bar{W}_t,\ t \in [0,T].$$

So if we choose $h_t = \frac{r-\mu}{\sigma(S_{t-b})}$ then the discounted share price will be a martingale under $\mathbb{Q}$. All that we must check is that this choice of h satisfies the necessary conditions for the Girsanov Theorem to hold. We must first verify that $\int_0^T |h_u|^2 du < \infty$ a.s. This follows from the fact that S is a.s continuous, and therefore a.s bounded on the interval $[0,T]$ for any $0 < T < \infty$. So that, for any $0 < T < \infty$, there exists a $\delta > 0$ such that

$$S_t < \delta,\ \text{for all } t \in [0,T] \text{ a.s .}$$

Now we recall that $\sigma : \mathbb{R} \to \mathbb{R}$ was assumed to be strictly positive, therefore $\frac{1}{\sigma(x)}$ is bounded on the interval $(0,\delta)$, meaning that there exists an $\epsilon > 0$ such that

$$\frac{1}{\sigma(x)} < \epsilon \text{ for all } x \in (0,\delta)$$

using this we have that

$$\begin{aligned}\int_0^T |h_u|^2 du &= \int_0^T |\frac{r-\mu}{\sigma(S_{u-b})}|^2 du \\ &< \int_0^T |(r-\mu)\epsilon|^2 du \\ &= \Big((r-\mu)\varepsilon\Big)^2 (T) < \infty.\end{aligned}$$

Next we must check whether it holds that $\mathbb{E}_P(L_t) = 1$ where

$$L_t = e^{\int_0^t h_u dW_u - \frac{1}{2}\int_0^t |h_u|^2 du},\ t \in [0,T],$$

in order to prove this we will follow the argument given in [2]. First we notice that, for every $u \in [0,T]$, h_u is known at time $u-b$ since knowing h_u only requires knowledge of the stock price up to time $u-b$. Mathematically, this means that h_u is $\mathcal{F}_{u-b}$ measurable. Therefore given the information up to time $T-b$ the integral $\int_{T-b}^T h_u dW_u$ is simply the stochastic integral of a known function since h_u is known over the whole interval $[T-b,T]$. The integral is, therefore, a Gaussian variable with mean zero and variance

$$\begin{aligned} var\Big(\int_{T-b}^T h_u dW_u\Big) &= \mathbb{E}^{\mathbb{P}}\Big[\big(\int_{T-b}^T h_u dW_u\big)^2\Big] \\ &= \int_{T-b}^T \mathbb{E}^{\mathbb{P}}(h_u)^2 du,\ \text{by the itô isometry} \\ &= \int_{T-b}^T h_u^2 du.\end{aligned}$$

By the theory of moment generating functions of normally distributed random variables we get

$$\mathbb{E}^{\mathbb{P}}\big(e^{\int_{T-b}^{T} h_u dW_u} \big| \mathcal{F}_{T-b}\big) = e^{\frac{1}{2}\int_{T-b}^{T} h_u^2 du}$$

and, as a consequence,

$$\mathbb{E}^{\mathbb{P}}\left(e^{\int_{T-b}^{T} h(u)dW(u) - \frac{1}{2}\int_{T-b}^{T} h^2(u)du} \Big| \mathcal{F}_{T-b}\right) = e^{-\frac{1}{2}\int_{T-b}^{T} h^2(u)du}\mathbb{E}^{\mathbb{P}}\left(e^{\int_{T-b}^{T} h(u)dW(u)} \Big| \mathcal{F}_{T-b}\right) = 1$$

we can use this to derive the following proposition by induction.

Proposition 4.3 *The following relation holds for all $k \in \mathbb{Z}^+$*

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}\left(e^{\int_0^T h_u dW_u - \frac{1}{2}\int_0^T h_u^2 du} \Big| \mathcal{F}_{T-kb}\right)\\ &= e^{\int_0^{T-kb} h_u dW_u - \frac{1}{2}\int_0^{T-kb} h_u^2 du} \qquad (4.19)\end{aligned}$$

Proof Let us consider the case when $k = 1$,

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}\left(e^{\int_0^T h_u dW_u - \frac{1}{2}\int_0^T h_u^2 du} \Big| \mathcal{F}_{T-b}\right)\\ &= \mathbb{E}^{\mathbb{P}}\left(e^{\int_{T-b}^T h_u dW_u + \int_0^{T-b} h_u dW_u - \frac{1}{2}\int_{T-b}^T h_u^2 du - \frac{1}{2}\int_0^{T-b} h_u^2 du} \Big| \mathcal{F}_{T-b}\right)\\ &= e^{\int_0^{T-b} h_u dW_u - \frac{1}{2}\int_0^{T-b} h_u^2 du}\mathbb{E}^{\mathbb{P}}\left(e^{\int_{T-b}^T h_u dW_u - \frac{1}{2}\int_{T-b}^T h_u^2 du} \Big| \mathcal{F}_{T-b}\right), \text{ by measurability}\\ &= e^{\int_0^{T-b} h_u dW_u - \frac{1}{2}\int_0^{T-b} h_u^2 du}. \qquad (4.20)\end{aligned}$$

Clearly the proposition holds for $k = 1$. Assume that is holds for $k = n \in \mathbb{Z}^+$. We then have that

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}\left(e^{\int_0^T h_u dW_u - \frac{1}{2}\int_0^T h_u^2 du} \Big| \mathcal{F}_{T-nb}\right)\\ &= e^{\int_0^{T-nb} h_u dW_u - \frac{1}{2}\int_0^{T-nb} h_u^2 du}\end{aligned}$$

Taking conditional expectation of the above with respect to $\mathcal{F}_{T-(n+1)b}$ gives, by the tower property,

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}\left(e^{\int_0^T h_u dW_u - \frac{1}{2}\int_0^T h_u^2 du} \Big| \mathcal{F}_{T-(n+1)b}\right)\\ &= \mathbb{E}^{\mathbb{P}}\left(e^{\int_0^{T-nb} h_u dW_u - \frac{1}{2}\int_0^{T-nb} h_u^2 du} \Big| \mathcal{F}_{T-(n+1)b}\right)\\ &= \mathbb{E}^{\mathbb{P}}\left(e^{\int_{T-(n+1)b}^{T-nb} h_u dW_u + \int_0^{T-(n+1)b} h_u dW_u - \frac{1}{2}\int_{T-(n+1)b}^{T-nb} h_u^2 du - \frac{1}{2}\int_0^{T-(n+1)b} h_u^2 du} \Big| \mathcal{F}_{T-(n+1)b}\right)\\ &= e^{\int_0^{T-(n+1)b} h_u dW_u - \frac{1}{2}\int_0^{T-(n+1)b} h_u^2 du}\mathbb{E}^{\mathbb{P}}\left(e^{\int_{T-(n+1)b}^{T-nb} h_u dW_u - \frac{1}{2}\int_{T-(n+1)b}^{T-nb} h_u^2 du} \Big| \mathcal{F}_{T-(n+1)b}\right)\\ &= e^{\int_0^{T-(n+1)b} h_u dW_u - \frac{1}{2}\int_0^{T-(n+1)b} h_u^2 du}.\end{aligned}$$

So the proposition holds for $n+1$ and therefore it holds for all $k \in \mathbb{Z}^+$ by induction.

Let $k \in \mathbb{Z}^+$ be such that $0 \leq T - kb \leq b$. We then make use of the tower property of conditional expectation again by taking expected values of equation (4.19) conditional on $\mathcal{F}_0$ remembering that $\mathcal{F}_0 \subseteq \mathcal{F}_{T-kb}$. This gives

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left(e^{\int_0^T h_u dW_u - \frac{1}{2}\int_0^T h_u^2 du} \big| \mathcal{F}_0\right) \\
&= \mathbb{E}^{\mathbb{P}}\left(e^{\int_0^{T-kb} h_u dW_u - \frac{1}{2}\int_0^{T-kb} h_u^2 du} \big| \mathcal{F}_0\right) \\
&= 1, \text{ since } T - kb \leq b.
\end{aligned}$$

Then, finally, taking unconditional expected values of the above gives

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}\left[\mathbb{E}_p\left(e^{\int_0^T h_u dW_u - \frac{1}{2}\int_0^T h_u^2 du} \big| \mathcal{F}_0\right)\right] = \mathbb{E}_P(1) \\
&\Rightarrow \mathbb{E}^{\mathbb{P}}\left[e^{\int_0^T h_u dW_u - \frac{1}{2}\int_0^T h_u^2 du}\right] = 1 \\
&\Rightarrow \mathbb{E}^{\mathbb{P}}(L_t) = 1
\end{aligned}$$

and so we have proven that h satisfies the conditions for the Girsanov Theorem and the process defined in (4.18) is, in fact, a standard Brownian motion under the measure $\mathbb{Q}$ defined by $d\mathbb{Q} = L_t d\mathbb{P}$.

This means that the dynamics of S_t under $\mathbb{Q}$ follow the SDE

$$dS_t = rS_t dt + \sigma(S_{t-b}) S_t dW_t, \ t \in [0, T]$$

which, by setting $\mu = r$ in (4.17), has the solution

$$S_t = S_s e^{r(t-s) + \int_s^t \sigma(Su-b) dW_u - \frac{1}{2}\int_s^t g^2(S_{u-b}) du} \tag{4.21}$$

for any $s, t \in [0, T]$ with $s < t$.

Clearly we have found a probability measure under which the discounted stock price is a martingale. It was also shown, in the previous subsection, that a model of this form is complete, which implies that our martingale measure is unique. We can, therefore, use the martingale pricing formula (see, for example, Theorem 10.19 in [3]) to deduce that the unique price of any T-claim paying $\Psi(S_T)$ (for some integrable $\Psi(\cdot)$) at time t must be of the form

$$V(t, S_t) = \mathbb{E}^{\mathbb{Q}}\big(e^{-r(T-t)}\Psi(S_T)|\mathcal{F}_t\big).$$

For a European call option with strike K and maturity T, $\Psi(S_T) = (S_T - K)^+$. We, therefore, have that

$$V(t, S_t) = \mathbb{E}^{\mathbb{Q}}\big(e^{-r(T-t)}(S_T - K)^+|\mathcal{F}_t\big). \tag{4.22}$$

Let us consider the case, as in [2], when $t \in [T-b, T]$, in other words, we are trying to price a call option within the last observation window. From equation (4.21)

$$S_T = S_t e^{r(T-t)+\int_t^T \sigma(S_{u-b}))dW_u - \frac{1}{2}\int_t^T \sigma^2(S_{u-b})du}$$

Obviously S_t is $\mathcal{F}_t$-measurable. We also have that, because of the limits of the integrals and the assumption that $t \in [T-b, T]$,

$$\begin{aligned} &t \leq u \leq T \\ \Rightarrow &t - b \leq u - b \leq T - b \leq t. \end{aligned}$$

This implies that $-\frac{1}{2}\int_t^T \sigma^2(S_{u-b}))du$ is also $\mathcal{F}_t$-measurable since it is a Riemann integral of a known function. The fact that S_{u-b} is a known over $[t, T]$ also implies that $\sigma(S_{u-b})$ is $\mathcal{F}_t$-measurable and so the integral $\int_t^T \sigma(S_{u-b})dW_u$, given $\mathcal{F}_t$, has a normal distribution with mean zero and variance $\Sigma^2 = \int_t^T \sigma^2(S_{u-b})du$. If we take z to represent a standard normal variable and define $m := -\frac{1}{2}\int_t^T \sigma^2(S_{u-b})du$, $x := S_t$ then (4.22) implies that

$$\begin{aligned} V(t, x) &= \mathbb{E}^{\mathbb{Q}}\big(e^{-r(T-t)}(xe^{r(T-t)+\Sigma z+m} - K)^+|\mathcal{F}_t\big) \\ &= \frac{e^{-r(T-t)}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\big(xe^{r(T-t)+\Sigma z+m} - K\big)^+ e^{-\frac{z^2}{2}}dz \end{aligned} \tag{4.23}$$

Now, consider the inequality

$$\begin{aligned} &xe^{r(T-t)+\Sigma z+m} - K \geq 0 \\ \Rightarrow &z \geq \frac{\ln\left(\frac{K}{x}\right) - r(T-t) - m}{\Sigma} := a \end{aligned}$$

which implies, by (4.23), that

$$\begin{aligned} V(t, x) &= \frac{e^{-r(T-t)}}{\sqrt{2\pi}}\int_a^{\infty}\big(xe^{r(T-t)+\Sigma z+m} - K\big)e^{-\frac{z^2}{2}}dz \\ &= \frac{e^{-r(T-t)}}{\sqrt{2\pi}}\left(xe^{r(T-t)+m}\int_a^{\infty}e^{\Sigma z}e^{-\frac{z^2}{2}}dz - K\int_a^{\infty}e^{-\frac{z^2}{2}}dz\right) \\ &= \frac{e^{-r(T-t)}}{\sqrt{2\pi}}\left(xe^{r(T-t)+m+\frac{\Sigma^2}{2}}\int_a^{\infty}e^{-\frac{(z-\Sigma)^2}{2}}dz - K\int_a^{\infty}e^{-\frac{z^2}{2}}dz\right) \\ &= xe^{m+\frac{\Sigma^2}{2}}\Phi(-a+\Sigma) - e^{-r(T-t)}K\Phi(-a) \end{aligned}$$

so that, finally, we have,

$$V(t, S_t) = S_t\Phi(d_1) - e^{-r(T-t)}K\Phi(d_2) \tag{4.24}$$

where,

$$d_1 = \frac{\ln\left(\frac{S_t}{K}\right) + r(T-t) + \frac{1}{2}\int_t^T \sigma^2(S_{u-b})du}{\sqrt{\int_t^T \sigma^2(S(u-b))du}} \tag{4.25}$$

$$d_2 = \frac{\ln\left(\frac{S_t}{K}\right) + r(T-t) - \frac{1}{2}\int_t^T \sigma^2(S_{u-b})du}{\sqrt{\int_t^T \sigma^2(S(u-b))du}}. \tag{4.26}$$

We have found a closed form solution for the price of a European call option when $t \in [T-b, T]$, this is the main appeal of this model, and it demonstrates that under certain conditions the classical theory of option pricing can be preserved in path-dependent volatility models.There are, however, some issues.

If we again assumed that $t \in [T-b, T]$, and approached the problem by constructing a self-financing replicating portfolio consisting of the risk-free and risky asset (or alternatively use the Feynmann-Kac formula and (4.22)) we would have arrived at the conclusion that the discounted price of the contingent claim, F, must obey the following Black-Scholes style PDE [2]

$$\frac{\partial F(t,x)}{\partial t} = -\frac{1}{2}\sigma(S_{t-b})^2x^2\frac{\partial^2 F(t,x)}{\partial x^2} - rx\frac{\partial F(t,x)}{\partial x} + rF(t,x) \tag{4.27}$$

$$F(T,x) = (x-K)^+ x > 0. \tag{4.28}$$

If one wanted to construct a numerical scheme for the above terminal value problem the challenges are obvious. The dependence of the problem on S_{t-b} makes it impossible to set up a discrete scheme on some fixed domain in $\mathbb{R}^2$. In addition to this, when we consider the problem when $t \in [0, T-b]$ it becomes impossible to derive a similar terminal value problem because in this region the solution is "anticipating" or "forward looking" with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$, meaning it is not adapted to the filtration. The goal of this dissertation, as is clear from the title, is to identify and study models which allow for a dependence on the past but still preserve partial differential equation approach to the pricing of derivatives. With that goal in mind, this simple case has shown that models of the form of (4.4)-(4.5) and the theory of stochastic functional differential equations will not serve our purpose. We will, therefore, not consider another model of this form in the remaining sections.

4.3 The Hobson and Rogers model

Before beginning with our next model we will try to understand the possible thought process and inspiration behind its creation. Again, we recall the purpose of this dissertation: to identify a model that allows the volatility to depend on past values, while still preserving the PDE approach to derivative pricing. A key step in the PDE approach is when we assume that the price on a contingent claim at time t has the form $F(t, S_t)$. This amounts to assuming that the contingent claim admits a Markovian realisation and this assumption will only hold if the underlying process has the Markov property.

Clearly then, if we wish to conserve the PDE approach, it is crucial that our underlying model has the Markov property. We can motivate the intuition of what follows with the following simple example: suppose we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a discrete (possibly non-Markovian) process $X = \{X_n, n = 1, 2, \dots\}$ taking values in $\mathbb{R}$. We let the filtration $\big(\mathcal{F}_t, t \in \mathbb{N}\big)$ be the filtration generated by X so that X is $(\mathcal{F}_t)$ adapted. Now, if we define the process $Y = \big\{Y_n = \big(X_{n-1}, X_{n-2}, \dots, X_1\big), n = 1, 2, \dots\big\}$ then the joint process $\big(X, Y\big)$ has the following property

$$\begin{aligned}\mathbb{E}\big((X_n, Y_n)|\mathcal{F}_{n-1}\big) &= \mathbb{E}\big((X_n, Y_n)|X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_1 = x_1\big)\\ &= \mathbb{E}\big((X_n, Y_n)|X_{n-1} = x_{n-1}, Y_{n-1} = y_{n-1}\big),\ \text{where } y_{n-1} = (x_{n-2}, \dots, x_1)\\ &= \mathbb{E}\big((X_n, Y_n)|(X_{n-1}, Y_{n-1})\big).\end{aligned}$$

This shows that the joint process is Markov. What we have illustrated is that, by defining a random variable that summarises the entire path of some (possibly non-Markovian) process, we have constructed a joint process that has the Markov property.

The question becomes, how do we implement the above logic for a continuous time process describing a stock price? This question was first answered by David Hobson and Leonard Rogers in [27]. Their results, and a discussion thereof, will now be presented.

4.3.1 The HR model

The model starts as follows: suppose we have a stochastic process S_t representing the price of a risky asset at time t, and a risk-free asset B that earns interest at a constant rate r. We define the discounted log price of the asset as $Z_t = \ln\big(e^{-rt}S_t\big)$. Next, in what is the crucial innovation of this model, we define the deviation function of order m, denoted $D_t^{(m)}$, by

$$D_t^{(m)} := \int_{-\infty}^{t} \lambda e^{\lambda(u-t)}(Z_t - Z_u)^m du. \tag{4.29}$$

Here λ is a constant parameter, and so $D^{(m)}$ represents the exponentially weighted m^{th} moment of the historical log price. λ essentially describes the rate at which past information is forgotten. The next logical step will be to make some assumptions concerning the dynamics of our stock price, in line with more standard models we assume that the discounted log price solves the SDE

$$dZ_t = \mu(D_t^{(1)}, \ldots, D_t^{(n)})dt + \sigma(D_t^{(1)}, \ldots, D_t^{(n)})dW_t \tag{4.30}$$

for some $n \in \mathbb{N}$. This is the dynamics under a measure $\mathbb{P}$. Where $\sigma(\cdot)$ and $\mu(\cdot)$ are Lipschitz functions taking values in $\mathbb{R}$, $\sigma(\cdot)$ is strictly positive and bounded, and W is a standard Wiener process. A key feature is that there is still only one driving Wiener process. This, along with the fact that we have one tradable risky asset, means that the model will retain completeness. Next we prove a crucial result that appears as Lemma 3.1 in [27].

Theorem 4.4 $(Z_t, D_t^{(1)}, \ldots, D_t^{(n)})$ *forms a Markov process. The* m^{th} *deviation process* $D_t^{(m)}$ *satisfies the coupled SDEs*

$$dD_t^{(m)} = mD_t^{(m-1)}dZ_t + \frac{m(m-1)}{2}D_t^{(m-2)}d\langle Z\rangle_t - \lambda D_t^{(m)}dt \tag{4.31}$$

Proof 4.4 Let us begin by proving the second part of the lemma, the first will follow form this. We can write $D_t^{(m)}$ in the following form

$$\begin{aligned} e^{\lambda t}D_t^{(m)} &= \int_{-\infty}^{t} \lambda e^{\lambda(u)}(Z_t - Z_u)^m du \\ &= \sum_{k=0}^{m} \binom{m}{k} (Z_t)^k \int_{-\infty}^{t} \lambda e^{\lambda u}(-Z_u)^{m-k}du. \end{aligned} \tag{4.32}$$

Then, by taking differentials on both the left and right of equation (4.32) we have

$$d\big(e^{\lambda t}D_t^{(m)}\big) = \lambda e^{\lambda t}D_t^{(m)}dt + e^{\lambda t}dD_t^{(m)} \tag{4.33}$$

and

$$
\begin{aligned}
&d\bigg(\sum_{k=0}^{m}\binom{m}{k}(Z_t)^k\int_{-\infty}^{t}\lambda e^{\lambda u}(-Z_u)^{m-k}du\bigg)\\
=&\sum_{k=0}^{m}\binom{m}{k}d\bigg((Z_t)^k\int_{-\infty}^{t}\lambda e^{\lambda u}(-Z_u)^{m-k}du\bigg)\\
=&\sum_{k=0}^{m}\binom{m}{k}\bigg\{d\big((Z_t)^k\big)\int_{-\infty}^{t}\lambda e^{\lambda u}(-Z_u)^{m-k}du+(Z_t)^k d\bigg(\int_{-\infty}^{t}\lambda e^{\lambda u}(-Z_u)^{m-k}du\bigg)\\
&+d\big((Z_t)^k\big)d\bigg(\int_{-\infty}^{t}\lambda e^{\lambda u}(-Z_u)^{m-k}du\bigg)\bigg\}\\
=&\sum_{k=0}^{m}\binom{m}{k}\bigg\{\bigg(kZ_t^{k-1}dZ_t+\frac{k(k-1)}{2}Z_t^{k-2}d\langle Z\rangle_t\bigg)\int_{-\infty}^{t}\lambda e^{\lambda u}(-Z_u)^{m-k}du\\
&+(Z_t)^k\lambda e^{\lambda t}(-Z_t)^{m-k}dt+\bigg(kZ_t^{k-1}dZ_t+\frac{k(k-1)}{2}Z_t^{k-2}d\langle Z\rangle_t\bigg)\lambda e^{\lambda t}(-Z_t)^{m-k}dt\bigg\}\\
=&\sum_{k=0}^{m}\binom{m}{k}\bigg\{\bigg(kZ_t^{k-1}dZ_t+\frac{k(k-1)}{2}Z_t^{k-2}d\langle Z\rangle_t\bigg)\int_{-\infty}^{t}\lambda e^{\lambda u}(-Z_u)^{m-k}du\\
&+(Z_t)^k\lambda e^{\lambda t}(-Z_t)^{m-k}dt+0\bigg\},\ \text{by the Itô multiplication table}\\
=&\sum_{k=0}^{m}\binom{m}{k}\bigg\{kZ_t^{k-1}\int_{-\infty}^{t}\lambda e^{\lambda u}(-Z_u)^{m-k}du\bigg\}dZ_t\\
&+\sum_{k=0}^{m}\binom{m}{k}\bigg\{\frac{k(k-1)}{2}Z_t^{k-2}\int_{-\infty}^{t}\lambda e^{\lambda u}(-Z_u)^{m-k}du\bigg\}d\langle Z\rangle_t\\
&+\bigg\{\sum_{k=0}^{m}\binom{m}{k}(-1)^{m-k}\bigg\}\lambda e^{\lambda t}Z_t^m dt\\
=&m\sum_{k=1}^{m}\binom{m-1}{k-1}\bigg\{\int_{-\infty}^{t}\lambda e^{\lambda u}(-Z_u)^{m-1-(k-1)}du\bigg\}Z_t^{k-1}dZ_t\\
&+\frac{m(m-1)}{2}\sum_{k=2}^{m}\binom{m-2}{k-2}\bigg\{\int_{-\infty}^{t}\lambda e^{\lambda u}(-Z_u)^{(m-2)-(k-2)}du\bigg\}Z_t^{k-2}d\langle Z\rangle_t\\
&+\{0\}\lambda e^{\lambda t}Z_t^m dt\\
=&e^{\lambda t}\bigg\{mD_t^{(m-1)}dZ_t+\frac{m(m-1)}{2}D_t^{(m-2)}d\langle Z\rangle_t\bigg\}. \qquad (4.34)
\end{aligned}
$$

Equating (4.33) and (4.34), multiplying by $e^{-\lambda t}$ and solving for $dD_t^{(m)}$ will give us (4.31). In order to prove that the system $(Z_t, D_t^{(1)}, \ldots D_t^{(m)})$ is a Markov process, it will be enough to show that the diffusion and drift coefficients in equation (4.31) are smooth enough for each m. Substituting the SDE for Z_t into equation (4.31) gives

$$\begin{aligned}dD_t^{(m)} =& mD_t^{(m-1)}\left(\mu(D_t^{(1)}, \ldots, D_t^{(n)})dt + \sigma(D_t^{(1)}, \ldots, D_t^{(n)})dW_t\right) \\ &+ \frac{m(m-1)}{2}D_t^{(m-2)}\left(\sigma^2(D_t^{(1)}, \ldots, D_t^{(n)})dt\right) - \lambda D_t^{(m)}dt \\ =& \left(mD_t^{(m-1)}\mu(D_t^{(1)}, \ldots, D_t^{(n)}) - \lambda D_t^{(m)} + \frac{m(m-1)}{2}D_t^{(m-2)}\sigma^2(D_t^{(1)}, \ldots, D_t^{(n)})\right)dt \\ &+ mD_t^{(m-1)}\sigma(D_t^{(1)}, \ldots, D_t^{(n)})dW_t \end{aligned} \tag{4.35}$$

since, by assumption, μ and σ are Lipschitz the drift and diffusion coefficients above will be sufficiently smooth in order to guarantee that the joint process is Markov. ∎

Another important observation of the process $D^{(m)}$ is that it is adapted to the filtration generated by Z, since knowing Z_s for all $s < t$ allows us to calculate the integral in (4.29). The filtration generated by Z will be denoted by $\mathcal{F}_t$.

Now that we have our process in a workable form and we are sure that the joint process is Markov, we can proceed with the problem of option pricing in this model.

4.3.2 Option pricing in the HR model

Let us consider the case where $n = 1$ in (4.30). Since we are only considering the first order deviation function we can adopted the notation $D^{(1)} = D$ without risking any ambiguity. Theorem 4.4 and equation (4.30) then gives us

$$dB_t = rB_tdt \tag{4.36}$$
$$dD_t = dZ_t - \lambda D_tdt \tag{4.37}$$
$$dZ_t = \mu(D_t)dt + \sigma(D_t)dW_t \tag{4.38}$$

we can then substitute (4.30) into the second of our three equations to get

$$dB_t = rB_tdt \tag{4.39}$$
$$dD_t = (\mu(D_t) - \lambda D_t)dt + \sigma(D_t)dW_t \tag{4.40}$$
$$dZ_t = \mu(D_t)dt + \sigma(D_t)dW_t. \tag{4.41}$$

Throughout the rest of this dissertation we will refer to the above set of equations as the HR model. As in [27], we will use the martingale approach to the problem of pricing a T-claim with payoff $\Phi(S_t)$. Therefore, our first goal will be to find a measure under which the discounted option price is a martingale or, equivalently, a measure under which the process S_t has a drift coefficient r. Let us consider the $\mathbb{P}$ dynamics of S_t. By the definition of Z_t we have that $S_t = e^{Z_t + rt}$and so, using the Itô formula applied to the function $f(t, z) = e^{z+rt}$ gives us

$$df(t, Z_t) = dS_t = f_1 dt + f_2 dZ_t + \frac{1}{2} f_{22} d\langle Z \rangle_t \tag{4.42}$$

$$= S_t\big(r + \mu(D_t) + \frac{1}{2}\sigma^2(D_t)\big)dt + S_t\sigma(D_t)dW_t. \tag{4.43}$$

Therefore, the appropriate Girsanov kernel for us to consider in Theorem 2.14 is the function

$$h(t) = \frac{-\mu(D_t) - \frac{1}{2}\sigma^2(D_t)}{\sigma(D_t)}$$

and we define a new measure $\mathbb{Q}$ by $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_t} = L_t$, where

$$L_t = e^{\int_0^t h(u)dW(u) - \frac{1}{2}\int_0^t |h(u)|^2 du}, \ t \in [0, T].$$

i.e

$$dL_t = h_t L_t dW_t, \text{ on } [0, T]$$
$$L_0 = 1.$$

In order to ensure that this is a feasible choice, we must consider the conditions in Theorem 2.14. We remember that the solution function will be a.s continuous and that we are considering values on the bounded interval $[0, T]$. That $\int_0^T |h(u)|du < \infty$ follows from the strict positivity of $\sigma(\cdot)$, and that the functions $\mu(\cdot)$ and $\sigma(\cdot)$ are Lipschitz. The non-explosiveness of (D_t, S_t) also guarantees that the conditions of corollary 2.18 are satisfied and that our new measure is indeed a probability measure.

Since the discounted price process is a martingale under the new measure $\mathbb{Q}$ we can use the Martingale Pricing Theorem to conclude that the time t price of our T-claim is [27]

$$V(T - t, S_t, D_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\Phi(S_t)|\mathcal{F}_t]$$

assuming that $\Phi(\cdot)$ is integrable enough. Then the Feynman-Kac formula (see for example [29] p366) implies that the solution satisfies the deterministic PDE

$$
\begin{aligned}
0 &= \big(V_1 - rSV_2 + rV + \lambda D V_3\big) \\
&\quad - \frac{\sigma^2(D)}{2}\Big(- V_3 + S^2 V_{22} + V_{33} + SV_{23}\Big) \qquad (4.44)\\
V(0,S,D) &= \Phi(S). \qquad (4.45)
\end{aligned}
$$

Clearly this equation is much more suitable for numerical methods than (4.27). In fact, it is guaranteed a unique solution under certain condition on $\Phi(\cdot)$ [27]. In a later section we construct and analyse a numerical scheme that approximates the solution of the above system, but for now we will highlight some issues.

The main concern raised regarding the HR model in the integration over an infinite time horizon. The definition of $D^{(m)}$ suggests that we have the following initial condition for $D^{(m)}$,

$$D_0^{(m)} = \int_{-\infty}^{0} \lambda e^{\lambda u}(Z_0 - Z_u)^m du$$

Since there is no conceivable situation where we have price data over an infinite time horizon, this raises obvious concerns. One possible suggestion could be to assume that the share has grown at the risk free rate from its inception up to a price Z_0. So that at any time $s < 0$ we can assume that:

$$Z_s = \ln(e^{-rs}S_0) = -rs + Z_0$$

so that

$$
\begin{aligned}
D_0^{(m)} &= \int_{-\infty}^{0} \lambda e^{\lambda u}(Z_0 - Z_0 + ru)^m du \\
&= r\lambda \int_{-\infty}^{0} e^{\lambda u} u^m du.
\end{aligned}
$$

Alternatively if share data is available up to some finite point in time before the valuation time, say $-a$ for some $a > 0$, then one could calculate the integral in the initial condition explicitly over the interval $[-a, 0]$ and use the above approximation over $[-\infty, -a]$ with $-a$ as the reference price.

This solution seems unsatisfying, and for this reason the next model we will discuss involves a clever side step of the problem.

4.4 Generalized averaging

In order to solve the issue of the integration over an infinite horizon, as well as having more flexibility in the averaging process, a more general model was proposed by Paolo Foschi and Andrea Pascucci in [17]. They introduce a weighting function, φ, that is strictly positive on $[0,T]$, integrable on $[-\infty,T]$ and non-negative on $[-\infty,T]$. In the language of measure theory, the weight of the whole space is then

$$\phi(t) = \int_{-\infty}^{t} \varphi(u)du. \tag{4.46}$$

Key in the above definition is that φ my have compact support on the interval $[-\infty,0)$ and so we may have a bounded interval of integration. Clearly there is considerably more freedom in choosing our weighting function φ here compared to the HR model in which we are restricted to exponential weightings. For consistency in notation, the discounted log price process is again denoted by Z. We then define the process

$$M_t = \frac{1}{\phi(t)} \int_{-\infty}^{t} \varphi(u)Z_u du, \; t \in [0,T] \tag{4.47}$$

which will be refereed to as the averaging process. Taking differentials we have

$$\begin{aligned} dM_t &= d\left(\frac{1}{\phi(t)}\int_{-\infty}^{t}\varphi(u)Z_u du\right) \\ &= d\left(\frac{1}{\phi(t)}\right)\int_{-\infty}^{t}\varphi(u)Z_u du + \frac{1}{\phi(t)}d\left(\int_{-\infty}^{t}\varphi(u)Z_u du\right) + d\left(\frac{1}{\phi(t)}\right)d\left(\int_{-\infty}^{t}\varphi(u)Z_u du\right) \\ &= -\frac{\phi'(t)}{\phi^2(t)}dt\int_{-\infty}^{t}\varphi(u)Z_u du + \frac{1}{\phi(t)}\varphi(t)Z_t dt - \frac{\phi'(t)}{\phi^2(t)}\varphi(t)Z_t(dt)^2 \\ &= -\frac{\varphi(t)}{\phi^2(t)}dt\int_{-\infty}^{t}\varphi(u)Z_u du + \frac{\varphi(t)}{\phi(t)}Z_t dt \\ &= \frac{\varphi(t)}{\phi(t)}(Z_t - M_t)dt. \end{aligned} \tag{4.48}$$

In order to complete our new model some assumptions must be made about the dynamics of the price process. As before, it is assumed that the discounted log price process solves the SDE

$$dZ_t = \mu(Z_t - M_t)dt + \sigma(Z_t - M_t)dW_t, \tag{4.49}$$

and again, in order to guarantee a solution to the system (4.48)-(4.49) as well as for simplification of some calculations we assume that $\mu(\cdot)$, $\sigma(\cdot)$ are Lipschitz continuous, bounded and that $\sigma(\cdot)$ is strictly positive [17]. This will also ensure, as is desired, that the process (Z_t, M_t) is Markovian.

An important questions to ask at this stage, and one that will be important in our quest to derive a pricing PDE, is whether or not the process

(Z_t, M_t) has a density function and, assuming that the density exists, where is it positive? In order to answer this question it will be useful to rewrite equations (4.48) and (4.49) as the following single vector equation:

$$d\begin{bmatrix} M_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \frac{\varphi(t)}{\phi(t)}(Z_t - M_t) \\ \mu(Z_t - M_t) \end{bmatrix} dt + \begin{bmatrix} 0 & 0 \\ 0 & \sigma(Z_t - M_t) \end{bmatrix} \begin{bmatrix} dW_t^1 \\ dW_t^2 \end{bmatrix} \tag{4.50}$$

where W^1 and W^2 are independent Wiener processes. This is clearly in the form of equation (2.8), and it can be shown that this equation satisfies the parabolic Hörmander's condition. To see this, notice that in the notation of (2.8) we have that

$$V_0(x,y) = \begin{bmatrix} \frac{\varphi(t)}{\phi(t)}(y-x) \\ \mu(y-x) \end{bmatrix} \tag{4.51}$$

$$V_1(x,y) = \begin{bmatrix} 0 \\ \sigma(y-x) \end{bmatrix} \tag{4.52}$$

We then have that

$$\begin{aligned} \mathcal{V}_0 &= \{V_1\} \\ \mathcal{V}_1 &= \mathcal{V}_0 \cup \{[V_1, V_0], [V_1, V_1]\}. \end{aligned} \tag{4.53}$$

It is obvious from the definition of the Lie bracket that $[V_1, V_1] = 0$, and we can calculate $[V_1, V_0]$,

$$\begin{aligned} [V_1, V_0] &= DV_0 V_1 - DV_1 V_0 \\ &= \begin{bmatrix} \frac{\partial V_0^1}{\partial x} & \frac{\partial V_0^1}{\partial y} \\ \frac{\partial V_0^2}{\partial x} & \frac{\partial V_0^2}{\partial y} \end{bmatrix} \begin{bmatrix} 0 \\ \sigma(y-x) \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ \frac{\partial}{\partial x}\sigma(y-x) & \frac{\partial}{\partial y}\sigma(y-x) \end{bmatrix} \begin{bmatrix} \frac{\varphi(t)}{phi(t)}(y-x) \\ \mu(y-x) \end{bmatrix} \\ &= \begin{bmatrix} -\frac{\varphi(t)}{\phi(t)} & \frac{\varphi(t)}{\phi(t)} \\ -\mu_x(y-x) & \mu_y(y-x) \end{bmatrix} \begin{bmatrix} 0 \\ \sigma(y-x) \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ -\sigma_x(y-x) & \sigma_y(y-x) \end{bmatrix} \begin{bmatrix} \frac{\varphi(t)}{\phi(t)}(y-x) \\ \mu(y-x) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\varphi(t)}{\phi(t)}\sigma(x-y) \\ \mu_y(y-x)\sigma(y-x) + \sigma_x(y-x)\frac{\varphi(t)}{\phi(t)}(y-x) - \sigma_y(y-x)\mu(y-x) \end{bmatrix} \end{aligned}$$

now we can define the vector field

$$\begin{aligned} \mathcal{V}_1(x) &= span\{V(x) : V \in \mathcal{V}_1\} \\ &= span\left\{ \begin{bmatrix} 0 \\ \sigma(y-x) \end{bmatrix}, \begin{bmatrix} \frac{\varphi(t)}{\phi(t)}\sigma(x-y) \\ \mu_y(y-x)\sigma(y-x) + \sigma_x(y-x)\frac{\varphi(t)}{\phi(t)}(y-x) - \sigma_y(y-x)\mu(y-x) \end{bmatrix} \right\} \end{aligned}$$

and since we have assumed that $\phi(t)$ is strictly positive for all t we can conclude that $\cup_{k\geq 1}\mathcal{V}_k(x) = \mathcal{V}_1(x) = \mathbb{R}^2$. This means that our SDE satisfies the parabolic Hörmander condition and therefore, by Hörmanders Theorem, the joint process (M_t, Z_t) has a smooth density on $\mathbb{R}^2$.

We have shown that the density exist and is smooth, next we want to show that this density is strictly positive on $\mathbb{R}^2$.

This can be done by first considering the density of Z. If we define a measure $\mathbb{Q}$ as in the Girsanov Theorem with $h(t) = -\frac{\mu(Z_t)}{\sigma Z_t}$, then the dynamics of Z under $\mathbb{Q}$ are

$$dZ_t = \sigma(Z_t - M_t)d\bar{W}_t.$$

This satisfies the conditions of proposition (2.11), and we can conclude that Z_t has a strictly positive density under $\mathbb{Q}$. Clearly the boundedness of $\mu(\cdot)$ and strict positivity of $\sigma(\cdot)$ ensures that $\mathbb{Q}$ is well defined.

Regarding the density of M_t we notice, for any $t > 0$ and $b > 0$, that

$$\begin{aligned}\mathbb{Q}\big(|M_t| > b\big) &= \mathbb{Q}\bigg(\Big|\frac{1}{\phi(t)}\int_{-\infty}^{t} \varphi(u) Z_u du\Big| > b\bigg)\\ &\geq \mathbb{Q}\bigg(\Big|\frac{Z_m}{\phi(t)}\int_{-\infty}^{t} \varphi(u) du\Big| > b\bigg), \text{ where } Z_m = \inf_{0\leq s\leq t} Z_s\\ &= \mathbb{Q}(|Z_m| > b) > 0. \end{aligned} \tag{4.54}$$

The final step is a result of the previous argument holding for any $t > 0$, and in particular it holds for times where the infimum is achieved.

By construction $\mathbb{Q} << \mathbb{P}$ and so if an event has a non-zero probability under $\mathbb{Q}$ it must have a non-zero probability under $\mathbb{P}$, and we have shown that (Z_t, M_t) has a strictly positive density under both measures. This fact shall be used later. For now, we consider some convenient notation.

Clearly an important quantity in the above model is the process describing the deviation from the mean, $D_t := Z_t - M_t$. Using this definition for D, the definition of Z, and the Itô formula, the system can be written in the form

$$dD_t = \Big(\mu(D_t) - \frac{\varphi(t)}{\phi(t)} D_t\Big)dt + \sigma(D_t)dW_t \tag{4.55}$$

$$dS_t = S_t\Big(r + \mu(D_t) + \frac{1}{2}\sigma^2(D_t)\Big)dt + S_t\sigma(D_t)dW_t. \tag{4.56}$$

At this point it is worth noting that choosing $\varphi(t) = e^{\lambda t}$ will give us that $\frac{\varphi(t)}{\phi(t)} = \lambda$, and as a consequence this system of SDEs reduces to the system described in (4.37)-(4.38). Another useful averaging function is $\varphi(t) = 1$

for $t \in [0, T]$ and zero elsewhere, which will give a geometric average and is useful for the pricing of Asian options.

As we have done before, after studying the underlying properties of our model, we will consider the problem of pricing and hedging a T-claim.

4.4.1 A Classical approach to option pricing

In all the preceding cases in this chapter we have obtained a pricing formula by finding a martingale measure and then making use of the martingale pricing theory. In this instance we will consider a more classical approach by considering a self financing replicating portfolio. This is not only for the sake of variety, but it also highlights why the Markovian property of this model is so appealing.

The approach deviates slightly from what is done in [17] in that we consider the state variables (S_t, D_t), whereas in [17] they use the pair (Z_t, M_t). This approach is preferred since the variables representing the price process and deviation from the mean are more intuitively clear and instructive than those representing the discounted log price and the mean function.

As is the case in the derivation of the seminal Black-Scholes equation, we start with a self financing portfolio V_t that, at time t, consists of α_t shares and β_t of the risk free asset, meaning

$$V_t = \alpha_t S_t + \beta_t B_t.$$

The self-financing condition then implies that

$$dV_t = \alpha_t dS_t + \beta_t dB_t.$$

Substituting the dynamics of both the risk-free asset and risky asset, as described in (4.36) and (4.56) respectively, we get

$$dV_t = \left(\alpha_t S_t r + \alpha_t S_t \mu(D_t) + \alpha_t S_t \frac{1}{2}\sigma^2(D_t) + r\beta_t B_t\right)dt + \alpha_t S_t \sigma(D_t) dW_t.$$

Using the definition of V_t we have that $r\beta_t B_t = r(V_t - \alpha_t S_t)$ substituting this into the above SDE gives

$$\begin{aligned} dV_t =& \left(\alpha_t S_t r + \alpha_t S_t \mu(D_t) + \alpha_t S_t \frac{1}{2}\sigma^2(D_t) + r(V_t - \alpha_t S_t)\right)dt \\ &+ \alpha_t S_t \sigma(D_t) dW_t. \end{aligned} \tag{4.57}$$

Now, let us consider the price of a T-claim at time t denoted by F_t. We assume that the price is of the form $F_t = F(t, S_t, D_t)$, this is a safe assumption precisely because of the Markov nature of our process (4.55)-(4.56). We can then use the generalised Itô formula, along with (4.55)-(4.56), to show that

$$\begin{aligned}
dF_t &= F_1 dt + F_2 dS_t + F_3 dD_t + \frac{1}{2}(F_{22} d\langle S\rangle_t + 2F_{23} dS_t dD_t + F_(33) d\langle D_t\rangle_t) \\
&= \left\{ V_1 + V_2(r + \mu(D_t) + \frac{1}{2}\sigma^2(D_t))S_t + V_3(\mu - \frac{\varphi(t)}{\phi(t)}D_t) + \frac{1}{2}V_{22}S_t^2\sigma^2(D_t) \right. \\
&\left. + V_{23}S_t\sigma^2(D_t) + \frac{1}{2}V_{33}\sigma^2(D_t) \right\} dt \\
&+ (V_2 S_t \sigma(D_t) + V_3\sigma(D_t))dW_t. \qquad (4.58)
\end{aligned}$$

At this point we make the assumption that the self-financing portfolio V_t replicates the option price process F_t. We also assume that it is of the form $V_t = \alpha(t, S_t, D_t)S_t + \beta(t, S_t, D_t)dW_t$, again this is a safe assumption since the process (S_t, D_t) is Markov. Since V replicates F we have that $V_t = F_t$, and as a consequence $dV_t = dF_t$. By the uniqueness of the Itô expansion we can equate the drift and diffusion coefficients in the right hand sides of equations (4.57) and (4.58) to get the set off equations, that hold for each $t \in [0, T]$

$$V_2 S_t\sigma(D_t) + V_3\sigma(D_t) = \alpha(t, S_t, D_t)S_t\sigma(D_t) \qquad (4.59)$$

$$\begin{aligned}
&V_1 + V_2(r + \mu(D_t) + \frac{1}{2}\sigma^2(D_t))S_t + V_3(\mu - \frac{\varphi(t)}{\phi(t)}D_t) + \frac{1}{2}V_{22}S_t^2\sigma^2(D_t) \\
&+ V_{23}S_t\sigma^2(D_t) + \frac{1}{2}V_{33}\sigma^2(D_t) = \alpha(t, S_t, D_t)S_t r + \alpha(t, S_t, D_t)S_t\mu(D_t) \\
&+ \alpha(t, S_t, D_t)S_t\frac{1}{2}\sigma^2(D_t) + r(V_t - \alpha(t, S_t, D_t)S_t). \qquad (4.60)
\end{aligned}$$

At this point we recall the assumption that $\sigma(\cdot)$ is strictly positive, this allows us to divide through by $\sigma(D_t)$ and solve for α_t in (4.59) to get

$$\alpha(t, S_t, D_t) = \frac{V_2(t, S_t, D_t)S_t + V_3(t, S_t, D_t)}{S_t} \text{ for each } t \in [0, T]. \qquad (4.61)$$

We can then substitute this expression into (4.60) to get the equation

$$\begin{aligned}
&V_1(t, S_t, D_t) + V_2(t, S_t, D_t)(r + \mu(D_t) + \frac{1}{2}\sigma^2(D_t))S_t + V_3(t, S_t, D_t)(\mu - \frac{\varphi(t)}{\phi(t)}D_t) \\
&+ \frac{1}{2}V_{22}(t, S_t, D_t)S_t^2\sigma^2(D_t) + V_{23}(t, S_t, D_t)S_t\sigma^2(D_t) + \frac{1}{2}V_{33}(t, S_t, D_t)\sigma^2(D_t) = \\
&\big(V_2(t, S_t, D_t)S_t + V_3(t, S_t, D_t)\big)r + \big(V_2(t, S_t, D_t)S_t + V_3(t, S_t, D_t)\big)\mu(D_t) \\
&+ \big(V_2(t, S_t, D_t)S_t + V_3(t, S_t, D_t)\big)\frac{1}{2}\sigma^2(D_t) + r(V - V_2(t, S_t, D_t)S_t - V_3(t, S_t, D_t)).
\end{aligned}$$

Simplification gives,

$$\begin{aligned} &V_1(t,S_t,D_t)+rV_2(t,S_t,D_t)S_t-rV(t,S_t,D_t)-\frac{\varphi(t)}{\phi(t)}D_tV_3(t,S_t,D_t)\\ &+\frac{\sigma^2(D_t)}{2}\Big(V_{22}(t,S_t,D_t)S_t^2+2V_{23}S_t+V_{33}(t,S_t,D_t)-V_3(t,S_t,D_t)\Big)=0. \end{aligned}$$

We have shown that the pair (S_t, D_t) has a strictly positive transition density over $\mathbb{R}^2$, and so the above equation must hold for all real pairs $(s, d) \in \mathbb{R}^2$ and for all $t \in [0, T]$. We therefore consider V to be of the form $V = V(t, s, d)$. Finally then, for a T-claim with payoff $\Psi(S_T, D_T)$, we arrive at the following deterministic terminal value problem

$$\begin{aligned} &V_1+rV_2s-rV-\frac{\varphi(t)}{\phi(t)}dV_3+\frac{\sigma^2(d)}{2}\Big(V_{22}s^2+2V_{23}s+V_{33}-V_3\Big)=0\\ &V(T,s,d)=\Psi(s,d). \end{aligned}$$

If we instead solve this equation over the backward time $\tau := T - t$, then it is in the numerically convenient form of an initial value problem in $\mathbb{R}^2$,

$$\begin{aligned} &V_1-rV_2s+rV+\frac{\varphi(t)}{\phi(t)}V_3d-\frac{\sigma^2(d)}{2}\Big(V_{22}s^2+2V_{23}s+V_{33}-V_3\Big)=0\\ &V(0,s,d)=\Psi(s,d). \end{aligned} \tag{4.62}$$

The existence and uniqueness of a solution to (4.62) equation is guaranteed by results found in the study of Kolmogorov type degenerate parabolic partial differential equations [17].

We recall that choosing $\phi(t) = e^{\lambda t}$ implies that $\frac{\varphi(t)}{\phi(t)} = \lambda$, and in this case the system (4.62) reduces exactly to the Hobson and Rogers PDE described in (4.44). It is also worth noting that if we restrict V to only depend on s and t, and set $\sigma(\cdot)$ to a positive constant constant σ then (4.62) reduces to

$$\begin{aligned} &V_1-rV_2s+rV-\frac{\sigma^2}{2}V_{22}s^2=0\\ &V(0,s,d)=\Psi(s,d), \end{aligned}$$

which is the familiar Black-Scholes PDE for the price of a T-claim, so we have comfort that the new model agrees with the traditional theory. Upon reflection, we have managed to show that classical theory and methods can be preserved under the introduction of a dependency on the past if this is done in a certain way. As a result, we only need to slightly adapt the well understood numerical techniques used to solve pricing problems under classical theory. We have managed to introduce past price information without sacrificing much, a clear strength of this model.

In the next chapter we seek to tie together what was done here and the work presented in the chapter on local volatility.

5 A fitted model

The preceding work has highlighted the numerous advantages of path-dependent volatility models, they give rich volatility dynamics, completeness is not lost and some models yield a pricing PDE for a contingent claim.

We now turn our attention to the problem of fitting these models to the observed market data. We have seen that, given a continuum of European call prices, we can find the unique local volatility function that produces these prices when used in a diffusion model. This local volatility was also shown to be the conditional expectation of some more complicated process.

Inspired by the work of Julien Guyon [20] and Foschi [17], we propose the following stock price model

$$dM_t = \frac{\varphi(t)}{\phi(t)}\Big(Z_t - M_t\Big)dt \tag{5.1}$$

$$dZ_t = \mu dt + \sigma(Z_t - M_t)\ell(t, Z_t)dW_t \tag{5.2}$$

where $Z_t := \ln(e^{-rt}S_t)$ is the discounted log price. The function $\ell : [0,T] \times \mathbb{R} \to \mathbb{R}$ is called the leverage function, and we assume that it is such that the model (5.1),(5.2) produces option prices that exactly fit the observed market smile. M_t is defined as in the Foschi model [17], and μ is some constant representing the drift.We also assume that the function $\sigma : \mathbb{R} \to \mathbb{R}$ is strictly positive, that it satisfies the Lipschitz and standard growth condition, and that it is bounded as in [17].

We will return to the question of how to calculate, or at least estimate, the leverage function, but for now we examine the pricing dynamics that this model implies.The first important point to consider is whether the joint process (M_t, Z_t) has the Markov property, as in the case of [17]. This would be guaranteed if the coefficient functions $f(t, S_t, M_t) := \mu$ and $g(t, S_T, M_t) := \sigma(Z_t - M_t)\ell(t, Z_t)$ satisfied the usual Lipschitz and growth conditions. That f satisfies this is obvious, proving this for the function $g : [0,T] \times \mathbb{R}^2 \to \mathbb{R}$ is more difficult. We start by assuming that the function $\ell : [0,T] \times \mathbb{R} \to \mathbb{R}$ is itself strictly positive, Lipschitz and bounded. This seems like a restrictive and arbitrary move at this point, however we will motivate this assumption when we return to the problem of calculating the leverage function.

With this assumption we have that our function g is a product of two Lipschitz functions, we then use the following fact

Proposition 5.1 *Assume that the functions $\ell : [0,T] \times \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}$, are both Lipschitz in the second argument, and that they are bounded. In* *addition, assume that the function σ, satisfies the standard growth condition. This implies there is a $C < \infty$ such that*

$$|\sigma(x) - \sigma(y)| \leq C|x - y|, \ \forall x, y \in \mathbb{R} \tag{5.3}$$
$$|\ell(t,x) - \ell(t,y)| \leq C|x - y|, \ \forall x, y \in \mathbb{R} \text{ and } t\epsilon[0,T] \tag{5.4}$$
$$|\sigma(x)| \leq C(1 + |x|), \ \forall x, y \in \mathbb{R} \tag{5.5}$$
$$|\sigma(x)| \leq C, \ \forall x, y \in \mathbb{R} \tag{5.6}$$
$$|\ell(t,x)| \leq C, \ \forall x, y \in \mathbb{R} \text{ and } t \in [0,T]. \tag{5.7}$$

Then the function $g : [0,T] \times \mathbb{R} \to \mathbb{R}$, defined by $g(t,x) = \sigma(x)\ell(t,x)$, is Lipschitz in the second argument and satisfies the standard growth condition.

Proof Let $x, y \in \mathbb{R}$ be arbitrary, then for any $t \in [0,T]$ we have,

$$\begin{aligned}
|g(t,x) - g(t,y)| &= |\sigma(x)\ell(t,x) + \sigma(y)\ell(t,y)| && (5.8)\\
&= |\sigma(x)\ell(t,x) + \sigma(y)\ell(t,x) \\
&\quad - \sigma(y)\ell(t,x) + \sigma(y)\ell(t,y)| && (5.9)\\
&= |\ell(t,x)(\sigma(x) - \sigma(y)) + \sigma(y)(\ell(t,y) - \ell(t,x))| && (5.10)\\
&\leq |\ell(t,x)||(\sigma(x) - \sigma(y))| + |\sigma(y)||(\ell(t,y) - \ell(t,x))| && (5.11)\\
&\leq C^2|x - y| + C^2|y - x| && (5.12)\\
&= 2C^2|x - y|. && (5.13)
\end{aligned}$$

Where, in line (5.12), we used the boundedness and the Lipschitz assumptions. So we have shown that the g is Lipschitz. That g satisfies the standard growth condition is obvious since, from (5.5) and (5.6),

$$|g(t,x)| = |\sigma(x)\ell(t,x)| \leq C^2(1 + |x|) \tag{5.14}$$

∎

So, under certain assumptions, the model (5.1),(5.2) has a Markovian solution. This will allow us to safely assume that the price process for a contingent claim has a Markovian realisation. We can, therefore, safely assume that the value of any contingent claim at time t has the form $V_t = V(t, Z_t, M_t)$.

5.1 Option pricing theory: A classical approach

The goal of this subsection will be to derive a pricing PDE for a contingent claim with a payoff $V_t = \Psi(S_T, D_T)$. Where we define the quantity $D_t := Z_t - M_t$, and S_t is the price of the underlying at time t. D_t will be referred to as the deviation from the mean. Clearly we will need the dynamics for

S_t and D_t. Remembering that $Z_t = \ln(e^{-rt}S_t) \Rightarrow S_t = e^{Z_t+rt}$, we define the function $h(t,x) = e^{rt+x}$. This gives

$$\frac{\partial h}{\partial t} := h_1 = rh \tag{5.15}$$

$$\frac{\partial h}{\partial x} := h_2 = h \tag{5.16}$$

$$\frac{\partial^2 h}{\partial x^2} := h_{22} = h \tag{5.17}$$

$$, \tag{5.18}$$

and so by the Itô formula

$$dS_t = dh(t,Z_t) = h_1(t,Z_t)dt + h_2(t,Z_t)dZ_t + h(t,Z_t)_{22}(dZ_t)^2 \tag{5.19}$$

$$= \big(h_1 + \mu h_2 + \frac{1}{2}\sigma^2(D_t)\ell^2(t,S_t)h_{22}\big)dt + \sigma(D_t)\ell(t,S_t)h_2 dW_t \tag{5.20}$$

$$= \big(rh + \mu h + \frac{1}{2}\sigma^2(D_t)\ell^2(t,S_t)h\big)dt + \sigma(D_t)\ell(t,S_t)h dW_t \tag{5.21}$$

$$\Rightarrow dS_t = \left(r + \mu + \frac{1}{2}\sigma^2(D_t)\ell^2(t,S_t)\right)S_t dt + \sigma(D_t)\ell(t,S_t)S_t dW_t. \tag{5.22}$$

For D_t we have

$$dD_t = dZ_t - dM_t \tag{5.23}$$

$$\Rightarrow dD_t = \left(\mu - \frac{\varphi(t)}{\phi(t)}D_t\right)dt + \sigma(D_t)\ell(t,S_t)dW_t. \tag{5.24}$$

Equations (5.22) and (5.24) give the dynamics of the Markov process (S_t, D_t), we now assume that the time t price of a contingent claim has the form $V_t = V(t, S_t, D_t)$. By the Itô formula the dynamics of this process are

$$dV = V_1 dt + V_2 dS_t + V_3 dD_t$$

$$+ \frac{1}{2}\Big(V_{22}(dS_t)^2 + 2V_{23}(dD_t)(dS_t) + V_{33}(dD_t)^2\Big) \tag{5.25}$$

$$= \Big(V_1 + V_2(r + \mu + \frac{1}{2}\sigma^2\ell^2)S_t + V_3(\mu - \frac{\phi}{\Phi}D_t) + \frac{1}{2}V_{22}S_t^2\sigma^2\ell^2 + V_{23}S_t\sigma^2\ell^2 + \frac{1}{2}V_{33}\sigma^2\ell^2\Big)dt +$$

$$+ \Big(V_2 S_t \sigma\ell + V_3\sigma\ell\Big)dW_t, \tag{5.26}$$

where we have dropped the arguments of $\sigma(D_t)$, $\ell(t,S_t)$, $\phi(t)$, $\Phi(t)$ and the derivatives of V, for example $V_1(t,S_t)$, for brevity. Now consider a self-financing portfolio $\Pi_t = \alpha(t,S_t,D_t)S_t + \beta(t,S_t,D_t)B_t$ with $\alpha(t,S_t,D_t)$ shares and $\beta(t,S_t,D_t)$ risk free bonds, i.e the asset with dynamics $dB_t = rB_t dt$. The self financing condition implies that

$$d\Pi = \alpha_t dS_t + \beta_t dB_t \tag{5.27}$$

$$= (\alpha_t\big(r + \mu + \frac{1}{2}\sigma^2(D_t)\ell^2(t,S_t)\big)S_t + r\beta_t B_t)dt$$

$$+ \alpha_t\sigma(D_t)\ell(t,S_t)S_t dW_t, \tag{5.28}$$

where we have dropped the arguments of $\alpha(t, S_t, D_t)$ and $\beta(t, S_t, D_t)$ for brevity. We now assume that the above portfolio perfectly replicates the contingent claim so that $\Pi_t = V_t$ and $d\Pi_t = dV_t$. By the uniqueness of the representation of an Itô process and equations (5.26) and (5.28) we have

$$\alpha_t\bigg(r + \mu + \frac{1}{2}\sigma^2(D_t)\ell^2(t, S_t)\bigg)S_t + r\beta_t B_t = \tag{5.29}$$
$$V_1 + V_2\bigg(r + \mu + \frac{1}{2}\sigma^2\ell^2\bigg)S_t + V_3(\mu - \frac{\phi}{\Phi}D_t) + \frac{1}{2}V_{22}S_t^2\sigma^2\ell^2 + V_{23}S_t\sigma^2\ell^2 + \frac{1}{2}V_{33}\sigma^2\ell^2$$
$$\alpha(t, S_t, D_t)\sigma(D_t)\ell(t, S_t)S_t = V_2 S_t\sigma(D_t)\ell(t, S_t) + V_3\sigma(D_t)\ell(t, S_t). \tag{5.30}$$

Since, by assumption, we had that $\phi(t) > 0$ for all $t \in [0, T]$, all the conditions for the Hörmander's Theorem are satisfied and the process (S_t, D_t) has a strictly positive density on $\mathbb{R}^2$ for $t > 0$. Therefore equation (5.30) holds for all real numbers, also remembering that $\sigma(\cdot)$ and $\ell(t, \cdot)$ are nowhere 0, we have that

$$\alpha(t, s, d) = \frac{V_2(t, s, d)s + V_3(t, s, d)}{s}, \tag{5.31}$$

using this, as well as the fact that $\beta(t, S_t, D_t)rB_t = r(V_t - \alpha(t, S_t, D_t)S_t)$, equation (5.29) gives

$$(V_2S_t + V_3)\big(r + \mu + \frac{1}{2}\sigma^2(D_t)\ell^2(t, S_t)\big) + r(V - V_2S_t - V_3) = \tag{5.32}$$
$$V_1 + V_2(r + \mu + \frac{1}{2}\sigma^2\ell^2)S_t + V_3(\mu - \frac{\phi}{\Phi}D_t) + \frac{1}{2}V_{22}S_t^2\sigma^2\ell^2 + V_{23}S_t\sigma^2\ell^2 + \frac{1}{2}V_{33}\sigma^2\ell^2$$
$$\Rightarrow V_1 + rS_tV_2 - \frac{\phi(t)}{\Phi(t)}D_tV_3 - rV$$
$$+ \frac{\sigma^2(D_t)\ell^2(t, S_t)}{2}\Big(S_t^2V_{22} + 2S_tV_{23} + V_{33} - V_3\Big) = 0. \tag{5.33}$$

Now, if we define the solution as a function of the backward time $\tau := T - t$ then we have the following deterministic initial value problem

$$V_1 - rsV_2 + \frac{\phi(t)}{\Phi(t)}dV_3 + rV$$
$$- \frac{\sigma^2(d)\ell^2(t, s)}{2}\Big(s^2V_{22} + 2sV_{23} + V_{33} - V_3\Big) = 0 \tag{5.34}$$
$$V(0, s, d) = \Psi(s, d). \tag{5.35}$$

This describes the evolution of the price of a contingent claim paying $\Psi(S_T, D_T)$ at maturity time T, through backward time. If we choose $\phi(t) = e^{\lambda t}$ in equation (5.1) our model reduces to the Hobson and Rogers model with order 1 offset [17]. To see that this is indeed the case, we notice that for this choice of $\phi(t)$ we have that $\frac{\phi(t)}{\Phi(t)} = \lambda$ and the initial value problem (5.34)-(5.35) reduces to

$$
\begin{aligned}
&V_1 - rsV_2 + \lambda d V_3 + rV \\
&- \frac{\sigma^2(d)\ell^2(t,s)}{2}\Big(s^2 V_{22} + 2sV_{23} + V_{33} - V_3\Big) = 0. && (5.36)\\
&V(0,s,d) = \Psi(s,d), && (5.37)
\end{aligned}
$$

which agrees exactly with the PDE found in the original paper by Hobson and Rogers [27], with the additional factor of the leverage function appearing. In the non-leveraged case, i.e when $\ell(t, S_t) = 1$, the equations are identical.

5.2 A martingale approach

The above result can also be obtained via martingale pricing theory. By Girsanov's Theorem, see for example Theorem 11.3 in [3], the process defined by

$$d\bar{W}_t = h_t dt + dW_t \tag{5.38}$$

where h_t is any adapted process such that $\int_0^T |h(u)|du < \infty$ a.s, is a Brownian motion under the measure $\mathbb{Q}$ defined by

$$d\mathbb{Q} = L_t d\mathbb{P} \tag{5.39}$$

where

$$L_t = e^{\int_0^t h_s dW_s - \frac{1}{2}\int_0^t h_s^2 ds}. \tag{5.40}$$

In order to get the dynamics of (S_t, D_t) under the new measure $\mathbb{Q}$ we substitute (5.38) into (5.22) and (5.24) to get

$$
\begin{aligned}
dD_t &= \left(\mu - \frac{\phi(t)}{\Phi(t)}D_t + h_t\sigma(D_t)\ell(t,S_t)\right)dt + \sigma(D_t)\ell(t,S_t)d\bar{W}_t && (5.41)\\
dS_t &= \left(r + \mu + \frac{1}{2}\sigma^2(D_t)\ell^2(t,S_t) + h_t\sigma(D_t)\ell(t,S_t)\right)S_t dt \\
&\quad + S_t\sigma(D_t)\ell(t,S_t)d\bar{W}_t. && (5.42)
\end{aligned}
$$

We want the discounted stock price process to be a martingale under our new measure $\mathbb{Q}$, and so we set $r+\mu+\frac{1}{2}\sigma^2(D_t)\ell^2(t,S_t)+h_t\sigma(D_t)\ell(t,S_t) = r$. This gives that,

$$h_t = \frac{-\mu - \frac{1}{2}\sigma^2(D_t)\ell^2(t,S_t)}{\sigma(D_t)\ell(t,S_t)} \tag{5.43}$$

which defines our new measure $\mathbb{Q}$ and ensures that $\mathbb{Q}$ is the risk neutral measure. The risk neutral dynamics of the model are then

$$
\begin{aligned}
dD_t &= \left(-\frac{\phi(t)}{\Phi(t)}D_t - \frac{1}{2}\sigma^2(D_t)\ell^2(t,S_t)\right)dt + \sigma(D_t)\ell(t,S_t)d\bar{W}_t && (5.44)\\
dS_t &= rS_t dt + \sigma(D_t)\ell(t,S_t)d\bar{W}_t. && (5.45)
\end{aligned}
$$

Then, since the discounted stock price is a martingale under this measure, the Martingale Pricing Theorem ensures that the price of a contingent claim with the payoff $\Psi(S_T)$ at time t is the discounted expected value of the payoff under the risk-neutral measure given the information up to t. In other words we have that

$$V_t = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\big[\Psi(S_T)|\mathcal{F}_t\big]. \tag{5.46}$$

If we again assume that the price is of the form $V_t = V(t, S_t, D_t)$ then we can use the Feynman-Kac formula to show that the solution must satisfy the PDE (5.33).

5.3 Calculating the leverage function

Up to this point we have side stepped the issue of calculating the leverage function by making assumptions about its form. A method for calculating the leverage function, as done by Guyon in [20], will now be presented and we will use this to motivate the assumptions made up to this point.

The key property of the leverage function is that it ensures that the model (5.1)-(5.2), or equivalently (5.44)-(5.45), exactly fits the market smile. We now recall Gyöngy's Theorem, which describes the construction of a 1 dimensional Itô process, with non-random coefficients, with the same marginal distribution (i.e option prices) as some, more complex and possibly random, multidimensional model. In the case of a model for an underlying stock, we call the resulting volatility term that appears in the new one dimensional process the local volatility. If we consider the model (5.44)-(5.45) as the complex multidimensional stock prices process in the theorem, we can extract a simple process of the form

$$d\bar{S}_t = a(t, \bar{S}_t)dt + \sigma_{LV}(t, \bar{S}_t)d\bar{W}_t \tag{5.47}$$

with the same European call option prices. The coefficient $\sigma_{LV} : [0, T] \times \mathbb{R} \to \mathbb{R}$ has the interpretation

$$\sigma_{LV}(t, x) = \sqrt{\mathbb{E}[\sigma^2(D_t)\ell^2(t, S_t)|S_t = x]}, \tag{5.48}$$

now using the fact the $\ell^2(t, S_t)$ is an S_t measurable random variable we have, by the properties of conditional expectation, that

$$\sigma_{LV}(t, x) = \ell(t, x)\sqrt{\mathbb{E}[\sigma^2(D_t)|S_t = x]}. \tag{5.49}$$

We recall that since, by definition, $\mathbb{E}[\sigma^2(D_t)|S_t = x]$ is a $S_t = x$ measurable random variable, it can be written as a function of x by the Doob-Dynkin

Lemma. Now since $\sigma(x) > 0$ for all $x \in \mathbb{R}$ we know that $\sqrt{\mathbb{E}[\sigma^2(D_t)|S_t = x]} > 0$ for all $x\epsilon\mathbb{R}$, and we therefore can safely divide to get

$$\ell(t,x) = \frac{\sigma_{LV}(t,x)}{\sqrt{\mathbb{E}[\sigma^2(D_t)|S_t = x]}} \tag{5.50}$$

this gives us a method for calculating the leverage function. In the local volatility section we were able to construct $\sigma_{LV}(t,x)$ from the implied volatility surface. The conditional expectation can be approximated to any degree of accuracy using a Monte Carlo method. We can, therefore, use (5.50) to construct the leverage function.

We now return our attention to the assumptions made earlier in the subsection. The first of these assumption was that the leverage function is nowhere 0. Considering (5.50), this would be true if the numerator stays away from zero and the denominator is bounded. Since $\sigma(\cdot)$ was bounded by construction the later is true, for the former we recall that the local volatility surface must be positive by no arbitrage arguments, therefore our first assumption is valid.

The second assumption was that the leverage function was Lipschitz in the second argument, this is more difficult to motivate. The following facts will be important

1. If a function is continuously differentiable on a subset $X \subseteq \mathbb{R}$ then it is Lipschitz on X
2. If the function $g : \mathbb{R} \to \mathbb{R}$ is strictly positive on $\mathbb{R}$ then the number $\varepsilon = \inf_{x\in\mathbb{R}}\{g(x)\}$ exists and $\varepsilon > 0$.
3. If the function $g : \mathbb{R} \to \mathbb{R}$ is strictly positive and Lipschitz on $X \subseteq \mathbb{R}$, then the function defined by $h(x) = \sqrt{g(x)}$ is Lipschitz on X.

We now consider the following proposition,

Proposition 5.2 *Let $g : \mathbb{R} \to \mathbb{R}$ be a strictly positive, bounded and Lipschitz function, also let $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ be bounded and Lipschitz in the second argument, then the function $h : [0,T] \times \mathbb{R} \to \mathbb{R}$ defined by $h(t,x) = \frac{f(t,x)}{\sqrt{g(x)}}$ is Lipschitz in the second argument. i.e there exists a $C < \infty$ such that*

$$|h(t,x) - h(t,y)| \leq C|x-y|, \ \forall x,y \in \mathbb{R} \text{ and } t \in [0,T] \tag{5.51}$$

Proof

$$
\begin{aligned}
|h(t,x) - h(t,y)| &= \left| \frac{f(t,x)}{\sqrt{g(x)}} - \frac{f(t,y)}{\sqrt{g(y)}} \right| && (5.52)\\
&= \left| \frac{f(x)\sqrt{g(y)} - f(y)\sqrt{g(x)}}{\sqrt{g(x)g(y)}} \right| && (5.53)\\
&\leq \frac{1}{\varepsilon}\left| f(x)\sqrt{g(y)} - f(y)\sqrt{g(x)} \right|, && (5.54)\\
&\text{where } \varepsilon = \inf\{g(x)\}\\
&= \frac{1}{\varepsilon}\left| f(x)\sqrt{g(y)} - f(y)\sqrt{g(y)} \right.\\
&\left. + f(y)\sqrt{g(y)} - f(y)\sqrt{g(x)} \right| && (5.55)\\
&= \frac{1}{\varepsilon}\left| \sqrt{g(y)}[f(x) - f(y)] + f(y)\left[\sqrt{g(y)} - \sqrt{g(x)}\right] \right| && (5.56)\\
&\leq \frac{1}{\varepsilon}\left\{ \left| \sqrt{g(y)}[f(x) - f(y)] \right| \right.\\
&\left. + \left| f(y)\left[\sqrt{g(y)} - \sqrt{g(x)}\right] \right| \right\} && (5.57)\\
&\leq \frac{1}{\varepsilon}\left\{ \left| \sqrt{U_1}[x - y] \right| + \left| U_2[y - x] \right| \right\}, && (5.58)\\
&\text{for some } U_1, U_2 \in \mathbb{R}\\
&= \left(\frac{\sqrt{U}_1 + U_2}{\varepsilon} \right)|x - y| && (5.59)
\end{aligned}
$$

∎

This proposition tells us that if σ_{LV} is a bounded and Lipschitz function, and if $\mathbb{E}[\sigma^2(D_t)|S_t = x]$ is strictly positive, bounded and Lipschitz, then the leverage function will also be Lipschitz. In the local volatility section we constructed σ_{LV} using derivatives of the option price surface, it was therefore require that the surface was continuously differentiable, which in turn means that σ_{LV} is continuously differentiable and hence Lipschitz. It is also, therefore, bounded if considered on a bounded domain. That $\mathbb{E}[\sigma^2(D_t)|S_t = x]$ is strictly positive and bounded is obvious since it inherits these properties directly from $\sigma(\cdot)$, that it is Lipschitz is less so.

The question becomes whether the conditional expectation of a Lipschitz function is itself Lipschitz. Although we were not able to directly prove this, it seems as if it must hold. Intuitively, the conditional expectation operator smooths out the function σ via integration, which should not effect the differentiability of the function and hence its Lipschitz property. Obviously this argument is purely "hand waving", and this property requires further investigation, but for now we accept that the assumptions made about the leverage function are indeed reasonable.

At this point we reflect on what has been achieved. Under certain assumptions, we were able to derive a deterministic pricing PDE for a contingent claim with a general payoff. This PDE involved a leverage function, a deterministic function that ensures that the prices we eventually calculate using our PDE are consistent with the market prices of European options. A method for calculating this leverage function was presented, and the assumptions that allowed the derivation of the PDE were motivated. In the paper by Guyon [20] in which the leverage function methodology was originally proposed, a Monte Carlo style algorithm was suggested when pricing options. In the above framework, once the leverage function has been calculated, it is possible to generate an entire surface of option prices by solving the PDE. This has obvious advantages over a straight Monte Carlo, especially with application to hedging. However, our method requires that we know the leverage function over the whole S_t domain and so, we will have to run multiple Monte Carlo simulations to calculate the necessary conditional expectation. This will be an obvious drag on any proposed numerical scheme, but there may be a way to bypass this issue. Guyon suggests a Monte Carlo method which calculates the leverage function while simulating the paths of the underlying process. We will put this method to the test for a specific case and see what there is to learn about the form of the leverage function.

5.4 The Particle method

In our fitted model, the system (D_t, Z_t) has the following dynamics under the martingale measure

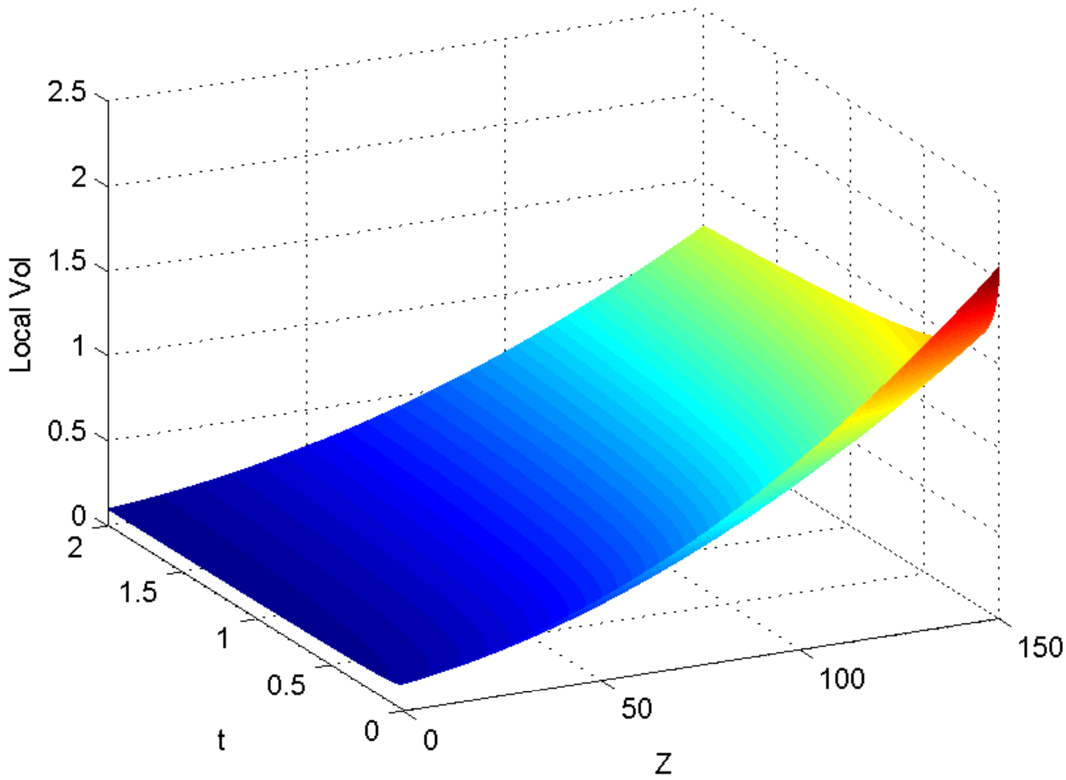
$$dD(t) = dZ(t) - \lambda D_t dt \tag{5.60}$$

$$dZ(t) = -\frac{1}{2}\sigma^2(D_t)\frac{\sigma_{LV}^2(t, Z_t)}{\mathbb{E}\big(\sigma^2(D_t)|Z_t\big)}dt + \sigma(D_t)\frac{\sigma_{LV}(t, Z_t)}{\sqrt{\mathbb{E}\big(\sigma^2(D_t)|Z_t\big)}}dW_t \tag{5.61}$$

in this section we consider the case where,

$$\sigma(x) = \sqrt{1 + \epsilon x^2} \wedge \eta.$$

and we use the following parametrically constructed local volatility surface.



We then consider the following N copies of the system of SDEs

$$dD^i(t) = dZ^i(t) - \lambda D^i_t dt$$

$$dZ^i_t = -\frac{1}{2}\sigma^2(D^i_t)\frac{\sigma^2_{LV}(t, Z^i_t)}{\mathbb{E}\big(\sigma^2(D^i_t)|Z^i_t\big)}dt + \sigma(D^i_t)\frac{\sigma_{LV}(t, Z^i_t)}{\sqrt{\mathbb{E}\big(\sigma^2(D^i_t)|Z^i_t\big)}}dW_t$$

with $i = 1, 2, \ldots, N$. We then use the following discretisation of our N copies:

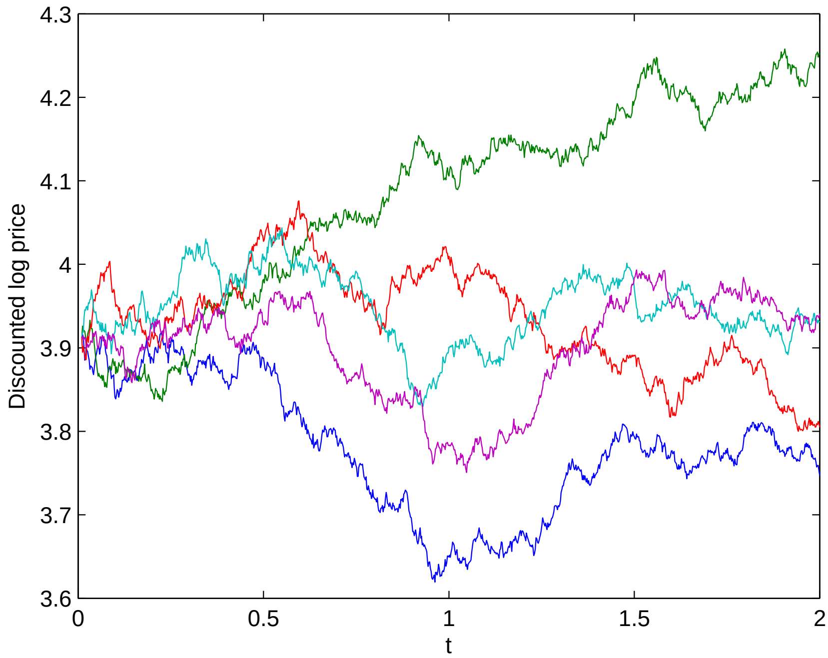
$$D^i_{t+1} = D^i_t + Z^i_{t+1} - Z^i_t - \lambda(D^i_t)(\Delta_t) \tag{5.62}$$

$$\begin{aligned} Z^i_{t+1} = Z^i_t - \frac{1}{2}\sigma^2(D^i_t)\frac{\sigma^2_{LV}(t, Z^i_t)}{\left(\sum_{j\in A}\frac{1}{n(A_j)}\sigma^2(D^j_t)\right)}\Delta_t \\ + \sigma(D^i_t)\frac{\sigma_{LV}(t, Z^i_t)}{\left(\sum_{j\in A}\frac{1}{n(A_j)}\sigma^2(D^j_t)\right)^{\frac{1}{2}}}z\sqrt{\Delta_t} \end{aligned} \tag{5.63}$$

where $A_i = \{j : Z_i - \varepsilon < Z_j < Z_i + \varepsilon\}$, and we have approximated the martingale measure $\mathbb{Q}$ by the empirical measure. z is a standard normal random variable. Each Z^i is know as a particle, and each particle interacts with other particles through the approximation of the conditional expectation.

To prove that the system (5.62)-(5.63) converges to (5.60)-(5.61) requires sophisticated arguments involving propagation chaos, a detailed discussion of this can be found in [21]. We will assume that it does converge as $N \to \infty$.

The scheme was run for $Z_0 = \ln(50)$, $D_0 = 0$, $\eta = 0.2$, $\epsilon = 5$, $T = 2$, $\Delta_t = 0.002$ (1000 time steps), $N = 5000$ and $\varepsilon = 0.005$. Below is a plot of the paths for 5 of our Z^i's.



We then considered some key statistical quantities. We remember that since e^Z is a martingale we would expect $\mathbb{E}(e^{Z_t}) = \mathbb{E}(e^{Z_0})$, for our scheme it was found that $\frac{1}{N}\sum_{i=1}^{N} e^{Z_T^i} = 50.0955 \approx 50$ supporting the consistency of the discretisation

Understanding the behaviour of the random variable $\mathbb{E}\big(\sigma^2(D_t)|Z_t\big)$ is important if we are going to estimate the leverage function. It was found that this random variable does not deviate significantly from a long run mean, and this suggests that modelling it as a constant is not unreasonable. A rigorous proof of this is not presented, rather we will provide statistical evidence.

We recall that we essentially approximated the above conditional expectation using the empirical measure and by grouping paths in buckets of width 2ε. For each time step t and particle j we can define the quantity

$$\bar{\sigma}(t,j)^2 = \left(\sum_{j \in A} \frac{1}{n(A_j)} \sigma^2(D_t^j) \right)^2$$

of course this quantity was used and recorded in the scheme itself. For each time set we can then calculate the mean and variance of $\bar{\sigma}$ over the j's. This is plotted below:

Figure 1: The empirical mean as a function of time

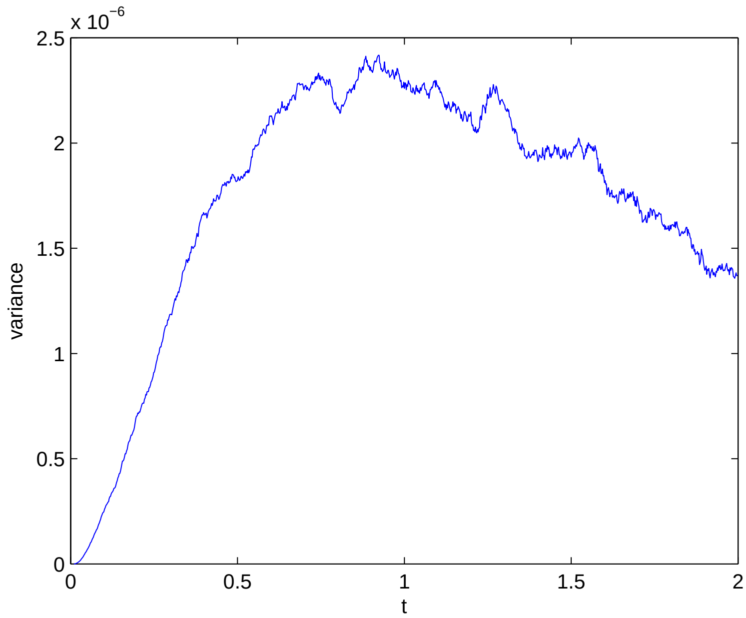


Figure 2: The variance as a function of time

Clearly both the variance and mean increase with time, but as in terms of absolutes the mean does not deviate much from its original value and the variance remains low throughout, maximising at a value of 2.4×10^{-6}.This supports the hypothesis that the conditional expectation for this particular case can be treated as a constant.

This treatment can then be applied to the calculation of the leverage function

$$\ell(t, S_t) = \frac{\sigma_{LV}(t, S_t)}{\sqrt{\mathbb{E}(\sigma^2(D_t)|Z_t)}} \tag{5.64}$$

$$\approx \frac{\sigma_{LV}(t, S_t)}{\sigma(D_0)} \tag{5.65}$$

Using this estimate we can avoid running a Monte Carlo to calculate the leverage function for use in our partial differential equation. Of course this type of evidence can only be considered relevant to this particular case, and it will most likely need to be verified for any other special case before it is used in the implementation of a PDE scheme.

However it gives further evidence that leverage functions seem to flatten out over time as hypothesised by Guyon [20], and gives insight into a possible method of approximation.

6 A numerical implementation of the HR model

In the preceding sections we have continuously mentioned the advantages of having a Markovian model that admits a terminal value problem for the price of a T-claim. In this section we revisit a scheme, constructed in fulfilment of MSc course work requirements, which was designed for the non-leveraged HR model. We then attempt to adapt this numerical scheme to the leveraged case, and make observations about the results.

6.1 The non-leveraged model revisited

At this point we recall the theory set out in subsection 4.3. If we set $n = 1$ in equation (4.30) we have shown that our model is then governed by the following system of SDEs,

$$dB_t = rB_t dt \tag{6.1}$$

$$dD_t = (\mu(D_t) - \lambda D_t)dt + \sigma(D_t)dW_t \tag{6.2}$$

$$dZ_t = \mu(D_t)dt + \sigma(D_t)dW_t \tag{6.3}$$

here we have written $D_t^{(1)} := D_t$ since no ambiguity is possible. The process (Z_t, D_t) is Markov and so we can assume that the price of a contingent claim u has the form $u = u(T - t, S_t, D_t)$. The system yields the following PDE for the price of a contingent claim u, paying $\psi(S_T)$ at time T for some measurable $\psi(.)$,

$$\begin{aligned} 0 &= \left(rS\frac{\partial u}{\partial S} - ru - \lambda D\frac{\partial u}{\partial D} - \frac{\partial u}{\partial t} \right) \\ &+ \left(-\frac{1}{2}\frac{\partial u}{\partial D} + \frac{1}{2}S^2\frac{\partial^2 u}{\partial S^2} + \frac{1}{2}\frac{\partial^2 u}{\partial D^2} + S\frac{\partial^2 u}{\partial S \partial D} \right)\sigma^2(D). \\ &:= \mathbb{L}u. \end{aligned} \tag{6.4}$$

We will attempt to price a European put option, with strike K and maturity T, in the framework of the above model, i.e $\Phi(S_T) = \max(0, K - S_T)$.

At this point it is common practice to employ a change of variables in equation (6.4) as in [12], however, in this dissertation we will endeavour to construct a scheme without a change of variables in order to investigate the efficiency of such a scheme. When pricing an option the boundary conditions are of crucial importance, by not changing variables we have the advantage of a real world framework and so financially intuitive boundary conditions can be easily constructed.

6.1.1 European put boundary and initial conditions

As was discussed in the previous subsection, the boundary conditions used are of vital importance in the construction of our scheme and the eventual solution. For the final condition we simply have that the option is equal to its payoff. We also remember at this point that our solution u is of the form $u = u(T-t, S_t, D_t)$ and so the condition at time $t = T$ is interpreted as the following initial condition:

$$u(0, S_T, D_T) = max(0, K - S_T). \tag{6.5}$$

Next we consider the situation where S is very large. In the case of a put option this means that the probability of exercise is very low, and therefore the value of the put option is very low. As a consequence, at the boundary where $S \to \infty$, we will require that

$$u(t, S, D) = 0. \tag{6.6}$$

At the boundary where $S = 0$ we assume that the exercise of the option is guaranteed. For a put option this means the receipt of K in cash upon expiry, and so we have

$$u(t, 0, D) = Ke^{-r(T-t)}. \tag{6.7}$$

The situation is slightly more complex at boundaries where $D \to \pm\infty$. Here we remember that our volatility function $\sigma(\cdot)$ is assumed Lipschitz and bounded. It is also common practice to use a function that has the property $\sigma(x) = \sigma_{max}$ if x is such that $|x| \geq B$ for some predetermined B and σ_{max}. The function used in this dissertation will be $\sigma(x) = \min(\eta\sqrt{1+\epsilon x}, \sigma_{max})$, which clearly has the aforementioned property. In this context it is easy to see that for large absolute value deviations from the mean, the volatility behaves like a constant, and so a natural condition for these extreme points is that the option price agrees with the Black-Scholes framework detailed in [4]. A boundary condition of this form is used in [12], the authors insist that for large D values the price of the option must equal the Black-Scholes price. We modify this slightly by insisting that, at the boundaries where D has large absolute value, the option price must obey the Black-Scholes PDE. This gives

$$0 = rSu_s - ru - u_t + \frac{1}{2}\sigma_{max}^2 S^2 u_{ss}. \tag{6.8}$$

The advantage of using the Black-Scholes PDE instead of the Black-Scholes formula is in the implementation. It is possible to incorporate this PDE into the system of equations we will eventually solve at each time step.

6.2 Finite difference approximations

We consider the problem in the following discrete space: we divide the state space S into points $(S_1, S_2, \ldots, S_M)$, where $S_1 = 0$ and S_M is some sufficiently large number. Similarly we use the points $(D_1, D_2, \ldots, D_L)$, and $(t_1, t_2, \ldots, t_N)$ as representations of the D and t space respectively. For each time we then have a 2 dimensional grid, with S on the y-axis and D on the x-axis, with $M \times L$ points. We then define the grid function $f = f(t_n, S_i, D_j) := f^n_{i,j}$ that takes a value on each of the grid nodes. f is the function that we will use to approximate u. We will use the following second order finite difference approximations for the derivatives in our numerical scheme,

$$\frac{\partial u}{\partial t} \approx \frac{f^{n+1}_{i,j} - f^n_{i,j}}{k} \tag{6.9}$$

$$\frac{\partial u}{\partial S} \approx \frac{1}{2}\left[\frac{f^n_{i+1,j} - f^n_{i-1,j}}{2h} + \frac{f^{n+1}_{i+1,j} - f^{n+1}_{i-1,j}}{2h}\right] \tag{6.10}$$

$$\frac{\partial u}{\partial D} \approx \frac{1}{2}\left[\frac{f^n_{i,j+1} - f^n_{i,j-1}}{2d} + \frac{f^{n+1}_{i,j+1} - f^{n+1}_{i,j-1}}{2d}\right] \tag{6.11}$$

$$\frac{\partial^2 u}{\partial S^2} =\approx \frac{1}{2}\left[\frac{f^n_{i+1,j} - 2f^n_{i,j} + f^n_{i-1,j}}{h^2} + \frac{f^{n+1}_{i+1,j} - 2f^{n+1}_{i,j} + f^{n+1}_{i-1,j}}{h^2}\right] \tag{6.12}$$

$$\frac{\partial^2 u}{\partial D^2} \approx \frac{1}{2}\left[\frac{f^n_{i,j+1} - 2f^n_{i,j} + f^n_{i,j-1}}{d^2} + \frac{f^{n+1}_{i,j+1} - 2f^{n+1}_{i,j} + f^{n+1}_{i,j-1}}{d^2}\right] \tag{6.13}$$

$$\begin{aligned}\frac{\partial^2 u}{\partial S \partial D} \approx \frac{1}{2}\Bigg[&\frac{f^n_{i+1,j+1} - f^n_{i+1,j-1} - f^n_{i-1,j+1} + f^n_{i-1,j-1}}{4hd} \\ &+ \frac{f^{n+1}_{i+1,j+1} - f^{n+1}_{i+1,j-1} - f^{n+1}_{i-1,j+1} + f^{n+1}_{i-1,j-1}}{4hd}\Bigg].\end{aligned} \tag{6.14}$$

6.3 Discretised PDE

If we substitute the approximations from equations (6.9)-(6.14) into our PDE (6.38), we get the following discrete equation,

$$
\begin{aligned}
0 = &\left(\frac{rS_i}{2}\left[\frac{f^n_{i+1,j}-f^n_{i-1,j}}{2h}+\frac{f^{n+1}_{i+1,j}-f^{n+1}_{i-1,j}}{2h}\right] - rf^n_{i,j} - \frac{\lambda D_j}{2}\left[\frac{f^n_{i,j+1}-f^n_{i,j-1}}{2d}+\frac{f^{n+1}_{i,j+1}-f^{n+1}_{i,j-1}}{2d}\right] - \right. \\
&\left. -\left[\frac{f^{n+1}_{i,j}-f^n_{i,j}}{k}\right]\right) + \left\{ -\frac{1}{4}\left[\frac{f^n_{i,j+1}-f^n_{i,j-1}}{2d}+\frac{f^{n+1}_{i,j+1}-f^{n+1}_{i,j-1}}{2d}\right]+ \right. \\
&+\frac{S_i^2}{4}\left[\frac{f^n_{i+1,j}-2f^n_{i,j}+f^n_{i-1,j}}{h^2}+\frac{f^{n+1}_{i+1,j}-2f^{n+1}_{i,j}+f^{n+1}_{i-1,j}}{h^2}\right]+ \\
&+\frac{1}{4}\left[\frac{f^n_{i,j+1}-2f^n_{i,j}+f^n_{i,j-1}}{d^2}+\frac{f^{n+1}_{i,j+1}-2f^{n+1}_{i,j}+f^{n+1}_{i,j-1}}{d^2}\right]+ \\
&\left. +S_i\frac{1}{2}\left[\frac{f^n_{i+1,j+1}-f^n_{i+1,j-1}-f^n_{i-1,j+1}+f^n_{i-1,j-1}}{4hd}+\frac{f^{n+1}_{i+1,j+1}-f^{n+1}_{i+1,j-1}-f^{n+1}_{i-1,j+1}+f^{n+1}_{i-1,j-1}}{4hd}\right]\right\}\sigma_j^2 \\
&:= \mathbb{L}_{[k,h,d]}f.
\end{aligned} \tag{6.15}
$$

Notice that we have also defined the linear operator $\mathbb{L}_{[k,h,d]}$ here. The above equation can be written in a more convenient way,

$$
\begin{aligned}
\Rightarrow 0 = &\frac{rS_i}{4h}[f^n_{i+1,j}-f^n_{i-1,j}+f^{n+1}_{i+1,j}-f^{n+1}_{i-1,j}] - rf^n_{i,j} - \frac{\lambda D_j}{4d}[f^n_{i,j+1}-f^n_{i,j-1}+f^{n+1}_{i,j+1}-f^{n+1}_{i,j-1}] \\
&-\frac{1}{k}[f^{n+1}_{i,j}-f^n_{i,j}] - \frac{\sigma_j^2}{8d}[f^n_{i,j+1}-f^n_{i,j-1}+f^{n+1}_{i,j+1}-f^{n+1}_{i,j-1}] \\
&+\frac{S_i^2\sigma_j^2}{4h^2}[f^n_{i+1,j}-2f^n_{i,j}+f^n_{i-1,j}+f^{n+1}_{i+1,j}-2f^{n+1}_{i,j}+f^{n+1}_{i-1,j}]+ \\
&+\frac{\sigma_j^2}{4d^2}[f^n_{i,j+1}-2f^n_{i,j}+f^n_{i,j-1}+f^{n+1}_{i,j+1}-2f^{n+1}_{i,j}+f^{n+1}_{i,j-1}]+ \\
&+\frac{S_i\sigma_j^2}{8hd}[f^n_{i+1,j+1}-f^n_{i+1,j-1}-f^n_{i-1,j+1}+f^n_{i-1,j-1}+f^{n+1}_{i+1,j+1}-f^{n+1}_{i+1,j-1}-f^{n+1}_{i-1,j+1}+f^{n+1}_{i-1,j-1}]
\end{aligned} \tag{6.16}
$$

$$\begin{aligned}
&\Rightarrow \Big[-\frac{S_i\sigma_j^2}{8hd}\Big] f^{n+1}_{i+1,j+1} + \Big[-\frac{rS_i}{4h} - \frac{S_i^2\sigma_j^2}{4h^2}\Big] f^{n+1}_{i+1,j} + \Big[\frac{\lambda D_j}{4d} + \frac{\sigma_j^2}{8d} - \frac{\sigma_j^2}{4d^2}\Big] f^{n+1}_{i,j+1} + \\
&+ \Big[\frac{1}{k} + \frac{S_i^2\sigma_j^2}{2h^2} + \frac{\sigma_j^2}{2d^2}\Big] f^{n+1}_{i,j} + \Big[\frac{rS_i}{4h} - \frac{S_i^2\sigma_j^2}{4h^2}\Big] f^{n+1}_{i-1,j} + \Big[-\frac{\lambda D_j}{4d} - \frac{\sigma_j^2}{8d} - \frac{\sigma_j^2}{4d^2}\Big] f^{n+1}_{i,j-1} + \\
&+ \Big[\frac{S_i\sigma_j^2}{8hd}\Big] f^{n+1}_{i+1,j-1} + \Big[\frac{S_i\sigma_j^2}{8hd}\Big] f^{n+1}_{i-1,j+1} + \Big[-\frac{S_i\sigma_j^2}{8hd}\Big] f^{n+1}_{i-1,j-1} \qquad (6.17)\\
&= \Big[\frac{S_i\sigma_j^2}{8hd}\Big] f^{n}_{i+1,j+1} + \Big[\frac{rS_i}{4h} + \frac{S_i^2\sigma_j^2}{4h^2}\Big] f^{n}_{i+1,j} + \Big[\frac{-\lambda D_j}{4d} - \frac{\sigma_j^2}{8d} + \frac{\sigma_j^2}{4d^2}\Big] f^{n}_{i,j+1} + \\
&+ \Big[\frac{1}{k} - \frac{S_i^2\sigma_j^2}{2h^2} - \frac{\sigma_j^2}{2d^2} - r\Big] f^{n}_{i,j} + \Big[-\frac{rS_i}{4h} + \frac{S_i^2\sigma_j^2}{4h^2}\Big] f^{n}_{i-1,j} + \Big[\frac{\lambda D_j}{4d} + \frac{\sigma_j^2}{8d} + \frac{\sigma_j^2}{4d^2}\Big] f^{n}_{i,j-1} + \\
&+ \Big[-\frac{S_i\sigma_j^2}{8hd}\Big] f^{n}_{i+1,j-1} + \Big[-\frac{S_i\sigma_j^2}{8hd}\Big] f^{n}_{i-1,j+1} + \Big[\frac{S_i\sigma_j^2}{8hd}\Big] f^{n}_{i-1,j-1}
\end{aligned}$$

$$\begin{aligned}
&\Rightarrow [\mathcal{A}_{i,j}] f^{n+1}_{i+1,j+1} + [\mathcal{B}_{i,j}] f^{n+1}_{i+1,j} + [\mathcal{C}_{i,j}] f^{n+1}_{i,j+1} + [\mathcal{D}_{i,j}] f^{n+1}_{i,j} + [\mathcal{B}_{i,j} + \frac{rS_i}{2h}] f^{n+1}_{i-1,j} + [-\mathcal{C}_{i,j} - \frac{\sigma_j^2}{2d^2}] f^{n+1}_{i,j-1} + \\
&+ [-\mathcal{A}_{i,j}] f^{n+1}_{i+1,j-1} + [-\mathcal{A}_{i,j}] f^{n+1}_{i-1,j+1} + [\mathcal{A}_{i,j}] f^{n+1}_{i-1,j-1} = \\
&= -[\mathcal{A}_{i,j}] f^{n}_{i+1,j+1} - [\mathcal{B}_{i,j}] f^{n}_{i+1,j} - [\mathcal{C}_{i,j}] f^{n}_{i,j+1} + \qquad (6.18)\\
&+ \big([-\mathcal{D}_{i,j}] + \frac{2}{k} - r\big) f^{n}_{i,j} + [-\mathcal{B}_{i,j} - \frac{rS_i}{2h}] f^{n}_{i-1,j} + [\mathcal{C}_{i,j} + \frac{\sigma_j^2}{2d^2}] f^{n}_{i,j-1} - \\
&- [-\mathcal{A}_{i,j}] f^{n}_{i+1,j-1} - [-\mathcal{A}_{i,j}] f^{n}_{i-1,j+1} - [\mathcal{A}_{i,j}] f^{n}_{i-1,j-1}
\end{aligned}$$

where,

$$\mathcal{A}_{i,j} = -\frac{S_i\sigma_j^2}{8hd} \qquad (6.19)$$

$$\mathcal{B}_{i,j} = -\frac{rS_i}{4h} - \frac{S_i^2\sigma_j^2}{4h^2} \qquad (6.20)$$

$$\mathcal{C}_{i,j} = \frac{\lambda D_j}{4d} + \frac{\sigma_j^2}{8d} - \frac{\sigma_j^2}{4d^2} \qquad (6.21)$$

$$\mathcal{D}_{i,j} = \frac{1}{k} + \frac{S_i^2\sigma_j^2}{2h^2} + \frac{\sigma_j^2}{2d^2}. \qquad (6.22)$$

Equation (6.18) holds for $i = 2, ..., M-1$ and $j = 2, ..., L-1$ and so gives us a set of $(M-2)(L-2)$ equations. This equation describes the behaviour of all of the interior points on our grid. We must now turn our attention to the boundary points in order to get a complete set of equations.

6.4 Discretised boundary conditions

In order to complete the scheme and ensure we have a system with a sufficient number of equations to generate a unique solution we have to discretise the boundary conditions described in equations (6.5)-(6.8). The discrete version of equation (6.5) is

$$f_{i,j}^{1} = \max(0, K - S_i), \forall i = 1, \dots, L \text{ and } j = 1, \dots, M. \tag{6.23}$$

Equations (6.6) and (6.7) give,

$$f_{M,j}^{n+1} = f_{M,j}^{n} = \dots = f_{M,j}^{1} = 0, \forall j = 1, 2, \dots, L \tag{6.24}$$
$$f_{1,j}^{n+1} = e^{-rk} f_{1,j}^{n}, f_{1,j}^{1} = K, \forall n = 2, \dots, N \text{ and } j = 1, \dots, L. \tag{6.25}$$

(6.24) and (6.25) describe $2L$ equations. Finally we approximate equation (6.8) using a Crank-Nicolson scheme similar to the one used to discretisation our main PDE. This gives the following two sets of equations,

$$\begin{aligned} 0 = & \frac{rS_i}{4h}[f_{i+1,L}^{n+1} - f_{i-1,L}^{n} + f_{i+1,L}^{n} - f_{i-1,L}^{n}] - rf_{i,L}^{n} - \frac{f_{i,L}^{n+1} - f_{i,L}^{n}}{k} + \\ & + \frac{\sigma_{max}^2 S_i^2}{4h^2}[f_{i+1,L}^{n+1} - 2f_{i,L}^{n+1} + f_{i-1,L}^{n+1} + f_{i+1,L}^{n} - 2f_{i,L}^{n} + f_{i-1,L}^{n}] \\ 0 = & \frac{rS_i}{4h}[f_{i+1,1}^{n+1} - f_{i-1,1}^{n} + f_{i+1,1}^{n} - f_{i-1,1}^{n}] - rf_{i,1}^{n} - \frac{f_{i,1}^{n+1} - f_{i,1}^{n}}{k} + \\ & + \frac{\sigma_{max}^2 S_i^2}{4h^2}[f_{i+1,1}^{n+1} - 2f_{i,1}^{n+1} + f_{i-1,1}^{n+1} + f_{i+1,1}^{n} - 2f_{i,1}^{n} + f_{i-1,1}^{n}]. \end{aligned} \tag{6.26}$$

This true for all n, i. From (6.26) we have that

$$\begin{aligned} & \left[-\frac{rS_i}{4h} - \frac{\sigma_{max}^2 S_i^2}{4h^2} \right] f_{i+1,L}^{n+1} + \left[\frac{1}{k} + \frac{\sigma_{max}^2 S_i^2}{2h^2} \right] f_{i,L}^{n+1} + \left[\frac{rS_i}{4h} - \frac{\sigma_{max}^2 S_i^2}{2h^2} \right] f_{i-1,L}^{n+1} = \\ & = \left[\frac{rS_i}{4h} + \frac{\sigma_{max}^2 S_i^2}{4h^2} \right] f_{i+1,L}^{n} + \left[-r + \frac{1}{k} - \frac{\sigma_{max}^2 S_i^2}{2h^2} \right] f_{i,L}^{n} + \left[-\frac{rS_i}{4h} + \frac{\sigma_{max}^2 S_i^2}{2h^2} \right] f_{i-1,L}^{n} \end{aligned} \tag{6.27}$$

$$\begin{aligned} & \left[-\frac{rS_i}{4h} - \frac{\sigma_{max}^2 S_i^2}{4h^2} \right] f_{i+1,1}^{n+1} + \left[\frac{1}{k} + \frac{\sigma_{max}^2 S_i^2}{2h^2} \right] f_{i,1}^{n+1} + \left[\frac{rS_i}{4h} - \frac{\sigma_{max}^2 S_i^2}{2h^2} \right] f_{i-1,1}^{n+1} = \\ & = \left[\frac{rS_i}{4h} + \frac{\sigma_{max}^2 S_i^2}{4h^2} \right] f_{i+1,1}^{n} + \left[-r + \frac{1}{k} - \frac{\sigma_{max}^2 S_i^2}{2h^2} \right] f_{i,1}^{n} + \left[-\frac{rS_i}{4h} + \frac{\sigma_{max}^2 S_i^2}{2h^2} \right] f_{i-1,1}^{n}. \end{aligned}$$

Defining the coefficient vectors

$$\mathcal{K}_i = -\frac{rS_i}{4h} - \frac{\sigma_{max}^2 S_i^2}{4h^2} \tag{6.28}$$
$$\mathcal{L}_i = \frac{1}{k} + \frac{\sigma_{max}^2 S_i^2}{2h^2} \tag{6.29}$$
$$\mathcal{M}_i = \frac{rS_i}{4h} - \frac{\sigma_{max}^2 S_i^2}{2h^2} \tag{6.30}$$

we can rewrite (6.27) as

$$
\begin{aligned}
&[\mathcal{K}_i]f^{n+1}_{i+1,L} + [\mathcal{L}_i]f^{n+1}_{i,L} + [\mathcal{M}_i]f^{n+1}_{i-1,L} = \\
&= -[\mathcal{K}_i]f^{n}_{i+1,1} + [-\mathcal{L}_i + \frac{2}{k} - r]f^{n}_{i,1} + [\mathcal{M}_i]f^{n}_{i-1,1} \qquad (6.31) \\
&[\mathcal{K}_i]f^{n+1}_{i+1,1} + [\mathcal{L}_i]f^{n+1}_{i,1} + [\mathcal{M}_i]f^{n+1}_{i-1,1} = \\
&= -[\mathcal{K}_i]f^{n}_{i+1,1} + [-\mathcal{L}_i + \frac{2}{k} - r]f^{n}_{i,1} + [\mathcal{M}_i]f^{n}_{i-1,1}. \qquad (6.32)
\end{aligned}
$$

This being true $i = 2, \ldots, M-1$, so that (6.31) and (6.32) give us a set of $2M-4$ equations. Now, in total, we have $(M-2)(L-2)+2L+2M-4 = ML$ equations, and ML unknowns, and so we have a system of linear equations with a unique solution. The next step will be to construct an algorithm that solves these equations at every time step.

6.5 Matrix construction

We are now ready to represent our set of equations in matrix form. The goal will be to represent our scheme in the form

$$\underline{\mathfrak{L}}\vec{f}^{n+1} = \underline{\mathfrak{R}}\vec{f}^{n} \tag{6.33}$$

where the vector $\vec{f}^n$ is defined by $\vec{f}^n := (f^n_{1,1}, f^n_{1,2}, ..., f^n_{1,L}, f^n_{2,1}, ..., f^n_{2,L}, ..., f^n_{M,L-1}, f^n_{M,L})^T$. The matrices $\underline{\mathfrak{L}}$ and $\underline{\mathfrak{R}}$ are both of size $(ML \times ML)$, and will be constructed using equations (6.18), (6.23),(6.24),(6.25),(6.31) and (6.32). The most informative way to describe the construction of the matrices is to consider a specific calibration of the algorithm. Suppose we have that $M = 4$, and $L = 4$. In this case the following matrix equation will be solved at each time step,

$$
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathcal{M}_{2,1} & 0 & 0 & 0 & \mathcal{L}_{2,1} & 0 & 0 & 0 & \mathcal{K}_{2,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathcal{A}_{2,2} & (\mathcal{B}_{2,2} + \frac{rS_2}{2h}) & -\mathcal{A}_{2,2} & 0 & (-\mathcal{C}_{2,2} - \frac{\sigma_2^2}{2d^2}) & \mathcal{D}_{2,2} & \mathcal{C}_{2,2} & 0 & -\mathcal{A}_{2,2} & \mathcal{B}_{2,2} & \mathcal{A}_{2,2} & 0 & 0 & 0 & 0 & 0 \\
0 & \mathcal{A}_{2,3} & (\mathcal{B}_{2,3} + \frac{rS_2}{2h}) & -\mathcal{A}_{2,3} & 0 & (-\mathcal{C}_{2,3} - \frac{\sigma_3^2}{2d^2}) & \mathcal{D}_{2,3} & \mathcal{C}_{2,3} & 0 & -\mathcal{A}_{2,3} & \mathcal{B}_{2,3} & \mathcal{A}_{2,3} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \mathcal{M}_{2,4} & 0 & 0 & 0 & \mathcal{L}_{2,4} & 0 & 0 & 0 & \mathcal{K}_{2,4} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \mathcal{M}_{3,1} & 0 & 0 & 0 & \mathcal{L}_{3,1} & 0 & 0 & 0 & \mathcal{K}_{3,1} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \mathcal{A}_{3,2} & (\mathcal{B}_{3,2} + \frac{rS_3}{2h}) & -\mathcal{A}_{3,2} & 0 & (-\mathcal{C}_{3,2} - \frac{\sigma_2^2}{2d^2}) & \mathcal{D}_{3,2} & \mathcal{C}_{3,2} & 0 & -\mathcal{A}_{3,2} & \mathcal{B}_{3,2} & \mathcal{A}_{3,2} & 0 \\
0 & 0 & 0 & 0 & 0 & \mathcal{A}_{3,3} & (\mathcal{B}_{3,3} + \frac{rS_3}{2h}) & -\mathcal{A}_{3,3} & 0 & (-\mathcal{C}_{3,3} - \frac{\sigma_3^2}{2d^2}) & \mathcal{D}_{3,3} & \mathcal{C}_{3,3} & 0 & -\mathcal{A}_{3,3} & \mathcal{B}_{3,3} & \mathcal{A}_{3,3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathcal{M}_{3,4} & 0 & 0 & 0 & \mathcal{L}_{3,4} & 0 & 0 & 0 & \mathcal{K}_{3,4} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\vec{f}^{n+1} =
$$

$$
\begin{pmatrix}
e^{-rk} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & e^{-rk} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & e^{-rk} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & e^{-rk} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\mathcal{M}_{2,1} & 0 & 0 & 0 & -\mathcal{L}_{2,1} + \frac{2}{k} - r & 0 & 0 & 0 & -\mathcal{K}_{2,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\mathcal{A}_{2,2} & -(\mathcal{B}_{2,2} + \frac{rS_2}{2h}) & \mathcal{A}_{2,2} & 0 & -(-\mathcal{C}_{2,2} - \frac{\sigma_2^2}{2d^2}) & -\mathcal{D}_{2,2} + \frac{2}{k} - r & -\mathcal{C}_{2,2} & 0 & \mathcal{A}_{2,2} & -\mathcal{B}_{2,2} & -\mathcal{A}_{2,2} & 0 & 0 & 0 & 0 & 0 \\
0 & -\mathcal{A}_{2,3} & -(\mathcal{B}_{2,3} + \frac{rS_2}{2h}) & \mathcal{A}_{2,3} & 0 & -(-\mathcal{C}_{2,3} - \frac{\sigma_3^2}{2d^2}) & -\mathcal{D}_{2,3} + \frac{2}{k} - r & -\mathcal{C}_{2,3} & 0 & \mathcal{A}_{2,3} & -\mathcal{B}_{2,3} & -\mathcal{A}_{2,3} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\mathcal{M}_{2,4} & 0 & 0 & 0 & -\mathcal{L}_{2,4} + \frac{2}{k} - r & 0 & 0 & 0 & -\mathcal{K}_{2,4} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\mathcal{M}_{3,1} & 0 & 0 & 0 & -\mathcal{L}_{3,1} + \frac{2}{k} - r & 0 & 0 & 0 & -\mathcal{K}_{3,1} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\mathcal{A}_{3,2} & -(\mathcal{B}_{3,2} + \frac{rS_3}{2h}) & \mathcal{A}_{3,2} & 0 & -(-\mathcal{C}_{3,2} - \frac{\sigma_2^2}{2d^2}) & -\mathcal{D}_{3,2} + \frac{2}{k} - r & -\mathcal{C}_{3,2} & 0 & \mathcal{A}_{3,2} & -\mathcal{B}_{3,2} & -\mathcal{A}_{3,2} & 0 \\
0 & 0 & 0 & 0 & 0 & -\mathcal{A}_{3,3} & -(\mathcal{B}_{3,3} + \frac{rS_3}{2h}) & \mathcal{A}_{3,3} & 0 & -(-\mathcal{C}_{3,3} - \frac{\sigma_3^2}{2d^2}) & -\mathcal{D}_{3,3} + \frac{2}{k} - r & -\mathcal{C}_{3,3} & 0 & \mathcal{A}_{3,3} & -\mathcal{B}_{3,3} & -\mathcal{A}_{3,3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\mathcal{M}_{3,4} & 0 & 0 & 0 & -\mathcal{L}_{3,4} + \frac{2}{k} - r & 0 & 0 & 0 & -\mathcal{K}_{3,4} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\vec{f}^{n}
$$

.

In the algorithm for the construction of the left matrix, we break construction up into five parts. We construct four separate matrices corresponding to the four boundary conditions, and one matrix for the interior points. The left matrix is then obtain by adding these 5 matrices together. This procedure is modified and repeated for the right matrix. The construction algorithm is presented in detail in the form of a Matlab code, this is attached as appendix 2.

6.6 Numerical results

Now that we have designed the scheme and investigated its stability we can proceed with some numerical tests. We use the volatility function

$$\sigma(D) = \min(\eta\sqrt{1 + \epsilon D^2}, \sigma_{max}) \tag{6.34}$$

in our numerical experiments, as in [27]. For our first experiment we use the parameter values used in [27].

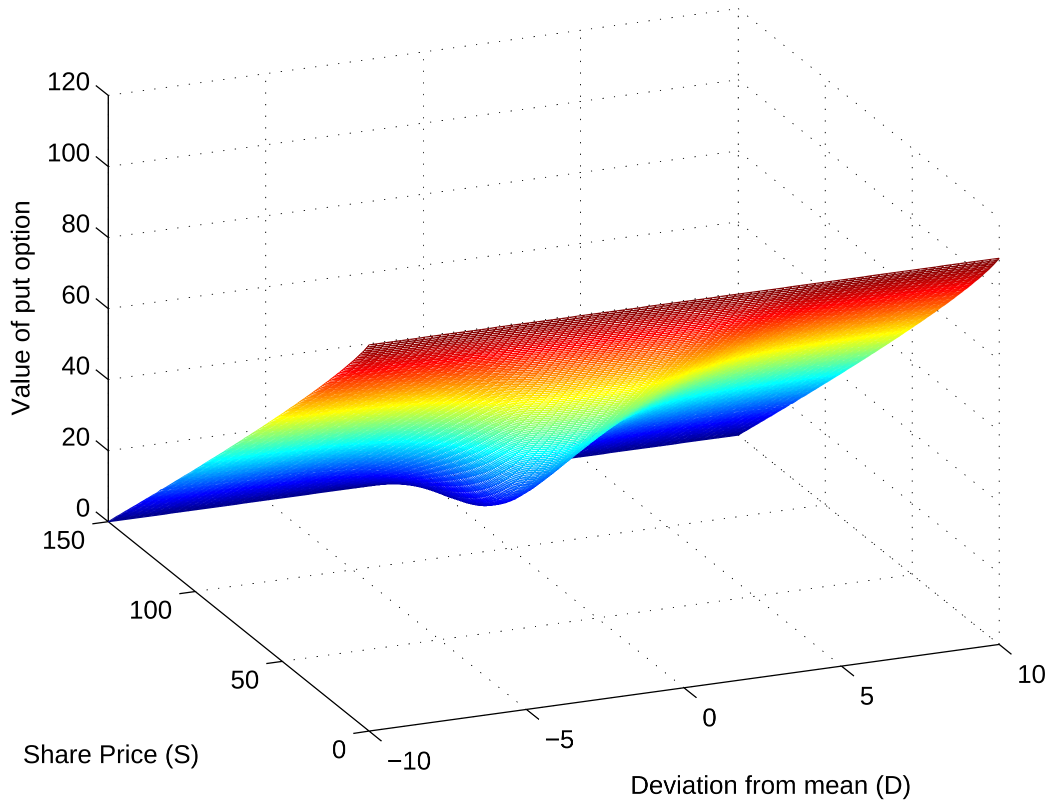


Figure 3: This image is a outcome of our scheme with parameters chosen to correspond with those in the original paper by Hobson and Rogers [27]. We use: $\epsilon = 5$, $\eta = 0.2$, $\sigma_{max} = 2$, $T = 1$, $K = 90$, $\lambda = 1$, $r = 0.1$, $L = 80$, $M = 150$ and $N = 100$

The above figure is an approximation to the solution of (6.38) at time 0 with the appropriate boundary conditions for a European put option. We see that, as expected, the option has higher value for lower share prices. The interesting behaviour occurs when the deviation from the mean (D) is close to zero, we see that when this is the case the value of the call option is lower than for high magnitude D. This is a result of the way we chose our volatility function, which is lower for lower magnitudes of D. As in the Black-Scholes framework, less volatility implies lower price, and so we see a depression in the value of the solution around $D = 0$.

A Crank-Nicolson scheme of this form can be shown to be unconditionally stable for equations of the form of (6.4). For a detailed stability analysis see [12]. This is, of course, an extremely desirable property and one we hope will be maintained under the introduction of our leverage function.

6.7 The leveraged model revisited

Introducing the leverage function into the model described above gives the following market model,

$$dB_t = rB_t dt \tag{6.35}$$

$$dD_t = (\mu(D_t) - \lambda D_t)dt + \sigma(D_t)\ell(t, Z_t)dW_t \tag{6.36}$$

$$dZ_t = \mu(D_t)dt + \sigma(D_t)dW_t \tag{6.37}$$

The system yields the following PDE for the price of a contingent claim V, paying $\psi(S_T)$ at time T for some measurable $\psi(.)$,

$$\begin{aligned} 0 = &\left(rS\frac{\partial V}{\partial S} - ru - \lambda D\frac{\partial V}{\partial D} - \frac{\partial V}{\partial t}\right) \\ &+ \left(-\frac{1}{2}\frac{\partial V}{\partial D} + \frac{1}{2}S^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\frac{\partial^2 V}{\partial D^2} + S\frac{\partial^2 V}{\partial S \partial D}\right)\sigma^2(D)\ell^2(t, S). \end{aligned} \tag{6.38}$$

Using our approximation for the leverage function from the previous sections we have that

$$\begin{aligned} 0 = &\left(rS\frac{\partial V}{\partial S} - ru - \lambda D\frac{\partial V}{\partial D} - \frac{\partial V}{\partial t}\right) \\ &+ \left(-\frac{1}{2}\frac{\partial V}{\partial D} + \frac{1}{2}S^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\frac{\partial^2 V}{\partial D^2} + S\frac{\partial^2 V}{\partial S \partial D}\right)\sigma^2_{LV}(t, S). \end{aligned} \tag{6.39}$$

We will attempt to price a European put option, with strike K and maturity T, in the framework of the above model, i.e $\Phi(S_T) = \max(0, K - S_T)$.

We follow a similar procedure for our boundary conditions, derivative approximations and discretisation and so we will not present this in detail. We consider the same local volatility surface as in the previous section.

6.8 Matrix construction

To demonstrate the final matrix equation it is, again, most useful to consider the case where $M = 4$ and $L = 4$. In this case we have the following

$$
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathcal{M}_{2,1} & 0 & 0 & 0 & \mathcal{L}_{2,1} & 0 & 0 & 0 & \mathcal{K}_{2,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathcal{A}^n_{2,2} & (\mathcal{B}^n_{2,2} + \frac{rS_2}{2h}) & -\mathcal{A}^n_{2,2} & 0 & (-\mathcal{C}^n_{2,2} - \frac{\sigma_2^2}{2d^2}) & \mathcal{D}^n_{2,2} & \mathcal{C}^n_{2,2} & 0 & -\mathcal{A}^n_{2,2} & \mathcal{B}^n_{2,2} & \mathcal{A}^n_{2,2} & 0 & 0 & 0 & 0 & 0 \\
0 & \mathcal{A}^n_{2,3} & (\mathcal{B}^n_{2,3} + \frac{rS_2}{2h}) & -\mathcal{A}^n_{2,3} & 0 & (-\mathcal{C}^n_{2,3} - \frac{\sigma_3^2}{2d^2}) & \mathcal{D}^n_{2,3} & \mathcal{C}^n_{2,3} & 0 & -\mathcal{A}^n_{2,3} & \mathcal{B}^n_{2,3} & \mathcal{A}^n_{2,3} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \mathcal{M}_{2,4} & 0 & 0 & 0 & \mathcal{L}_{2,4} & 0 & 0 & 0 & \mathcal{K}_{2,4} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \mathcal{M}_{3,1} & 0 & 0 & 0 & \mathcal{L}_{3,1} & 0 & 0 & 0 & \mathcal{K}_{3,1} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \mathcal{A}^n_{3,2} & (\mathcal{B}^n_{3,2} + \frac{rS_3}{2h}) & -\mathcal{A}^n_{3,2} & 0 & (-\mathcal{C}^n_{3,2} - \frac{\sigma_2^2}{2d^2}) & \mathcal{D}^n_{3,2} & \mathcal{C}^n_{3,2} & 0 & -\mathcal{A}^n_{3,2} & \mathcal{B}^n_{3,2} & \mathcal{A}^n_{3,2} & 0 \\
0 & 0 & 0 & 0 & 0 & \mathcal{A}^n_{3,3} & (\mathcal{B}^n_{3,3} + \frac{rS_3}{2h}) & -\mathcal{A}^n_{3,3} & 0 & (-\mathcal{C}^n_{3,3} - \frac{\sigma_3^2}{2d^2}) & \mathcal{D}^n_{3,3} & \mathcal{C}^n_{3,3} & 0 & -\mathcal{A}^n_{3,3} & \mathcal{B}^n_{3,3} & \mathcal{A}^n_{3,3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathcal{M}_{3,4} & 0 & 0 & 0 & \mathcal{L}_{3,4} & 0 & 0 & 0 & \mathcal{K}_{3,4} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix} \vec{f}^{n+1} =
$$

$$
\begin{pmatrix}
e^{-rk} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & e^{-rk} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & e^{-rk} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & e^{-rk} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\mathcal{M}_{2,1} & 0 & 0 & 0 & -\mathcal{L}_{2,1} + \frac{2}{k} - r & 0 & 0 & 0 & -\mathcal{K}_{2,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\mathcal{A}^n_{2,2} & -(\mathcal{B}^n_{2,2} + \frac{rS_2}{2h}) & \mathcal{A}^n_{2,2} & 0 & -(-\mathcal{C}^n_{2,2} - \frac{\sigma_2^2}{2d^2}) & -\mathcal{D}^n_{2,2} + \frac{2}{k} - r & -\mathcal{C}^n_{2,2} & 0 & \mathcal{A}^n_{2,2} & -\mathcal{B}^n_{2,2} & -\mathcal{A}^n_{2,2} & 0 & 0 & 0 & 0 & 0 \\
0 & -\mathcal{A}^n_{2,3} & -(\mathcal{B}^n_{2,3} + \frac{rS_2}{2h}) & \mathcal{A}^n_{2,3} & 0 & -(-\mathcal{C}^n_{2,3} - \frac{\sigma_3^2}{2d^2}) & -\mathcal{D}^n_{2,3} + \frac{2}{k} - r & -\mathcal{C}^n_{2,3} & 0 & \mathcal{A}^n_{2,3} & -\mathcal{B}^n_{2,3} & -\mathcal{A}^n_{2,3} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\mathcal{M}_{2,4} & 0 & 0 & 0 & -\mathcal{L}_{2,4} + \frac{2}{k} - r & 0 & 0 & 0 & -\mathcal{K}_{2,4} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\mathcal{M}_{3,1} & 0 & 0 & 0 & -\mathcal{L}_{3,1} + \frac{2}{k} - r & 0 & 0 & 0 & -\mathcal{K}_{3,1} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\mathcal{A}^n_{3,2} & -(\mathcal{B}^n_{3,2} + \frac{rS_3}{2h}) & \mathcal{A}^n_{3,2} & 0 & -(-\mathcal{C}^n_{3,2} - \frac{\sigma_2^2}{2d^2}) & -\mathcal{D}^n_{3,2} + \frac{2}{k} - r & -\mathcal{C}^n_{3,2} & 0 & \mathcal{A}^n_{3,2} & -\mathcal{B}^n_{3,2} & -\mathcal{A}^n_{3,2} & 0 \\
0 & 0 & 0 & 0 & 0 & -\mathcal{A}^n_{3,3} & -(\mathcal{B}^n_{3,3} + \frac{rS_3}{2h}) & \mathcal{A}^n_{3,3} & 0 & -(-\mathcal{C}^n_{3,3} - \frac{\sigma_3^2}{2d^2}) & -\mathcal{D}^n_{3,3} + \frac{2}{k} - r & -\mathcal{C}^n_{3,3} & 0 & \mathcal{A}^n_{3,3} & -\mathcal{B}^n_{3,3} & -\mathcal{A}^n_{3,3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\mathcal{M}_{3,4} & 0 & 0 & 0 & -\mathcal{L}_{3,4} + \frac{2}{k} - r & 0 & 0 & 0 & -\mathcal{K}_{3,4} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix} \vec{f}^{n}
$$

.

Where,

$$\mathcal{A}^n_{i,j} = -\frac{S_i\left(\Sigma^n_{i,j}\right)^2}{8hd} \tag{6.40}$$

$$\mathcal{B}^n_{i,j} = -\frac{rS_i}{4h} - \frac{S_i^2\left(\Sigma^n_{i,j}\right)^2}{4h^2} \tag{6.41}$$

$$\mathcal{C}^n_{i,j} = \frac{\lambda D_j}{4d} + \frac{\left(\Sigma^n_{i,j}\right)^2}{8d} - \frac{\left(\Sigma^n_{i,j}\right)^2}{4d^2} \tag{6.42}$$

$$\mathcal{D}^n_{i,j} = \frac{1}{k} + \frac{S_i^2\left(\Sigma^n_{i,j}\right)^2}{2h^2} + \frac{\left(\Sigma^n_{i,j}\right)^2}{2d^2}$$

$$\mathcal{K}_i = -\frac{rS_i}{4h} - \frac{\sigma^2_{max}S_i^2}{4h^2} \tag{6.43}$$

$$\mathcal{L}_i = \frac{1}{k} + \frac{\sigma^2_{max}S_i^2}{2h^2} \tag{6.44}$$

$$\mathcal{M}_i = \frac{rS_i}{4h} - \frac{\sigma^2_{max}S_i^2}{2h^2}. \tag{6.45}$$

We have used the notation $\Sigma^n_{i,j} := \sigma_{LV}(t_n, S_i)$. The construction algorithm is presented in detail in the form of a Matlab code, this is attached as appendix 3. We note that the most important difference between the two schemes is that the coefficient matrices now depend on time.

6.9 Numerical results

The above scheme was solved on the same mesh as in the previous case. The results of this numerical experiment are presented graphically below.

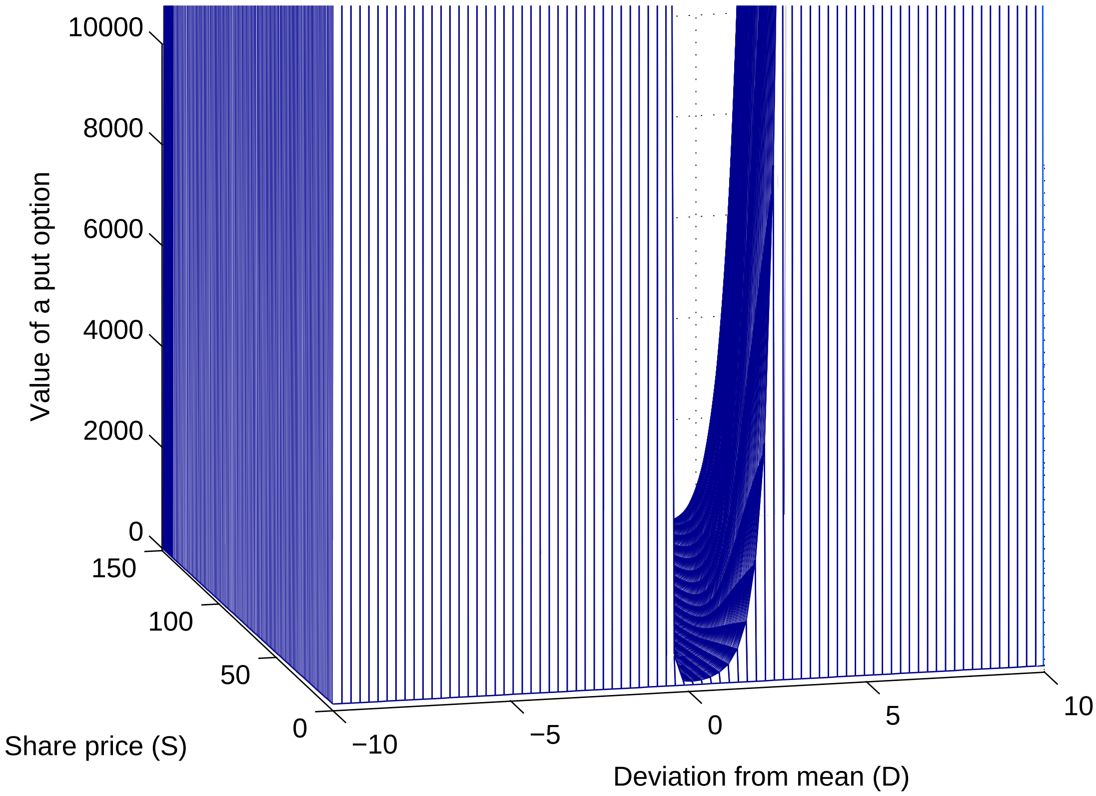


Figure 4: Again, we use: $\epsilon = 5$, $\eta = 0.2$, $\sigma_{max} = 2$, $T = 1$, $K = 90$, $\lambda = 1$, $r = 0.2$, $L = 80$, $M = 150$ and $N = 100$

Clearly we see an explosion of the approximate solution on this mesh. This shows that our scheme is not unconditionally stable when used for the leveraged case. Therefore, additional considerations need to be made when considering a scheme for this case. The time dependence in our volatility term that was introduced by the leverage function is the likely culprit in our method losing stability.

6.10 Efficiency

It is important to make a brief comment on the efficiency of our algorithm. The code in appendix 2 was run on a computer with 4GB of installed memory (RAM). For the mesh and parameter values chosen in order to generate figure 3, the MATLAB code takes approximately an hour to give an output. This is extremely slow and, in addition, any attempt to refine the mesh was unsuccessful due to a lack of memory.

Generating an entire implied volatility surface would involve running the code over thousands of maturities and strikes, and because of the slowness of the scheme, any such procedure remains impractical. This highly motivates the use of a change of variables, as this makes it possible to generate a solution for multiple strikes in one run by scaling. [12].

7 Conclusion

The goal of this dissertation was to find a fitted, complete and Markovian market model, for which we can derive a pricing PDE. By synthesizing the work of Hobson and Rogers, Foschi, Dupire, Gyöngy and Guyon we have achieved this goal in the theoretical setting constructed in the preliminary chapters.

Through this exercise we have demonstrated the appeal of path-dependent volatility models. However, despite the steady progress of the theory, much must still be done to develop the application of the models before they can be considered practical.

By demonstrating the particle method and a finite difference scheme for a particular case, we have demonstrated some of the challenges faced when attempting to design an efficient numerical scheme in the path-dependent volatility framework. We also suggested some possible considerations for making such schemes more efficient.

Further areas of study in this regard may include a more rigorous method for calculating leverage functions, and setting out clear methodology for calibrating such models to the market. A comprehensive study of the stability of numerical methods for fitted models will aid in the quest of finding a more appropriate scheme for these models.

The study of path-dependent volatility models is still in its infancy but the potential of this class of models is essentially limitless.

Appendix 1: Matlab Code -Particle Method

```
epsilon=5;
eta=0.2;
sigmax=5;
T=2;
K=90;
lambda=1;
r=0.1;
n=input('number of timesteps=');
N=input('number of simulations=');
t=linspace(0.01,T,n);
dt=t(2)-t(1);
vareps=0.005;
D=zeros(N,n);
Z=zeros(N,n);
Z(:,1)=log(50);
D(:,1)=0;
nA=zeros(N,n);
DA=zeros(N,n);
cntit=0;
for i=1:n-1;

    for j=1:N;
        for k=1:N;
            if Z(j,i)-vareps<=Z(k,i) && Z(k,i)<=Z(j,i)
                +vareps
                nA(j,i)=nA(j,i)+1;
                DA(j,i)=DA(j,i)+vol(eta,epsilon,sigmax
                    ,D(k,i))^2;
            else nA(j,i)=nA(j,i);
                DA(j,i)=DA(j,i);
            end
        end
        Z(j,i+1)=Z(j,i)+vol(eta,epsilon,sigmax,D(j,i))
            *(LV(t(i),Z(j,i))*normrnd(0,1)*sqrt(dt))/
            sqrt(((1/nA(j,i))*DA(j,i)))-(1/2)*(vol(eta,
            epsilon,sigmax,D(j,i))*LV(t(i),Z(j,i)))^2*
            dt/((1/nA(j,i))*DA(j,i));
        D(j,i+1)=D(j,i)+Z(j,i+1)-Z(j,i)-lambda*D(j,i)*
            dt;
    end
    cntit=cntit+1;
```

```
        disp(cntit)
end
S=zeros(N,n);
for p=1:N
S(p,:)=exp(Z(p,:)+r*t);
end
for c=1:30
        plot(t,S(c,:))
        hold on
end
plot(t,Z(1:20,:))
%%
Z1=0:0.1:150;
t1=0.01:0.001:2;
for k1=1:length(Z1)
        for k2=1:length(t1);
LV1(t1(k2),Z1(k1))=LV(t1(k2),Z1(k1));
        end
end
```

Appendix 2: Matlab Code - Non-leveraged finite difference

```
L=80;%number of points D axis
M=150;%number of points S axis
N=100;%number of time steps
T=1; %maturity
t0=0;%initial time
d0=-10; %min D value
dm=10; %max D valiue
s0=0; % min S
sm=150;%max S
VM=2; %max volatility
eta=0.2; %vol parameter
eps=5;%vol parameter
t=linspace(t0,T,N);
D=linspace(d0,dm,L); %D space
S=linspace(s0,sm,M); %S space
d=(dm-d0)/(L); %D step size
h=(sm-s0)/(M); %dS
k=(T-t0)/(N);
r=0.1;%risk free
lam=1;%model par
```

```
X=120; %strike
 sig=min(eta*sqrt(1+eps*D.^2),VM);%vol function
 f=zeros(L*M,2);
%%
%initial condition
 for p=1:M
       f((p-1)*L+1:p*L,1)=max(0,X-S(p));
 end

%coefficient matrices
 for i=1:M;
       for j=1:L;
 A11(i,j)=-(S(i)*sig(j)^2)/(8*h*d);
 B11(i,j)=-(r*S(i))/(4*h)-(S(i)^2*sig(j)^2)/(4*h^2);
 C11(i,j)=(lam*D(j))/(4*d)-sig(j)^2/(4*d^2)+(sig(j)^2)
     /8*d;
 D11(i,j)=1/k+((S(i)^2)*sig(j)^2)/(2*h^2)+sig(j)^2/(2*d
     ^2);
 K11(i,j)=-(r*S(i))/(4*h)-((VM^2)*S(i)^2)/(4*h^2);
 L11(i,j)=1/k+((VM^2)*S(i)^2)/(2*h^2);
 M11(i,j)=(r*S(i))/(4*h)-((VM^2)*S(i)^2)/(4*h^2);
       end
 end


%matrix construction
%LHS
LHS=zeros(M*L,M*L);
  %low S BC
  LowS=eye(L,M*L);
  LHS(1:L,1:M*L)=LowS;
  clearvars LowS
  %low D BC
  LowD=zeros(M*L,M*L);
  for c1=1:M-2;
         LowD(c1*L+1,(c1-1)*L+1)=M11(c1+1,1);
         LowD(c1*L+1,c1*L+1)=L11(c1+1,1);
         LowD(c1*L+1,(c1+1)*L+1)=K11(c1+1,1);
  end
  LHS=LHS+LowD;
  clearvars LowD
  %High D BC
  HighD=zeros(M*L,M*L);
  for c2=1:M-2;
```

```
        HighD((c2+1)*L,c2*L)=M11(c2+1,L);
        HighD((c2+1)*L,(c2+1)*L)=L11(c2+1,L);
        HighD((c2+1)*L,(c2+2)*L)=K11(c2+1,L);
    end
    LHS=LHS+HighD;
    clearvars HighD
    %High S BC
    highS=eye(L*M,L*M);
    LHS(M*L-(L-1):M*L,M*L-(L-1):M*L)=highS(M*L-(L-1):M*L,M*L-(L-1):M*L);
    clearvars highS

    %interior matrix

    Int=zeros(M*L,M*L);
    for c5=1:M-2; %block
        for c6=1:L-2; %row
            Int(L*c5+c6+1,(c5-1)*L+c6:(c5-1)*L+c6+2)=[A11(c5+1,c6+1) (B11(c5+1,c6+1)+(r*S(c5+1))/(2*h)) -A11(c5+1,c6+1)];
            Int(L*c5+c6+1,c5*L+c6:c5*L+c6+2)=[(-C11(c5+1,c6+1)-(sig(c6+1)^2)/(2*d^2)) D11(c5+1,c6+1) C11(c5+1,c6+1)];
            Int(L*c5+c6+1,(c5+1)*L+c6:(c5+1)*L+c6+2)=[-A11(c5+1,c6+1) B11(c5+1,c6+1) A11(c5+1,c6+1)];
        end
    end
    %Final Left hand matrix
    LHS=LHS+Int;
    clearvars Int

%RHS
RHS=zeros(M*L,M*L);
    %low S BC
    rlowS=eye(L,M*L)*exp(-r*k);
    RHS(1:L,1:M*L)=rlowS;
    clearvars rlowS
    %low D BC
    rLowD=zeros(M*L,M*L);
    for c1=1:M-2;
        rLowD(c1*L+1,(c1-1)*L+1)=-M11(c1+1,1);
        rLowD(c1*L+1,c1*L+1)=-L11(c1+1,1)+2/k-r;
        rLowD(c1*L+1,(c1+1)*L+1)=-K11(c1+1,1);
```

```
end
RHS=RHS+rLowD;
clearvars rLowD
%High D BC
rHighD=zeros(M*L,M*L);
for c2=1:M-2;
    rHighD((c2+1)*L,c2*L)=-M11(c2+1,L);
    rHighD((c2+1)*L,(c2+1)*L)=-L11(c2+1,L)+2/k-r;
    rHighD((c2+1)*L,(c2+2)*L)=-K11(c2+1,L);
end
RHS=RHS+rHighD;
clearvars rHighD
%High S BC
rhighS=eye(L*M,L*M);
RHS(M*L-(L-1):M*L,M*L-(L-1):M*L)=rhighS(M*L-(L-1):M*L,M*L-(L-1):M*L);
clearvars rhighS

%interior matrix

rInt=zeros(M*L,M*L);
for c5=1:M-2; %block
    for c6=1:L-2; %row
        rInt(L*c5+c6+1,(c5-1)*L+c6:(c5-1)*L+c6+2)=[-A11(c5+1,c6+1) -(B11(c5+1,c6+1)+(r*S(c5+1))/(2*h)) A11(c5+1,c6+1)];
        rInt(L*c5+c6+1,c5*L+c6:c5*L+c6+2)=[-(-C11(c5+1,c6+1)-(sig(c6+1)^2)/(2*d^2)) (-D11(c5+1,c6+1)+2/k-r) -C11(c5+1,c6+1)];
        rInt(L*c5+c6+1,(c5+1)*L+c6:(c5+1)*L+c6+2)=[A11(c5+1,c6+1) -B11(c5+1,c6+1) -A11(c5+1,c6+1)];
    end
end

%Final Right hand matrix
clearvars A11 B11 C11 D11 E11 K11 L11 M11
RHS=RHS+rInt;
clearvars rInt
%%
%matrix equation
check=0;
for n=1:N-1
f(:,2)=linsolve(LHS,RHS*f(:,1));
```

```
  f(:,1)=f(:,2);
  check=check+1;
  disp(check)
end
 %final value as matrix
f1=zeros(M,L);

 for p1=1:M
     f1(p1,:)=f((p1-1)*L+1:p1*L,2);
 end
%mesh(D,S,f1), axis([-20 20 0 200 -50 150]), drawnow
%end
%figure(1)
mesh(D,S,f1)
```

Appendix 3: Matlab Code - Leveraged finite difference

```
L=80;%number of points D axis
M=150;%number of points S axis
N=100;%number of time steps
T=1; %maturity
t0=0.001;%initial time
d0=-10; %min D value
dm=10; %max D valiue
s0=0; % min S
sm=150;%max S
eta=0.2; %vol parameter
eps=5;%vol parameter
t=linspace(t0,T,N);
D=linspace(d0,dm,L); %D space
S=linspace(s0,sm,M); %S space
d=(dm-d0)/(L); %D step size
h=(sm-s0)/(M); %dS
k=(T-t0)/(N);
r=0.1;%risk free
lam=1;%model par
X=90; %strike
VM=LV(t(N),exp(-r*t(N))*S(M)); %max volatility
sig=min(eta*sqrt(1+eps*D.^2),VM);%vol function
f=zeros(L*M,N);
%%
%initial condition
```

```matlab
for p=1:M
        f((p-1)*L+1:p*L,1)=max(0,X-S(p));
end

count=0;
for n=1:N-1
%coefficient matrices
for i=1:M;
        for j=1:L;
A11(i,j)=-(S(i)*LV(T-t(n),exp(-r*t(n))*S(i))^2)/(8*h*d);
B11(i,j)=-(r*S(i))/(4*h)-(S(i)^2*LV(T-t(n),exp(-r*t(n))*S(i))^2)/(4*h^2);
C11(i,j)=(lam*D(j))/(4*d)-LV(T-t(n),exp(-r*t(n))*S(i))^2/(4*d^2)+(LV(T-t(n),exp(-r*t(n))*S(i))^2)/8*d;
D11(i,j)=1/k+((S(i)^2)*LV(T-t(n),exp(-r*t(n))*S(i))^2)/(2*h^2)+LV(T-t(n),exp(-r*t(n))*S(i))^2/(2*d^2);
K11(i,j)=-(r*S(i))/(4*h)-((VM^2)*S(i)^2)/(4*h^2);
L11(i,j)=1/k+((VM^2)*S(i)^2)/(2*h^2);
M11(i,j)=(r*S(i))/(4*h)-((VM^2)*S(i)^2)/(4*h^2);
        end
end


%matrix construction
%LHS
LHS=zeros(M*L,M*L);
 %low S BC
 LowS=eye(L,M*L);
 LHS(1:L,1:M*L)=LowS;
 clearvars LowS
 %low D BC
 LowD=zeros(M*L,M*L);
 for c1=1:M-2;
        LowD(c1*L+1,(c1-1)*L+1)=M11(c1+1,1);
        LowD(c1*L+1,c1*L+1)=L11(c1+1,1);
        LowD(c1*L+1,(c1+1)*L+1)=K11(c1+1,1);
 end
 LHS=LHS+LowD;
 clearvars LowD
 %High D BC
 HighD=zeros(M*L,M*L);
 for c2=1:M-2;
        HighD((c2+1)*L,c2*L)=M11(c2+1,L);
```

```
        HighD((c2+1)*L,(c2+1)*L)=L11(c2+1,L);
        HighD((c2+1)*L,(c2+2)*L)=K11(c2+1,L);
    end
    LHS=LHS+HighD;
    clearvars HighD
    %High S BC
    highS=eye(L*M,L*M);
    LHS(M*L-(L-1):M*L,M*L-(L-1):M*L)=highS(M*L-(L-1):M*L,M*L-(L-1):M*L);
    clearvars highS

    %interior matrix

    Int=zeros(M*L,M*L);
    for c5=1:M-2; %block
        for c6=1:L-2; %row
            Int(L*c5+c6+1,(c5-1)*L+c6:(c5-1)*L+c6+2)=[A11(c5+1,c6+1) (B11(c5+1,c6+1)+(r*S(c5+1))/(2*h)) -A11(c5+1,c6+1)];
            Int(L*c5+c6+1,c5*L+c6:c5*L+c6+2)=[(-C11(c5+1,c6+1)-(sig(c6+1)^2)/(2*d^2)) D11(c5+1,c6+1) C11(c5+1,c6+1)];
            Int(L*c5+c6+1,(c5+1)*L+c6:(c5+1)*L+c6+2)=[-A11(c5+1,c6+1) B11(c5+1,c6+1) A11(c5+1,c6+1)];
        end
    end
    %Final Left hand matrix
    LHS=LHS+Int;
    clearvars Int

%RHS
RHS=zeros(M*L,M*L);
    %low S BC
    rlowS=eye(L,M*L)*exp(-r*k);
    RHS(1:L,1:M*L)=rlowS;
    clearvars rlowS
    %low D BC
    rLowD=zeros(M*L,M*L);
    for c1=1:M-2;
        rLowD(c1*L+1,(c1-1)*L+1)=-M11(c1+1,1);
        rLowD(c1*L+1,c1*L+1)=-L11(c1+1,1)+2/k-r;
        rLowD(c1*L+1,(c1+1)*L+1)=-K11(c1+1,1);
    end
```

```
RHS=RHS+rLowD;
clearvars rLowD
%High D BC
rHighD=zeros(M*L,M*L);
for c2=1:M-2;
      rHighD((c2+1)*L,c2*L)=-M11(c2+1,L);
      rHighD((c2+1)*L,(c2+1)*L)=-L11(c2+1,L)+2/k-r;
      rHighD((c2+1)*L,(c2+2)*L)=-K11(c2+1,L);
end
RHS=RHS+rHighD;
clearvars rHighD
%High S BC
rhighS=eye(L*M,L*M);
RHS(M*L-(L-1):M*L,M*L-(L-1):M*L)=rhighS(M*L-(L-1):M*L,M*L-(L-1):M*L);
clearvars rhighS

%interior matrix

rInt=zeros(M*L,M*L);
for c5=1:M-2; %block
      for c6=1:L-2; %row
            rInt(L*c5+c6+1,(c5-1)*L+c6:(c5-1)*L+c6+2)=[-A11(c5+1,c6+1) -(B11(c5+1,c6+1)+(r*S(c5+1))/(2*h)) A11(c5+1,c6+1)];
            rInt(L*c5+c6+1,c5*L+c6:c5*L+c6+2)=[-(-C11(c5+1,c6+1)-(sig(c6+1)^2)/(2*d^2)) (-D11(c5+1,c6+1)+2/k-r) -C11(c5+1,c6+1)];
            rInt(L*c5+c6+1,(c5+1)*L+c6:(c5+1)*L+c6+2)=[A11(c5+1,c6+1) -B11(c5+1,c6+1) -A11(c5+1,c6+1)];
      end
end

%Final Right hand matrix
%clearvars A11 B11 C11 D11 E11 K11 L11 M11
RHS=RHS+rInt;
clearvars rInt
%%
%matrix equation

f(:,n+1)=linsolve(LHS,RHS*f(:,n));
count=count+1;
disp(count)
```

```
end
 %%
%final value as matrix
f1=zeros(M,L);

 for p1=1:M
     f1(p1,:)=f((p1-1)*L+1:p1*L,n+1);
 end
 mesh(D,S,f1),axis([-10 10 0 150 -50 10000])

%mesh(D,S,f1), axis([-20 20 0 200 -50 150]), drawnow
%end
%figure(1)
```

References

[1] Carol Alexander and Leonardo M Nogueira. Stochastic local volatility. *Available at SSRN 1107685*, 2008.

[2] Mercedes Arriojas, Yaozhong Hu, Salah-Eldin Mohammed, and Gyula Pap. A delayed black and scholes formula. *Stochastic Analysis and Applications*, 25(2):471–492, 2007.

[3] Tomas Björk. *Arbitrage theory in continuous time*. Oxford university press, 2004.

[4] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The journal of political economy*, pages 637–654, 1973.

[5] Svetlana Borovkova and Ferry J Permana. Implied volatility in oil markets. *Computational Statistics & Data Analysis*, 53(6):2022–2039, 2009.

[6] D Brecher. Pushing the limits of local volatility in option pricing. *Wilmott Magazine, September*, 6:15, 2006.

[7] Gerard Brunick, Steven Shreve, et al. Mimicking an itô process by a solution of a stochastic differential equation. *The Annals of Applied Probability*, 23(4):1584–1628, 2013.

[8] Zdzislaw Brzezniak and Tomasz Zastawniak. *Basic stochastic processes: a course through exercises*. Springer Science & Business Media, 1999.

[9] Gearoid De Barra. *Measure theory and integration*. Elsevier, 2003.

[10] Emanuel Derman and Iraj Kani. Riding on a smile. *In Goldman Sachs Quantitative Strategies Research Notes*, 1994.

[11] Emanuel Derman, Iraj Kani, and Michael Kamal. Trading and hedging local volatility. *Journal of Financial Engineering*, 6:233–268, 1997.

[12] Marco Di Francesco, Paolo Foschi, and Andrea Pascucci. Analysis of an uncertain volatility model. *J. Appl. Math. Decis. Sci*, 17:15609, 2006.

[13] Bernard Dumas, Jeff Fleming, and Robert E Whaley. Implied volatility functions: Empirical tests. *The Journal of Finance*, 53(6):2059–2106, 1998.

[14] Bruno Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1993.

[15] Bruno Dupire. A unified theory of volatility. *Derivatives Pricing: The Classic Collection*, pages 185–196, 1996.

[16] Matthias R Fengler. Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance*, 9(4):417–428, 2009.

[17] Paolo Foschi and Andrea Pascucci. Path dependent volatility. *Decisions in Economics and Finance*, 31(1):13–32, 2008.

[18] Jim Gatheral. *The volatility surface: a practitioner's guide*, volume 357. John Wiley & Sons, 2011.

[19] Paul Glasserman and Qi Wu. Forward and future implied volatility. *International Journal of Theoretical and Applied Finance*, 14(03):407–432, 2011.

[20] Julien Guyon. Path-dependent volatility. 2014.

[21] Julien Guyon and Pierre Henry-Labordère. *Nonlinear option pricing*. CRC Press, 2013.

[22] István Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an itô differential. *Probability theory and related fields*, 71(4):501–516, 1986.

[23] Patrick S Hagan, Deep Kumar, Andrew S Lesniewski, and Diana E Woodward. Managing smile risk. *The Best of Wilmott*, page 249, 2002.

[24] Martin Hairer. On malliavin's proof of hörmander's theorem. *Bulletin des sciences mathematiques*, 135(6):650–666, 2011.

[25] Vera Hallulli and Tiziano Vargiolu. Financial models with dependence on the past: a survey. *SERIES ON ADVANCES IN MATHEMATICS FOR APPLIED SCIENCES*, 69:348, 2005.

[26] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2):327–343, 1993.

[27] David G Hobson, Leonard CG Rogers, et al. Complete models with stochastic volatility. *Mathematical Finance*, 8(1):27–48, 1998.

[28] Nabil Kahale. An arbitrage-free interpolation of volatilities. *Risk*, 17(5):102–106, 2004.

[29] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.

[30] Antonie Kotzé, Rudolf Oosthuizen, and Edson Pindza. Implied and local volatility surfaces for south african index and foreign exchange options. *Journal of Risk and Financial Management*, 8(1):43–82, 2015.

[31] Steven G Kou. A jump-diffusion model for option pricing. *Management science*, 48(8):1086–1101, 2002.

[32] NV Krylov. On the relation between differential operators of second order and the solutions of stochastic differential equations. In *Steklov Seminar*, volume 1985, 1984.

[33] Dilip B Madan, Peter P Carr, and Eric C Chang. The variance gamma process and option pricing. *European finance review*, 2(1):79–105, 1998.

[34] Xuerong Mao. *Stochastic differential equations and applications*. Elsevier, 2007.

[35] Philip E Protter. Stochastic differential equations. In *Stochastic Integration and Differential Equations*, pages 249–361. Springer, 2005.

[36] Malempati M Rao and Randall J Swift. *Probability theory with applications*, volume 582. Springer Science & Business Media, 2006.

[37] Riccardo Rebonato. *Volatility and correlation: the perfect hedger and the fox*. John Wiley & Sons, 2005.