Chapter 1

Introduction

An option is a financial contract between two parties. One party pays a premium to the other, giving them the right, but not the obligation, to buy or sell a specified asset (often a stock) at a certain time for a fixed price. Options can be used to manage risk, speculate on the stockmarket, generate additional income on shares, or provide leverage (the potential to earn a greater return from a smaller initial outlay). Since first being traded on an organised exchange in 1973, the market for options has experienced dramatic growth, and today huge numbers of options are traded daily. Consequently, option pricing has become an important area of current research in financial mathematics.

In order to price a general option, two features must first be decided upon. A model for the underlying asset on which the option relies must be specified, and a method for pricing the option must be chosen. In this thesis, a new model, called the “Switching Black-Scholes” model, for the underlying asset is proposed, and various option pricing methodologies are then applied. Before discussing in further detail the motivation for proposing this model and the option pricing methodologies used, we first give a brief introduction to the financial concepts relevant to this work. For a more thorough presentation of these ideas, see Elliott and Kopp ([29]).
1.1 Introduction to some Financial Concepts

As mentioned above, the holder of an option has the right, but not the obligation, to buy or sell a specified asset, $S$, at a certain time, $T$, for a fixed price, $K$. If the right is to buy the asset, the option is called a call option, whereas if the right is to sell the asset, it is called a put option. The date at which this transaction is to take place is called the expiry date of the option, and $K$ is known as the strike price. From these definitions, we see that the value of a call option at expiry is,

$$C(T) = (S(T) - K)^+ := \begin{cases} S(T) - K, & \text{if } S(T) > K, \\ 0, & \text{otherwise}, \end{cases}$$

and the value of a put option at expiry is,

$$P(T) = (K - S(T))^+ := \begin{cases} K - S(T), & \text{if } K > S(T), \\ 0, & \text{otherwise}. \end{cases}$$

The field of option pricing is concerned with finding the current value of call options, put options, and other more complicated options, using their known expiry values.

To discuss how a current option value might be found, consider a frictionless\footnote{“frictionless” means that there are no transaction costs, all assets are perfectly divisible, the assets do not pay dividends, and short sales (selling an asset you do not own) are permitted.} market consisting of one risky asset and one riskless asset (or bond), both defined on a complete probability space $(\Omega, \mathcal{F}, P)$. Suppose that these assets can be traded at discrete times $t_n = n\tau$ for $n = 0, 1, \ldots, N$, where $\tau > 0$ and $N \in \mathbb{N}$ are fixed, and let $T = N\tau$. Denote by $S(t)$ and $B(t)$ the values of the risky asset and bond, respectively, at a given time $t$, and define,

$$S_n = S(t_n), \quad B_n = B(t_n),$$

for $n = 0, 1, \ldots, N$. We assume that $B_0 = 1$ and both $(S_n)$ and $(B_n)$ are strictly positive and adapted to a filtration $\mathbb{F} = (\mathcal{F}_n)_{n=0,1,\ldots,N}$ of $\mathcal{F}$, which represents...
the evolution of information available in the market. The discounted value of
an asset is found by dividing the asset price by the bond value at that time. In
the current setting, the discounted asset price is,

\[ X_n = \frac{S_n}{B_n}, \]

and we write the change in the discounted price as,

\[ \Delta X_n = X_n - X_{n-1}. \]

A trading strategy in this market is an \( \mathbb{R}^2 \)-valued stochastic process
\( \phi = (\eta_n, \xi_n)_{n=1,2,\ldots,N} \), where \( \eta \) and \( \xi \) are predictable with respect to the fil-
tration \( \mathbb{F} \). The values \( \eta_n \) and \( \xi_n \) represent, respectively, the number of bonds
and the number of risky assets held in the time interval \([t_{n-1}, t_n]\), and the pair
\( \phi_n = (\eta_n, \xi_n) \) is called the portfolio of assets held during this time.

Given a trading strategy \( \phi \), the value at time \( t_n \) of the portfolio (before any
changes in the holdings are made) is,

\[ V_n(\phi) = \eta_n B_n + \xi_n S_n, \]

and the discounted value at time \( t_n \) is,

\[ V_n^*(\phi) = \eta_n + \xi_n X_n. \]

A trading strategy is said to be self-financing if,

\[ \eta_n B_n + \xi_n S_n = \eta_{n+1} B_n + \xi_{n+1} S_n, \]

for \( n = 0, 1, \ldots, N - 1 \), which means that no inflow or outflow of capital from
the portfolio is permitted. Straightforward arguments show that a trading stra-
thesis \( \phi \) is self-financing if and only if,

\[ V_n^*(\phi) = V_0(\phi) + \sum_{k=1}^{n} \xi_k \Delta X_k, \]

for all \( n \). The term \( V_0(\phi) \) can be interpreted as the initial investment at time
0 required to follow the strategy \( \phi \). The sum \( \sum_{k=1}^{n} \xi_k \Delta X_k \) represents the dis-
counted gains from trading according to this strategy, accumulated until time
\( t_n \).
A contingent claim in this market is a financial product depending on the risky asset, $S$, which is mathematically represented by an $\mathcal{F}_T$-measurable, non-negative random variable, $H$, satisfying $E[H] < \infty$. The claim is said to be attainable if $H$ coincides with the terminal value of a self-financing trading strategy $\phi$, called the replicating strategy. This means $V_N(\phi) = H$. Written in an alternative form, the claim $H$ is attainable if,

$$\frac{H}{B_N} = V_0(\phi) + \sum_{k=1}^{N} \xi_k \Delta X_k,$$

for some self-financing trading strategy, $\phi$. A financial market is said to be complete if every contingent claim is attainable. Otherwise, it is said to be incomplete.

The method most often used to price contingent claims, including options, is based on the principle of no-arbitrage, which states that arbitrage opportunities, in which risk-free gains from trading risky assets can be realised, should not be possible\(^2\). In the case of an attainable claim, the unique price of $H$, in the absence of arbitrage opportunities, is $H_0 = V_0(\phi)$, where $\phi$ is the replicating strategy. Any other price for $H$ would lead to an arbitrage opportunity, as the following argument shows:

- Suppose that the price of the claim is $H_0 + \epsilon$, for some $\epsilon > 0$.

- Sell the claim $H$ at time 0 and receive $H_0 + \epsilon$.

- Use $H_0 = V_0(\phi)$ to invest in the replicating strategy $\phi$ (which is self-financing), and follow this strategy until time $T$.

- At time $T$, we have $V_N(\phi) = H$ from trading, which covers the liability generated from selling the claim at time 0.

- Thus a risk-free profit of $\epsilon$ has been generated.

\(^2\)Mathematically, an arbitrage opportunity can be defined as a self-financing trading strategy, $\phi$, with $V_0(\phi) = 0$, $V_N(\phi) \geq 0$ with probability 1 and $V_N(\phi) > 0$ with positive probability.
• Similarly, if the price is assumed to be $H_0 - \epsilon$, a risk-free profit of $\epsilon$ can be generated by taking the opposite positions to those above.

The assumption of no-arbitrage is closely connected with the existence of a measure, $Q$, which is equivalent\(^3\) to $P$ and under which the discounted asset price is a martingale. Symbolically, this means that,

$$E^Q[X_n|\mathcal{F}_{n-1}] = X_{n-1},$$

for $n = 1, 2, \ldots, N$. Such a measure, $Q$, is called an equivalent martingale measure for the asset $S$, often simply abbreviated to martingale measure. Dalang, Morton and Willinger prove, in Theorem 3.3 of their paper [18], that there are no arbitrage opportunities in this type of market if and only if there exists a martingale measure for $S$. Furthermore, Taqqu and Willinger ([78]) show that the market is complete if and only if a unique martingale measure for $S$ exists.

This has an important consequence. Recall that the arbitrage-free price of an attainable claim, $H$, is given by $H_0 = V_0(\phi)$, where $\phi$ is the replicating strategy which satisfies,

$$\frac{H}{B_N} = V_0(\phi) + \sum_{k=1}^{N} \xi_k \Delta X_k.$$

As the process $X$ is a $(Q, \mathbb{F})$-martingale, the price of the claim at time 0 is,

$$H_0 = E^Q \left[ \frac{H}{B_N} \right],$$

and, more generally, the value at time $t_n$ of the claim is,

$$H_n = B_n E^Q \left[ \frac{H}{B_N} \bigg| \mathcal{F}_n \right],$$

or,

$$H_n = B_n E^Q \left[ \frac{H_k}{B_k} \bigg| \mathcal{F}_n \right],$$

\(^3\)Recall that two measures $P_1$ and $P_2$ are equivalent on $(\Omega, \mathcal{F})$ if $P_1(A) = 0$ iff $P_2(A) = 0$ for all $A \in \mathcal{F}$. 

for any \( k > n \). Hence, in complete markets, current values of attainable claims can be computed by taking expectations under the unique martingale measure for the risky asset.

In incomplete markets, pricing is not this straightforward, as there is no unique martingale measure. However, it can be shown, using no-arbitrage arguments similar to the one given above, that all arbitrage-free prices of a claim, \( H \), lie in the interval,

\[
\left[ \inf_Q E_Q^Q \left( \frac{H}{B_N} \right), \sup_Q E_Q^Q \left( \frac{H}{B_N} \right) \right],
\]

where \( Q \) runs over all possible martingale measures for \( S \). It is natural, therefore, to price contingent claims in incomplete markets by choosing a particular martingale measure in some economically reasonable way, and then proceeding as for the complete market case. Methods for choosing a martingale measure include mean-variance hedging ([71]), utility maximisation ([19]), local risk minimisation ([31]), entropy minimisation ([42]), Esscher transforms ([34]), and the method of Elliott/Madan ([28]). The specifics for some of these techniques will be discussed in later chapters.

This completes our introduction to the financial concepts required to understand this thesis. We now give a brief history of option pricing, and consider the motivation for proposing the Switching Black–Scholes model, which is the focus of this work.

### 1.2 History and Motivation

Option pricing was first studied in 1900 by the French mathematician Louis Bachelier. In his doctoral dissertation, *Théorie de la Spéculation* ([4]), he proposed that asset prices evolve as Brownian motion, and found the value of options based on this assumption. Despite the obvious fault of this model that
both asset prices and option prices can assume negative values\(^4\), little progress was made in this area until 1965, when Paul Samuelson ([67]) proposed that asset prices evolve according to geometric Brownian motion. This means that,

\[
dS(t) = \mu S(t) \, dt + \sigma S(t) \, dW(t), \tag{1.1}
\]

where \( S \) is the asset price, \( \mu \) is the expected rate of return on the asset, \( \sigma \) is the volatility of the asset price, and \( W \) is Brownian motion. Solving Equation (1.1) gives,

\[
S(t) = S(0) \exp \left\{ (\mu - \frac{1}{2} \sigma^2) t + \sigma W(t) \right\},
\]

from which it is clear that the asset price is always positive – an improvement on Bachelier’s model.

In 1973, Fischer Black, Myron Scholes and Robert Merton made a breakthrough in the valuation of options by developing a method of pricing based on this model (see [8], [60]). Using the idea that a call option can be perfectly replicated by the underlying asset and a bond, they found that the current price of a call option expiring at time \( T \) with strike price \( K \) is,

\[
C(0) = S(0) N(d_1) - Ke^{-rT} N(d_2), \tag{1.2}
\]

where,

\[
d_1 = \frac{\log(S(0)/K) + (r + \frac{1}{2} \sigma^2) T}{\sigma \sqrt{T}},
\]

\[
d_2 = d_1 - \sigma \sqrt{T}, \quad r \text{ is the risk-free, continuously compounding interest rate per annum, and } N \text{ is the normal distribution function, given by,}
\]

\[
N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}z^2} \, dz.
\]

The model (1.1) and pricing formula (1.2), known as the Black-Scholes model and pricing formula, are still widely used today, although its deficiencies have become a standard topic of research. To describe one major deficiency of the

\(^4\)Although Bachelier’s model is no longer used to model asset prices, it is sometimes used to model the spread (that is, the difference) between two asset prices, as this quantity can be both positive and negative.
Black-Scholes model, note from Equation (1.2) that the Black-Scholes call price is a function of five parameters: the initial asset price, $S(0)$, the strike price, $K$, the time to expiry, $T$, the interest rate, $r$, and the volatility, $\sigma$. Of these parameters, only the volatility of the underlying asset is not known with certainty. Thus, we can set the market price of a call option equal to the Black-Scholes price, and solve for the volatility parameter. A volatility obtained in this way is called the Black-Scholes implied volatility of the option.

If this process is repeated for varying strike prices of call options on the same asset, an implied volatility curve, or implied volatility smile, is observed (see, for example [9], [66]), which typically has a parabolic-like shape with a minimum when the call option is at-the-money (that is, the strike price and current asset price are identical). This is contrary to the assumption inherent in the Black-Scholes model of constant volatility.

Historically, various modifications to the Black-Scholes model have been proposed in order to eliminate this inconsistency with the market. For an excellent overview and comparison of these, see Bakshi, Cao and Chen ([5]). The asset price model,

$$dS(t) = a(t, S(t)) \, dt + b(t, S(t)) \, dW(t) + \lambda(t)S(t) \, dN(t), \tag{1.3}$$

where $a$, $b$ and $\lambda$ are given functions and $N$ is a standard Poisson process, incorporates many of these modifications. The case when $\lambda$ is zero and $b$ satisfies a second stochastic differential equation is known as a stochastic volatility model. This type of model has been studied by many authors, including Hull and White ([47]), Wiggins ([79]), Scott ([73]), Stein and Stein ([77]), and Heston ([46]).

When $\lambda$ is non-zero, the asset price is no longer continuous, and the model is called a jump diffusion model. Merton ([61]) and Bates ([6]) are among those who have studied option valuation using jump models.

Thirdly, there are asset price models incorporating both of these features, which are known as stochastic volatility, jump diffusion models. See Bates ([7]) and Scott ([74]) for examples of work in this area.
An alternative to the above approaches is to allow the drift and volatility parameters of the Black-Scholes model to vary according to a regime switching process. This method is much simpler to formulate and is more tractable than stochastic volatility and jump diffusion models, but still replicates the implied volatility smiles observed in the market, and produces option prices which are closer to market values than those obtained from the traditional Black-Scholes formula (see, for example, [2], [45], [68]). The Switching Black-Scholes model proposed in this thesis is a type of regime switching model in which the drift and volatility are assumed to switch values according to the state of a hidden Markov chain. In the next section, we examine the various types of regime switching models which have previously been studied in the literature, and the methods of option pricing employed.

1.3 Regime Switching Literature Overview

The concept of regime switching has been around for many years, with models gradually progressing from those in which the switching dates are known to the very general models now proposed. At first, the switches were assumed to be independent, but more recently, in the seventies, Lindgren ([58]) and Goldfeld and Quandt ([36]) introduced the notion of the switches being governed by a Markov chain, giving rise to the terms “Hidden Markov Model” and “Markov Switching Model”. It was not until 1989, when Hamilton’s paper, “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle” ([44]), appeared in *Econometrica* that regime switching models were first applied to finance.

The subject has subsequently attracted much interest, with applications to modelling the US short-term interest rate (this is a classic example – one reference is Hamilton [43]), exchange rates and stock prices. Diebold *et al* ([21]) and Gray ([39]) further advanced the field by allowing the transition probabilities of the Markov chain to vary with time - a concept not previously
considered.

Much of the research to date has been geared towards estimating parameters of regime switching models, while relatively little progress has been made regarding option pricing in such a setting. Some papers of importance in this area are those by Bollen ([10]), Buffington and Elliott ([12]), Campbell and Li ([13]), Di Masi et al. ([22]), Duan et al. ([23]), Duecker ([24]), Guo ([40]), Kim ([56]), Naik ([63]), and Yao et al. ([82]). Besides Duecker, Kim and Duan et al., who deal with Markov switching when the asset price follows a GARCH or ARIMA type model, these authors all consider a Black-Scholes asset price model in which the drift and/or volatility parameters can switch values according to a finite-state hidden Markov chain. However, the specific approach of each varies.

All except Naik assume that the drift, as well as the volatility, of the asset can switch. Naik conjectures that the drift is constant, and that at the volatility switching times, the asset can also randomly jump in price. With this model, and some additional assumptions which ensure a unique martingale measure exists, he obtains closed-form option pricing formulae, and analyses some numerical examples.

Guo has a similar approach, in that she incorporates some extra assets, related to the switching times, into the market in order to obtain a unique option pricing formula. She then discretises the asset price so as to numerically analyse the method, and extends the work to price Russian options in a further paper ([41]).

As a switching process incorporates a second source of uncertainty into the Black-Scholes model, a Black-Scholes market with regime switching is incomplete and no unique option pricing methodology exists. Naik and Guo deal with this incompleteness by ‘completing’ the market in some way; that is, by adding additional assets or assumptions into the market so as to force completeness. The papers by Bollen and by Campbell and Li take an even simpler approach. They assume that “regime risk” is not priced in the market, so the market is complete and option pricing is straightforward.
Campbell and Li price options directly under this assumption by noting that for pricing purposes, the return on the risky asset in each time period equals the riskless rate. Bollen also uses this observation, but in a discrete-time setting. He constructs a recombining lattice of asset price paths, and computes the option price as the discounted expected payoff at expiry. The final outcome is an option price which is dependent on the current regime of the market.

Only the papers by Di Masi et al and Yao et al use the traditional approach to pricing in incomplete markets by choosing a martingale measure in some economically reasonable way. Di Masi et al use the local risk minimisation method of Follmer and Schweizer ([31]), while Yao et al use a Girsanov type theorem to define a martingale measure under which to perform the pricing.

This thesis considers various other methods for choosing a martingale measure when the risky asset follows a Black-Scholes model incorporating regime switching.

1.4 Outline of the Thesis

The aim of this thesis is to study a “Switching Black-Scholes” model of a price process. This is a regime switching model in which the Black-Scholes drift and volatility parameters are permitted to vary between a finite number of possible values at known times, according to the state of a hidden Markov chain. The Markov chain is assumed to evolve in discrete-time, contrary to the assumption of continuous-time made for many regime switching models previously studied in the literature.

The thesis begins in Chapter 2 with a detailed mathematical description of the Switching Black–Scholes model. Also contained in this preliminary chapter are the details of some density functions and expectations associated with the model which are needed in later chapters.

Having defined the Switching Black–Scholes model, the bulk of the thesis is concerned with pricing options on an asset which evolves in this way. Due to the
additional source of randomness caused by incorporating parameter switching into the Black-Scholes model, the Switching Black-Scholes market is incomplete. Thus to price options, one martingale measure from the many that exist must be chosen. Chapters 3, 5 and 6 each focus on a different method for making this choice. Chapter 3 examines the mean-variance hedging method of pricing; Chapter 5, the Esscher method of pricing; and Chapter 6, the method of minimum entropy. Details of the exact procedures of these methods are given in the respective chapters.

The Esscher and minimum entropy prices computed in this thesis are found using a forward inductive procedure, whereas the mean-variance price given in Chapter 3 is computed using a backward inductive argument. A second chapter devoted to mean-variance hedging (Chapter 4) calculates this price by applying the forward inductive procedure used for the other two methods.

Chapter 7 focuses on some numerical studies of the option pricing formulae developed in Chapters 3, 5 and 6. Using C programs which compute the option prices, some concrete examples are given, and the effect of various parameters on the option price are examined. The examples also enable a quantitative comparison of the option pricing methodologies, which suggests some topics suitable for future research, and provides verification that the Switching Black-Scholes model does indeed produce implied volatility smiles.

Chapter 8 has a slightly different flavour, and contains work conducted jointly with Professor R.J. Elliott (University of Calgary, Canada) and Dr W.P. Malcolm (DSTO Edinburgh and University of Adelaide). The chapter addresses the question of how to estimate the parameters of the Switching Black-Scholes model. This involves constructing a semimartingale representation for the continuous-time version of the hidden Markov chain, and then computing filtered and smoothed estimates of the parameters and Markov chain state using reference probability methods. For the purposes of this chapter, the drift parameter of the Switching Black-Scholes model is assumed to be constant.
A conclusion to the thesis, including a discussion on further research possibilities, is given in Chapter 9.
Chapter 2

The Switching Black–Scholes Model

This chapter gives a detailed mathematical description of the Switching Black–Scholes model, which is the focus of this thesis, and considers some properties of the model which will be needed in later chapters. As explained in the Introduction, the Switching Black–Scholes model is based on the standard geometric Brownian motion model of a price process, $S$:

$$dS(t) = \mu S(t) \, dt + \sigma S(t) \, dW(t).$$

However, the drift and volatility parameters, $\mu$ and $\sigma$, are permitted to vary among a finite number of possible values at known times, according to the state of a hidden Markov process.

The chapter begins in Section 2.1 with a description of the market model which is used throughout the thesis. The market consists of a risky asset which evolves according to the Switching Black–Scholes model, and a riskless asset, or bond, which is used as a numeraire. Section 2.1.1 defines the riskless asset, Section 2.1.2 introduces the Markov chain used in the Switching Black–Scholes model, and Section 2.1.3 mathematically and intuitively describes the Switching Black–Scholes model.

Having specified the model, in Section 2.2 we define and compute some
associated density functions, both under the real world probability measure and under a martingale measure. Of particular importance is the relationship between the joint conditional density functions of \((S_n, Z_n)\) given \((S_{n-1}, Z_{n-1})\) and of \((S_n, Z_n)\) given \((S_0, Z_0)\), where \(S\) is the risky asset price and \(Z\) is the hidden Markov chain. Theorem 2.2.5 finds this relationship, which is analogous to the Jamshidian forward induction formula used to compute state prices in a binomial asset pricing model (see, for example, [50]). This relationship will be used in Chapters 4, 5 and 6 to calculate contingent claim prices.

Finally, in Section 2.3 we evaluate some expectations associated with the model which will be useful in later chapters.

In the following, all processes are defined on a complete probability space \((\Omega, \mathcal{F}, P)\), where \(P\) represents the real world probability measure, and \(W\) is used to denote Brownian motion on this space. This notation, and the notation introduced in Section 2.1, is used throughout the thesis.

### 2.1 Description of the Market

The market we consider consists of one risky asset and one riskless asset (or bond) which can be traded at discrete times \(t_n = n\tau\) for \(n = 0, 1, \ldots, N\), where \(\tau > 0\) and \(N \in \mathbb{N}\) are fixed. We assume that the market is “frictionless”, meaning that there are no taxes, transaction costs or restrictions on borrowing in the market, and short sales are permitted. The final time is \(T = N\tau\).

Let \(S(t)\) and \(B(t)\) denote the values of the risky asset and bond, respectively, at a given time \(t\), and define,

\[
S_n = S(t_n), \quad B_n = B(t_n),
\]

for \(n = 0, 1, \ldots, N\). The discounted asset price is defined as,

\[
X_n = \frac{S_n}{B_n},
\]

and we write the change in the discounted price as,

\[
\Delta X_n = X_n - X_{n-1}.
\]
2.1.1 The Riskless Asset

The bond is assumed to be deterministic with,

\[ B(t) = e^{rt}, \]

or \( B_n = e^{rn\tau} \), which corresponds to continuously compounding interest at a rate of \( r\% \) per annum. However, all results of this thesis can easily be modified to use other types of interest, such as simple interest \( (B_n = 1 + r n \tau) \) or interest compounding in each period \( (B_n = (1 + \frac{r}{N})^n). \)

2.1.2 The Hidden Markov Chain

To describe the evolution of the risky asset, we first need to define a Markov chain, whose state will determine the Switching Black–Scholes model parameters at each time. We take \( (Z_n)_{n=0,1,...,N} \) to be a stationary Markov chain with state space, \( \mathcal{H} \), of dimension \( M < \infty \). Without loss of generality, the state space is chosen to be,

\[ \mathcal{H} = \{ e_1, e_2, \ldots, e_M \}, \]

where \( e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T \) is the \( i \)th element of the standard basis for \( \mathbb{R}^M \). Occasionally, we identify the state \( e_j \) using only the integer \( j \).

The transition probabilities of this Markov chain are denoted by,

\[ A_{ij} = P(Z_n = e_i | Z_{n-1} = e_j), \]

for \( i, j \in \{1, 2, \ldots, M\} \) and \( n \in \{1, 2, \ldots, N\} \). Setting \( A = (A_{ij}) \), we can then write,

\[ Z_n = AZ_{n-1} + I_n, \]

where the process \( L = (I_n)_{n=1,2,\ldots,N} \) is a martingale increment with respect to the filtration generated by \( Z \).

We convert this Markov chain into a continuous-time process on \([0, T]\) by setting \( Z(t_n) = Z_n \) for \( n = 0, 1, \ldots, N \), and taking \( Z \) to be constant on each interval \([t_n, t_{n+1})\). The process \( Z \) is also known as the state process for the
Switching Black–Scholes market, and is used to describe the parameter switching in the risky asset price model.

2.1.3 The Risky Asset

The risky asset is assumed to evolve according to the “Switching Black–Scholes” model. This is modelled mathematically by the stochastic differential equation,

\[
\frac{dS(t)}{S(t)} = \mu(Z(t))dt + \sigma(Z(t))dW(t), \tag{2.1}
\]

where \(Z\) is the hidden Markov chain described in Section 2.1.2 and \(\mu\) and \(\sigma\) are functions \(\mathbb{R}^M \rightarrow (0, 1)\). It is clear from this equation that for a fixed value of \(Z(t)\), the risky asset evolves according to the ordinary Black–Scholes model.

By taking \(Z\) to be a Markov chain, we simulate a market in which the risky asset price follows geometric Brownian motion for the periods \([t_n, t_{n+1})\), but at the discrete times \(t_n, n = 0, 1, \ldots, N\), the drift and volatility parameters of the model can “switch”. In this way, the process \(Z\) can be regarded as an underlying variable which chooses one of \(M\) possible market states, and \(Z(t)\) as the market state at time \(t\).

To write Equation (2.1) in an alternative form, let \(\mu = (\mu_1, \mu_2, \ldots, \mu_M)^T\) and \(\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_M)^T\) contain the values among which the drift and volatility can switch. Here, \(\mu_j = \mu(e_j)\) and \(\sigma_j = \sigma(e_j)\) are the drift and volatility of the risky asset when the market (or hidden Markov chain) is in state \(e_j\). As the state space of the Markov chain was chosen to be the set of unit vectors in \(\mathbb{R}^M\), we can write,

\[
\frac{dS(t)}{S(t)} = \langle \mu, Z(t) \rangle dt + \langle \sigma, Z(t) \rangle dW(t), \tag{2.2}
\]

where \(\langle \cdot, \cdot \rangle\) denotes the inner product in \(\mathbb{R}^M\).

Using Itô calculus, it can be shown that for \(n = 1, 2, \ldots, N\), the solution of Equation (2.1) is,

\[
S_n = S_{n-1} \exp \left\{ \left( \mu(Z_n) - \frac{1}{2} \sigma(Z_n)^2 \right) \tau + \sigma(Z_n) \Delta W_n \right\}, \tag{2.3}
\]
where $\Delta W_n = W(t_n) - W(t_{n-1})$. By applying properties of Brownian motion such as,
\[
\frac{\Delta W_n}{\sqrt{\tau}} \sim N(0, 1) \text{ for all } n,
\]
this equation is used to analyse the behaviour of an asset evolving according to the Switching Black–Scholes model.

### 2.1.4 Information in the Market

At time $t_n$, the information known about the market can be represented by the $\sigma$-algebra,
\[
\mathcal{F}_n = \sigma\{S_k, Z_k : k = 0, 1, \ldots, n\}.
\]
The filtration $(\mathcal{F}_n)_{n=0,1,\ldots,N}$ of these $\sigma$-algebras is used constantly throughout the thesis.

As $Z$ is a Markov chain and the value of $S_n$ only depends on the values of $S_{n-1}$, $Z_n$ and $\Delta W_n$ (from Equation (2.3)), it is clear that the joint process $\{(S_n, Z_n)\}_{n=1,2,\ldots,N}$ is Markov with respect to the filtration $(\mathcal{F}_n)$. This means that,
\[
E[g(S_{n+1}, Z_{n+1}) | \mathcal{F}_n] = E[g(S_{n+1}, Z_{n+1}) | S_n, Z_n],
\]
for $n = 0, 1, \ldots, N - 1$ and for any function $g$ such that the expectation exists. Denoting the expectation in (2.4) by $h$, we can write $h = h(S_n, Z_n)$. This is an important property of the Switching Black–Scholes model which will be used repeatedly.

In particular, the joint conditional density function of $(S_n, Z_n)$ given $\mathcal{F}_{n-1}$ depends only on the values of $S_{n-1}$ and $Z_{n-1}$. We use this property in the next section.
2.2 Density Functions Associated with the Switching Black–Scholes Model

Throughout the thesis, we will encounter expectations of the form,

\[ E[f(S_n, Z_n) | \mathcal{F}_{n-1}], \]

for various functions \( f \). In order to evaluate these expectations, we now calculate the joint conditional density function of \((S_n, Z_n)\) given \( S_{n-1} = s \) and \( Z_{n-1} = e_i \) under the real world probability measure.

**Theorem 2.2.1.** Let \( n \in \{1, 2, \ldots, N\} \). Under the real world probability measure, \( P \), the joint conditional density function \( f_{s,e_i}(x, e_j) : \mathbb{R} \times \mathcal{H} \to \mathbb{R} \) of \((S_n, Z_n)\) given \( S_{n-1} = s \) and \( Z_{n-1} = e_i \) is,

\[
f_{s,e_i}(x, e_j) = \begin{cases} \frac{A_{ji}}{2\pi \sigma_j \sqrt{x}} \exp \left[ -\frac{1}{2} \left( \frac{\log(x/s) - (\mu_j - \frac{1}{2}\sigma_j^2)\tau}{\sigma_j \sqrt{\tau}} \right)^2 \right], & x > 0, \\
0, & x \leq 0. \end{cases}
\]

Here, \( S \) denotes the asset price and \( Z \) the Markov chain in the Switching Black–Scholes model. Note that this density function is independent of \( n \).

**Proof.** For any bounded Borel measurable function, \( g \), we have,

\[
E[g(S_n, Z_n) | S_{n-1} = s, Z_{n-1} = e_i] = \sum_{j=1}^{M} E[g(S_n, Z_n) | S_{n-1} = s, Z_{n-1} = e_i, Z_n = e_j] \times P(Z_n = e_j | Z_{n-1} = e_i).
\]

Using the expression for \( S_n \) given in Equation (2.3) and the transition probability definitions, this becomes,

\[
\sum_{j=1}^{M} A_{ji} E \left[ g \left( s e^{(\mu_j - \frac{1}{2}\sigma_j^2)\tau + \sigma_j \Delta W_n}, e_j \right) \right].
\]

As \( \Delta W_n/\sqrt{\tau} \) is normally distributed with mean 0 and variance 1, the expectation equals,

\[
\sum_{j=1}^{M} A_{ji} \int_{-\infty}^{\infty} g \left( s e^{(\mu_j - \frac{1}{2}\sigma_j^2)\tau + \sigma_j \sqrt{\tau} z}, e_j \right) \frac{e^{-z^2}}{\sqrt{2\pi}} dz,
\]
and changing variables to $x = se^{(\mu_j - \frac{1}{2}\sigma_j^2)\tau + \sigma_j\sqrt{\tau}z}$ gives,

$$\sum_{j=1}^{M} \int_{0}^{\infty} g(x, e_j) \frac{A_{ji}}{\sqrt{2\pi}\sigma_j\sqrt{\tau}} x \exp \left[ -\frac{1}{2} \left( \frac{\log(x/s) - (\mu_j - \frac{1}{2}\sigma_j^2)\tau}{\sigma_j\sqrt{\tau}} \right)^2 \right] dx.$$

Choosing $g(S_n, Z_n) = I(S_n \leq s', Z_n = e_j)$ proves the result. \qed

We also consider the individual density functions of $S_n$, and $Z_n$, given $S_{n-1} = s$ and $Z_{n-1} = e_i$ under the real world probability measure.

**Theorem 2.2.2.** Let $n \in \{1, 2, \ldots, N\}$. Under the real world probability measure, $P$, the conditional density function $g_{s,e_i} : \mathbb{R} \rightarrow \mathbb{R}$ of $S_n$ given $S_{n-1} = s$ and $Z_{n-1} = e_i$ is,

$$g_{s,e_i}(x) = \begin{cases} \sum_{j=1}^{M} \frac{A_{ji}}{\sqrt{2\pi}\sigma_j\sqrt{\tau}} x \exp \left[ -\frac{1}{2} \left( \frac{\log(x/s) - (\mu_j - \frac{1}{2}\sigma_j^2)\tau}{\sigma_j\sqrt{\tau}} \right)^2 \right], & x > 0, \\ 0, & x \leq 0. \end{cases}$$

Here, $S$ denotes the asset price and $Z$ the Markov chain in the Switching Black–Scholes model.

**Proof.** This is clear from the previous theorem and proof. \qed

**Theorem 2.2.3.** Let $n \in \{1, 2, \ldots, N\}$. Under the real world probability measure, $P$, the conditional density function $F_{s,e_i} : \mathcal{H} \rightarrow \mathbb{R}$ of $Z_n$ given $S_{n-1} = s$ and $Z_{n-1} = e_i$ is given by,

$$F_{s,e_i}(e_j) = A_{ji}$$

Here, $S$ denotes the asset price and $Z$ the Markov chain in the Switching Black–Scholes model.

**Proof.** Again, this follows from the result and proof of Theorem 2.2.1. \qed

In Chapter 1, we noted that the arbitrage-free price at time $m$ of an $\mathcal{F}_n$-measurable contingent claim, $H_n$, is,

$$H_m = E^Q [ H_n e^{-r(n-m)} | \mathcal{F}_m ],$$

for a martingale measure, $Q$. Consequently, we make the following definitions.
Definition 2.2.4. Let $n \in \{1, 2, \ldots, N\}$. Under a fixed martingale measure, $Q$, let $\phi_n$ and $\psi_n$ denote the joint conditional density functions of $(S_n, Z_n)$ given $\mathcal{F}_0$ and $\mathcal{F}_{n-1}$, respectively. As $\{(S_n, Z_n)\}_{n=1,2,\ldots,N}$ is Markov with respect to the filtration $(\mathcal{F}_n)$, for any bounded Borel measurable function, $g$, we have,

$$E^Q[g(S_n, Z_n) | S_0 = s, Z_0 = e_i] = \sum_{j=1}^{M} \int_{-\infty}^{\infty} \phi_n(s, e_i, x, e_j) g(x, e_j) \, dx,$$

and

$$E^Q[g(S_n, Z_n) | S_{n-1} = s, Z_{n-1} = e_i] = \sum_{j=1}^{M} \int_{-\infty}^{\infty} \psi_n(s, e_i, x, e_j) g(x, e_j) \, dx,$$

where $e_i \in \mathcal{H}$ and $s \in \mathbb{R}$.

In the next theorem, we show how $\phi_n$ can be calculated recursively from $\psi_n$ and $\phi_{n-1}$. This is a Jamshidian style forward induction formula, and allows us to compute the joint conditional density function of $(S_n, Z_n)$ given $\mathcal{F}_0$ under the martingale measure $Q$, for any value of $n \in \{1, 2, \ldots, N\}$, provided that we know $\phi_1$, the conditional density function of $(S_1, Z_1)$ given $\mathcal{F}_0$. Using these density functions, the price at time 0 of an $\mathcal{F}_n$-measurable contingent claim, $H_n$, is,

$$H_0(S_0, Z_0) = e^{-rn} E^Q[H_n | \mathcal{F}_0] = e^{-rn} \sum_{j=1}^{M} \int_{-\infty}^{\infty} \phi_n(S_0, Z_0, x, e_j) H_n(x, e_j) \, dx.$$

The density function $\phi_1$ can readily be computed using Bayes rule and the joint conditional density function of $(S_1, Z_1)$ given $\mathcal{F}_0$ under the real world probability measure (see Theorem 2.2.1). This method for calculating contingent claim prices is used in Chapters 4, 5 and 6 for various martingale measures.

We now state the theorem.

Theorem 2.2.5. Let $e_i, e_j \in \mathcal{H}$ and $x, s \in \mathbb{R}$. The sequence of density functions $(\phi_n)_{n=1,2,\ldots,N}$ defined in 2.2.4 can be calculated recursively using the equations,

$$\phi_1(s, e_i, x, e_j) = \psi_1(s, e_i, x, e_j),$$
and,
\[ \phi_n(s, e_i, x, e_j) = \sum_{l=1}^{M} \int_{-\infty}^{\infty} \psi_n(z, e_l, x, e_j) \phi_{n-1}(s, e_i, z, e_l) \, dz. \]

Proof. The first assertion follows immediately from Definition 2.2.4. To prove the second assertion, suppose that,
\[ H_0 = \mathbb{E}^{\mathbb{Q}}[H_n | \mathcal{F}_0], \]
for some \( \mathcal{F}_n \)-measurable random variable, \( H_n \). Using the definition of the joint conditional density function, \( \phi_n \), this can be written as,
\[ H_0(s, e_i) = \sum_{j=1}^{M} \int_{-\infty}^{\infty} \phi_n(s, e_i, x, e_j) H_n(x, e_j) \, dx. \] (2.5)

Alternatively, we could condition first on the \( \sigma \)-algebra \( \mathcal{F}_{n-1} \), so that,
\[ H_0 = \mathbb{E}^{\mathbb{Q}}[H_{n-1} | \mathcal{F}_0], \] (2.6)
where \( H_{n-1} = \mathbb{E}^{\mathbb{Q}}[H_n | \mathcal{F}_{n-1}] \). Using the joint conditional density function, \( \psi_n \), of \( (S_n, Z_n) \) given \( \mathcal{F}_{n-1} \) we have,
\[ H_{n-1}(z, e_l) = \sum_{j=1}^{M} \int_{-\infty}^{\infty} \psi_n(z, e_l, x, e_j) H_n(x, e_j) \, dx. \]

Substituting this into Equation (2.6) and applying the joint conditional density function, \( \phi_{n-1} \), of \( (S_{n-1}, Z_{n-1}) \) given \( \mathcal{F}_0 \), gives,
\[ H_0(s, e_i) = \sum_{l=1}^{M} \int_{-\infty}^{\infty} \phi_{n-1}(s, e_i, z, e_l) \sum_{j=1}^{M} \int_{-\infty}^{\infty} \psi_n(z, e_l, x, e_j) H_n(x, e_j) \, dx \, dz, \]
and changing the order of integration and summation, we have,
\[ H_0(s, e_i) = \sum_{j=1}^{M} \int_{-\infty}^{\infty} \sum_{l=1}^{M} \int_{-\infty}^{\infty} \phi_{n-1}(s, e_i, z, e_l) \psi_n(z, e_l, x, e_j) \, dz \, H_n(x, e_j) \, dx. \]

As all these equations hold for any \( \mathcal{F}_n \)-measurable random variable, \( H_n \), comparing the above equation with (2.5) gives the result:
\[ \phi_n(s, e_i, x, e_j) = \sum_{l=1}^{M} \int_{-\infty}^{\infty} \psi_n(z, e_l, x, e_j) \phi_{n-1}(s, e_i, z, e_l) \, dz. \]
2.3 Expectations Associated with the Model

In this section, we apply the density functions found in Section 2.2 to evaluate some expectations which will be required in later chapters. These results also make use of two lemmas given in Appendix 1, which calculate “normal density” type integrals.

The first result calculates the expectations of $S_N$, $S_N^2$, $\Delta X_N$ and $\Delta X_N^2$ given $S_{N-1} = s$ and $Z_{N-1} = e_i$.

**Lemma 2.3.1.** Let $(S_n)_{n=0,1,\ldots,N}$ denote an asset following the Switching Black–Scholes model, and let $X_n = e^{-rn} S_n$ denote the discounted value of this asset. Then we can calculate the expectations of $S_N$, $S_N^2$, $\Delta X_N$ and $\Delta X_N^2$ given $S_{N-1} = s$ and $Z_{N-1} = e_i$ can be calculated using the following formulae.

\[
E[S_N | S_{N-1} = s, Z_{N-1} = e_i] = \sum_{j=1}^{M} A_{ji} s e^{\mu_j \tau}; \quad (2.7)
\]

\[
E[S_N^2 | S_{N-1} = s, Z_{N-1} = e_i] = \sum_{j=1}^{M} A_{ji} s^2 e^{(2\mu_j + \sigma_j^2) \tau}; \quad (2.8)
\]

\[
E[\Delta X_N | S_{N-1} = s, Z_{N-1} = e_i] = \frac{s}{e^{\tau T}} \sum_{j=1}^{M} A_{ji} (e^{\mu_j \tau} - e^{r \tau}); \quad (2.9)
\]

\[
E[\Delta X_N^2 | S_{N-1} = s, Z_{N-1} = e_i] = \frac{s^2}{e^{2\tau T}} \sum_{j=1}^{M} A_{ji} \left( e^{(2\mu_j + \sigma_j^2) \tau} - 2e^{(r + \mu_j) \tau} + e^{2r \tau} \right). \quad (2.10)
\]

**Proof.** Using the conditional density function of $S_n$ given $S_{n-1} = s$, $Z_{n-1} = e_i$, determined in Theorem 2.2.2, we can write the expectation in Equation (2.7) as,

\[
E[S_N | S_{N-1} = s, Z_{N-1} = e_i] = \int_{-\infty}^{\infty} x g_{s,e_i}(x) \, dx
\]

\[
= \int_{0}^{\infty} x \sum_{j=1}^{M} A_{ji} \frac{1}{\sqrt{2\pi} \sigma_j \sqrt{T} x} \exp \left[ -\frac{1}{2} \left( \frac{\log(x/s) - (\mu_j - \frac{1}{2} \sigma_j^2) \tau}{\sigma_j \sqrt{T}} \right)^2 \right] \, dx.
\]
Changing variables to \( z = \frac{\log(x/s) - (\mu_j - \frac{1}{2} \sigma_j^2) \tau}{\sigma_j \sqrt{\tau}} \), we obtain,

\[
E[S_N | S_{N-1} = s, Z_{N-1} = e_i] = \sum_{j=1}^{M} A_{ji} \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} s e^{(\mu_j - \frac{1}{2} \sigma_j^2) \tau + \sigma_j \sqrt{\tau} z} e^{-\frac{1}{2} z^2} \, dz.
\]

This integral can be computed from Lemma A.1.1 of Appendix 1, which states that for a constant \( c \),

\[
\int_{-\infty}^{\infty} e^{cz - \frac{1}{2} z^2} \, dz = \sqrt{2\pi} e^{\frac{1}{2} c^2}.
\]

Setting \( c = \sigma_j \sqrt{\tau} \) gives Equation (2.7).

To prove Equation (2.8), we follow the same procedure to obtain,

\[
E [ S_N^2 | S_{N-1} = s, Z_{N-1} = e_i ]
= \int_{0}^{\infty} x^2 \sum_{j=1}^{M} A_{ji} \sqrt{\frac{2}{\pi}} \frac{1}{\sigma_j \sqrt{\tau}} \exp \left[ -\frac{1}{2} \left( \frac{\log(x/s) - (\mu_j - \frac{1}{2} \sigma_j^2) \tau}{\sigma_j \sqrt{\tau}} \right)^2 \right] \, dx
= \sum_{j=1}^{M} A_{ji} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\sigma_j^2}{\sigma_j \sqrt{\tau}} \left( x^2 e^{2(\mu_j - \frac{1}{2} \sigma_j^2) \tau + 2 \sigma_j \sqrt{\tau} z} e^{-\frac{1}{2} z^2} \right) \, dz.
\]

Again, application of Lemma A.1.1 gives the result.

For Equations (2.9) and (2.10), write,

\[
\Delta X_N = \frac{S_N}{e^{r_N \tau}} - \frac{S_{N-1}}{e^{r(N-1) \tau}} = \frac{1}{e^{r \tau}} (S_N - S_{N-1} e^{r \tau}).
\]

Then,

\[
E[\Delta X_N | S_{N-1} = s, Z_{N-1} = e_i] = \frac{1}{e^{r \tau}} \left( E[S_N | S_{N-1} = s, Z_{N-1} = e_i] - s e^{r \tau} \right).
\]

The identity \( \sum_{j=1}^{M} A_{ji} = 1 \) allows us to write,

\[
s e^{r \tau} = s \sum_{j=1}^{M} A_{ji} e^{r \tau}.
\]

Substituting Equations (2.13) and (2.7) into (2.12) proves Equation (2.9).

For the final equation, we also use Equation (2.11) to obtain,

\[
E[\Delta X_N^2 | S_{N-1} = s, Z_{N-1} = e_i] = \frac{1}{e^{2r \tau}} \left( E[S_N^2 | S_{N-1} = s, Z_{N-1} = e_i] - 2 s e^{r \tau} E[S_N | S_{N-1} = s, Z_{N-1} = e_i] + s^2 e^{2r \tau} \right).
\]
Substituting in Equations (2.7) and (2.8) gives,
\[
\frac{1}{e^{2rt}} \left( s^2 \sum_{j=1}^{M} A_{ji} e^{(2\mu_j + \sigma_j^2)\tau} - 2s^2 e^{rt} \sum_{j=1}^{M} A_{ji} e^{\mu_j \tau} + s^2 e^{2rt} \right).
\]

Using the identity \( \sum_{j=1}^{M} A_{ji} = 1 \) completes the proof. \( \square \)

The next lemma calculates a more general expectation, involving the normal distribution function.

**Lemma 2.3.2.** Let \( n \in \{1, 2, \ldots, N\} \). For any function \( g : \mathbb{R}^M \to \mathbb{R} \) and any real numbers \( \alpha, \delta \) and \( \gamma \), we have,
\[
E\left[ g(Z_n) S_n^\gamma N \left( \frac{\log S_n + \delta}{\alpha} \right) \Big| S_{n-1} = s, Z_{n-1} = e_i \right] = \sum_{j=1}^{M} A_{ji} g(e_j) s^\gamma \exp \left( \gamma (\mu_j - \frac{1}{2} \sigma_j^2) \tau + \frac{1}{2} \gamma^2 \sigma_j^2 \tau \right)
\times N \left( \frac{\log s + (\mu_j - \frac{1}{2} \sigma_j^2) \tau + \delta + \gamma \sigma_j^2 \tau}{\sqrt{\sigma_j^2 \tau + \alpha^2}} \right).
\]

Here, \( S \) denotes the asset price and \( Z \) the Markov chain in the Switching Black–Scholes model.

**Proof.** Using Theorem 2.2.1, which gives the joint conditional density function of \( (S_n, Z_n) \) given \( S_{n-1} = s \) and \( Z_{n-1} = e_i \) under the real world probability measure, the left hand side equals,
\[
\sum_{j=1}^{M} \int_0^{\infty} g(e_j) x^\gamma N \left( \frac{\log x + \delta}{\alpha} \right) \frac{A_{ji}}{\sqrt{2\pi} \sigma_j \sqrt{\tau x}} \frac{1}{\sqrt{2\pi} \sigma_j \sqrt{\tau x}}
\times \exp \left[ -\frac{1}{2} \left( \frac{\log(x/s) - (\mu_j - \frac{1}{2} \sigma_j^2) \tau}{\sigma_j \sqrt{\tau}} \right)^2 \right] dx.
\]

Changing variables to \( z = \frac{\log(x/s) - (\mu_j - \frac{1}{2} \sigma_j^2) \tau}{\sigma_j \sqrt{\tau}} \) gives,
\[
\sum_{j=1}^{M} A_{ji} \sqrt{2\pi} \sigma_j \sqrt{\tau} g(e_j) \int_{-\infty}^{\infty} s^\gamma e^{\gamma(\mu_j - \frac{1}{2} \sigma_j^2) \tau + \gamma \sigma_j \sqrt{\tau} z} e^{-\frac{1}{2} z^2} N (az + b) dz,
\]
where \( a = \sigma_j \sqrt{\tau} / \alpha \) and \( b = (\log s + (\mu_j - \frac{1}{2} \sigma_j^2) \tau + \delta) / \alpha \). Applying Lemma A.1.2 of Appendix 1 with \( \theta = \gamma \sigma_j \sqrt{\tau} \) gives,

\[
\sum_{j=1}^{M} A_{ji} g(e_j) s^{\gamma e^{\gamma(\mu_j - \frac{1}{2} \sigma_j^2) \tau}} e^{\frac{1}{2} \gamma^2 \sigma_j^2 \tau} N \left( \frac{\log s + (\mu_j - \frac{1}{2} \sigma_j^2) \tau + \delta + \gamma \sigma_j^2 \tau}{\sqrt{\sigma_j^2 \tau + \alpha^2}} \right),
\]

as required.

The final lemma deals with an expectation involving only the Markov chain.

**Lemma 2.3.3.** Let \( n \in \{1, 2, ..., N\} \) and \( X_n = (\mu(Z_n) - \frac{1}{2} \sigma(Z_n)^2) \tau + \sigma(Z_n) \Delta W_n \).

For any \( \mathcal{F}_{n-1} \)-measurable function, \( \beta \), the expectation of \( e^{\beta X_n} \) given \( S_{n-1} = s \) and \( Z_{n-1} = e_i \) is,

\[
E[e^{\beta X_n} | S_{n-1} = s, Z_{n-1} = e_i] = \sum_{j=1}^{M} A_{ji} e^{\beta(s, e_i)(\mu_j - \frac{1}{2} \sigma_j^2) \tau + \frac{1}{2} \gamma \sigma_j^2 \tau \beta(s, e_i)^2}.
\]

Here, \( S \) denotes the asset price and \( Z \) the Markov chain in the Switching Black–Scholes model.

**Proof.** We calculate the conditional expectation directly. Begin by conditioning also on \( Z_n = e_j \) to obtain,

\[
E[e^{\beta X_n} | S_{n-1} = s, Z_{n-1} = e_i] = \sum_{j=1}^{M} A_{ji} \mathbb{E}[e^{\beta(s, e_i)(\mu_j - \frac{1}{2} \sigma_j^2) \tau + \sigma_j \sqrt{\tau} \beta(s, e_i) \Delta W_n}].
\]

The only random part of this expression is \( \Delta W_n \), which is normally distributed with mean 0 and variance \( \tau \). Hence the expectation becomes,

\[
\sum_{j=1}^{M} A_{ji} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\beta(s, e_i)(\mu_j - \frac{1}{2} \sigma_j^2) \tau + \sigma_j \sqrt{\tau} \beta(s, e_i) z} e^{-\frac{1}{2} z^2} \, dz.
\]

Using Lemma A.1.1 of Appendix 1, the integral can be evaluated to give,

\[
E[e^{\beta X_n} | S_{n-1} = s, Z_{n-1} = e_i] = \sum_{j=1}^{M} A_{ji} e^{\beta(s, e_i)(\mu_j - \frac{1}{2} \sigma_j^2) \tau + \frac{1}{2} \beta(s, e_i)^2 \sigma_j^2 \tau},
\]

as required.
Chapter 3

Option Pricing via
Mean–Variance Hedging - Part 1

In this chapter and the next, the method of mean-variance hedging, as developed by Martin Schweizer in the papers [69] – [72], is applied to price a call option on an asset following the Switching Black–Scholes model. We discuss in Section 3.1 exactly what “mean-variance hedging” involves, and see that the mean-variance price, \( H_0 \), of a contingent claim, \( H \), can be computed in the usual way, as an expectation under a martingale measure of the discounted claim at expiry. In particular,

\[
H_0 = \mathbb{E}^Q \left[ \frac{H}{B_T} \right],
\]

where \( Q \) is the \( b \)-variance-optimal signed martingale measure, which will be defined in Section 3.1.3. In Section 3.1.4 the precise structure of this measure in discrete time is given.

The bulk of this chapter, and the whole of the next chapter, is devoted to computing the mean-variance price of a call option on an asset following the Switching Black–Scholes model. The two chapters find the same price, but using different procedures, which give rise to different representations of the price. In Chapter 4, we use conditional density functions of \( (S_n, Z_n) \) given \( \mathcal{F}_k \) (for \( n > k \)) under the measure \( Q \), which results in a forward inductive procedure, whereas
in the current chapter, we calculate the price directly by conditioning backwards in time.

After describing the method of mean-variance hedging in Section 3.1, and giving a formula for the b-variance-optimal signed martingale measure in terms of a predictable process \((\beta_n)_{n=1,2,\ldots,N}\), we discuss in Section 3.2 how this measure, and in particular the sequence \((\beta_n)\), can be computed, and explain the backwardisation procedure used to find the mean-variance price of a contingent claim.

To describe the remaining content, recall that the Black-Scholes pricing formula,

\[
H_0 = S_0 N(d_1) - K e^{-rT} N(d_2),
\]

(3.1)
gives the price of a call option with strike price, \(K\), on an asset, \(S\), evolving as geometric Brownian motion with constant drift, \(\mu\), and constant volatility, \(\sigma\). In this formula, \(N\) is the normal distribution function,

\[
d_1 = \frac{\log(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma \sqrt{T}},
\]

\[
d_2 = d_1 - \sigma \sqrt{T}, \text{ and } r \text{ is the risk-free, continuously compounding interest rate per annum.}
\]

We will see that the mean-variance price of a call option with strike price, \(K\), on an asset following the Switching Black–Scholes model is,

\[
H_0 = \frac{e^{-rT}}{D_0} \sum_{j_1,\ldots,j_N=1}^{M} f(j_1,\ldots,j_N) \left( S_0 \mathcal{R}(j_{1,N}) - K \mathcal{Q}(j_{1,N}) \right),
\]

(3.2)
where \(f(j_1,\ldots,j_N)\) and \(D_0\) are constants. This equation represents the mean-variance price as a sum of terms analogous to the Black-Scholes formula (3.1). The \(\mathcal{R}\) and \(\mathcal{Q}\) terms correspond to \(N(d_1)\) and \(N(d_2)\), and can be written as a sum of \(2^N\) components of the form \(c N(d)\), where \(c\) is a constant, and,

\[
d = \frac{\log(S_0/K) + b}{a} + \xi,
\]

(3.3)
for \(a = (\sigma_{j_1}^2 + \ldots + \sigma_{j_N}^2)^{1/2} \sqrt{T}\), \(b = \sum_{i=1}^{N-1} (\mu_{j_i} - \frac{1}{2}\sigma_{j_i}^2) \tau + (\mu_{j_N} + \frac{1}{2}\sigma_{j_N}^2) \tau\) and \(\xi\) a constant depending on the \(\sigma_{j_i} \sqrt{T}\) values. We see that the \(d\) values in Equation (3.3) have the same structure as \(d_1\) and \(d_2\) in the Black-Scholes formula.
In Section 3.3, we show how these $d$ values can be calculated, using a backward recursive method. This is followed in Section 3.4 by recursive definitions of the $Q$ and $R$ values, and a proof of the pricing formula (3.2).

In Section 3.5, we consider the special case of no switching (that is, $M = 1$), in which the $Q$ and $R$ values simplify considerably, becoming sums over $N + 1$ rather than $2^N$ terms.

We begin with an overview of the mean-variance hedging method of pricing.

### 3.1 What is Mean-Variance Hedging?

Consider a complete probability space $(\Omega, \mathcal{F}, P)$ with a filtration $(\mathcal{F}_t)_{t \in \mathbb{R}}$ for a given time index $T \subseteq [0, T]$, $T > 0$. Let $X = (X_t)_{t \in \mathbb{R}}$ be a real-valued, adapted process representing the discounted price of a risky asset, with numeraire $B = (B_t)_{t \in \mathbb{R}}$, a riskless bond whose value is deterministic at all times and has initial value 1. A trading strategy in such a market specifies the number of assets $X$ and $B$ held at each time $t \in T$. If, in addition, the trading strategy is self-financing, it can be completely described by the initial capital, and the number of asset $X$ held at each time. Let $\Theta$ denote the set of all possible self-financing trading strategies (so that $\Theta$ contains real-valued, $(\mathcal{F}_t)$-predictable processes $\theta = (\theta_t)_{t \in \mathbb{R}}$, with $\theta_t$ representing the number of $X$ held at time $t$), and let $G_T(\theta)$ denote the gain at time $T$ obtained from following the trading strategy $\theta$.

#### 3.1.1 Description of the Method

In this market, a contingent claim is a real-valued $\mathcal{F}_T$-measurable random variable, $H$, with $\mathbb{E}[|H|] < \infty$. Mean-variance hedging is a method for pricing and hedging such a claim, by choosing an initial wealth, $c$, and a self-financing trading strategy, $\theta$, such that the final wealth, $cB_T + G_T(\theta)$, is as close as possible.

---

1This means that $G_T(\theta) = \int_0^T \theta_t \, dX_t$. 
to $H$. A quadratic criterion is used to measure closeness.

Formulating this mathematically, mean-variance hedging gives an optimal pair $(H_0, \theta^*)$ which is the solution to the problem,

$$\text{minimise } E[(H - c B_T - G_T(\theta))^2] \text{ over all } c \in \mathbb{R} \text{ and } \theta \in \Theta. \quad (3.4)$$

A financial interpretation is: if we sell the claim $H$ at time 0 and use initial capital $H_0$ to follow the trading strategy $\theta^*$, then we are minimising the expected value of the squared net loss at time $T$. The number $H_0$ is a natural choice for the price of the claim. If $H$ is attainable, then $H_0$ agrees with the price that would be assigned in a complete market, which confirms that mean-variance hedging gives a consistent extension of pricing from complete to incomplete markets.

The paper [72] gives an alternative interpretation of mean-variance hedging. Suppose at time 0 we sell $H$ for an amount $c$, and using this as initial capital, execute a self-financing trading strategy, $\theta$, with outcome $G_T(\theta)$. At maturity of the claim, our wealth is,

$$\text{wealth} := c B_T + G_T(\theta) - H.$$

If we define,

$$\text{risk}(\theta, H) := \text{Var}[\text{wealth}] = \text{Var}[G_T(\theta) - H],$$

$$\text{profit}(c, \theta, H) := E[\text{wealth}] = c B_T + E[G_T(\theta) - H],$$

the mean-variance optimal pair $(H_0, \theta^*)$ is found by minimising $\text{risk}(\theta, H)$ over all trading strategies $\theta \in \Theta$ to give an optimal $\theta^*$, and then choosing $H_0$ such that $\text{profit}(H_0, \theta^*, H) = 0$. Thus,

$$H_0 = \frac{1}{B_T} E[H - G_T(\theta^*)].$$

This formulation of mean-variance hedging gives a second economic motivation for choosing $H_0$ as the price of the claim. To find $H_0$, we choose a self-financing trading strategy to minimise the risk (that is, the variance of the difference
between the value of this strategy and the claim), and then set $H_0$ to be the initial capital required to ensure the expected profit from following this strategy is zero. If this expectation were more than zero, others could sell the claim for a lower price, and if it were less than zero, we are likely to make a loss. Thus $H_0$ is just the right price.

### 3.1.2 Ensuring a Unique Solution Exists

When using mean-variance hedging for the valuation of contingent claims, we need to impose conditions to ensure the existence and uniqueness of solutions to the optimisation problems involved. The papers [69] – [72] of Martin Schweizer develop the theory of mean-variance hedging in detail when the discounted asset price and contingent claim are square integrable (that is, $X_t$ and $H$ belong to the space $L^2 = L^2(\Omega, \mathcal{F}, P)$ for all $t \in T$), and the set of admissible trading strategies, $\Theta$, is a linear space containing predictable processes, $\theta$, such that the gains process $G(\theta) = (G_t(\theta))_{t \in T}$ is square integrable. These conditions imply existence of the required expectations. Schweizer also assumes a frictionless market.

In the special case of discrete-time, $T = \{1, 2, \ldots, T\}$, a solution to (3.4) is guaranteed by assuming that the mean-variance tradeoff process,

$$
\tilde{K}_t := \sum_{j=1}^{t} \left( \frac{\mathbb{E}[\Delta X_j | \mathcal{F}_{j-1}]}{\text{Var} [\Delta X_j | \mathcal{F}_{j-1}]} \right)^2
$$

is $P$-a.s. bounded, \hspace{1cm} (3.5)

(uniformly in $\omega$ and $j$), as then the space $G_T(\Theta) = \{G_T(\theta) : \theta \in \Theta\}$ is closed in $L^2$, by Theorem 2.1 of [70]. (This paper also gives conditions equivalent to (3.5)). However, if we are only interested in pricing claims, and not in finding a hedging strategy, we can minimise over the closure of $G_T(\Theta)$ to obtain the price, $H_0$, of $H$. 

3.1.3 The Mean-Variance Price as an Expectation

In his papers, Schweizer shows how, under the conditions above, the mean-variance price can be expressed in the usual way as the expected value of the discounted claim under a martingale measure. To explain this, we make the following definitions, and use \( b \) to denote \( B_T \).

**Definition 3.1.1.** A signed measure\(^2\) \( \tilde{P} \) on \((\Omega, \mathcal{F})\) is called a signed \((\Theta, b)\)-martingale measure if \( \tilde{P}(\Omega) = 1, \tilde{P} \ll P, \frac{1}{b} \frac{d\tilde{P}}{dP} \in L^2 \) and,

\[
E \left[ \frac{d\tilde{P}}{dP} G_T(\theta) \right] = 0 \quad \text{for all } \theta \in \Theta. \tag{3.6}
\]

Denote by \( \mathbb{P}_s(\Theta) \) the set of all signed \((\Theta, b)\)-martingale measures.

Note that for many choices of \( \Theta \), the condition (3.6) is equivalent to,

\[
E \left[ \frac{d\tilde{P}}{dP} (X_t - X_s) \bigg| \mathcal{F}_s \right] = 0 \quad \text{for all } s, t \in \mathbb{T}, \ s \leq t, \tag{3.7}
\]

which explains the term, signed martingale measure.

**Definition 3.1.2.** If \( \mathbb{P}_s(\Theta) \) is non-empty, an element \( Q \in \mathbb{P}_s(\Theta) \) is said to be \( b \)-variance-optimal if it minimises,

\[
\left\| \frac{1}{b} \frac{d\tilde{P}}{dP} \right\|_{L^2},
\]

over all \( \tilde{P} \in \mathbb{P}_s(\Theta) \).

If we assume \( \mathbb{P}_s(\Theta) \) is non-empty, existence and uniqueness of a \( b \)-variance-optimal signed martingale measure follows immediately from the observation that \( \left\{ \frac{1}{b} \frac{d\tilde{P}}{dP} : \tilde{P} \in \mathbb{P}_s(\Theta) \right\} \) is a non-empty, closed, convex set. In his paper [72], Schweizer showed that the condition \( \mathbb{P}_s(\Theta) \) is non-empty is equivalent to assuming that \((\Theta, b)\) admits no approximate profits in \( L^2 \), defined to mean that the closure of the space \( G_T(\Theta) \) does not contain \( b \). Interpreted financially, this

\(^2\)A signed measure is a measure which may take negative values.
means that the riskless payoff $b$ cannot be hedged by a self-financing trading strategy with initial wealth $0$, which is a no-arbitrage condition.

Under the assumption that $(\Theta, b)$ admits no approximate profits in $L^2$, Schweizer showed that the mean-variance price of $H$ can be expressed as,

$$H_0 = E^Q \left[ \frac{H}{B_T} \right], \quad (3.8)$$

where $Q$ is the unique $b$-variance-optimal signed martingale measure.

### 3.1.4 Mean-Variance Hedging in Discrete-Time

In the discrete-time scenario, Schweizer gave a method for finding the $b$-variance-optimal signed martingale measure explicitly. To explain this procedure, let the time index be $\mathbb{T} = \{1, 2, \ldots, T\}$, and let,

$$\Theta = \{\text{predictable processes } \theta = (\theta_n)_{n=1,2,\ldots,T} \text{ such that } \theta_n \Delta X_n \in L^2 \text{ for } n = 1, 2, \ldots, T\}. \quad (3.9)$$

Note that for this choice of $\Theta$, conditions (3.6) and (3.7) are equivalent. Given these assumptions, the $b$-variance-optimal signed martingale measure, $Q$, is defined by the equation,

$$\frac{dQ}{dP} = \frac{1}{\bar{c}} \prod_{n=1}^{T} (1 - \beta_n \Delta X_n), \quad (3.10)$$

where $\bar{c} = E \left[ \prod_{n=1}^{T} (1 - \beta_n \Delta X_n) \right]$ and $(\beta_n)_{n=1,2,\ldots,T}$ is a predictable process, called the adjustment process of $X$, which will be defined in Section 3.2. Furthermore, if the mean-variance tradeoff process $\hat{K}$ is bounded, then the optimal trading strategy is defined inductively by,

$$\theta_n^* = \rho_n - \beta_n \left( H_0 + \sum_{j=1}^{n-1} \theta_j^* \Delta X_j \right),$$

where,

$$\rho_n := \frac{E \left[ \frac{H}{B_T} \Delta X_n \prod_{j=n+1}^{T} (1 - \beta_j \Delta X_j) \bigg| \mathcal{F}_{n-1} \right]}{E \left[ \Delta X_n^2 \prod_{j=n+1}^{T} (1 - \beta_j \Delta X_j)^2 \bigg| \mathcal{F}_{n-1} \right]},$$
for $n = 1, 2, \ldots, T$.

For the remainder of this chapter, we apply Schweizer’s interpretation of mean-variance hedging to find the price of a call option on an asset following the Switching Black–Scholes model. It is easily verified that for this model, $X$ and $H$ are square integrable, $(\Theta, b)$ admits no approximate profits in $L^2$, and $\tilde{K}$ is bounded, so use of this method is valid.

### 3.2 The $b$-Variance-Optimal Signed Martingale Measure

In this section, we calculate the $b$-variance-optimal signed martingale measure, $Q$, for the Switching Black–Scholes model, and explain the backwardisation procedure used to find the mean-variance price of a contingent claim. It was shown in Section 3.1.4 that,

$$
\frac{dQ}{dP} = \frac{1}{\tilde{c}} \prod_{n=1}^{N} (1 - \beta_n \Delta X_n),
$$

where $\tilde{c} = E\left[\prod_{n=1}^{N} (1 - \beta_n \Delta X_n)\right]$ and $(\beta_n)_{n=1,2,\ldots,N}$ is a predictable process. We now define the sequence $(\beta_n)$.

**Definition 3.2.1.** The adjustment process $(\beta_n)_{n=1,2,\ldots,N}$ of the discounted asset price, $X$, is defined on the probability space $(\Omega, \mathcal{F}, P)$ by setting,

\begin{equation}
\beta_n = \frac{E\left[\Delta X_n \prod_{j=n+1}^{N} (1 - \beta_j \Delta X_j) \bigg| \mathcal{F}_{n-1}\right]}{E\left[\Delta X_n^2 \prod_{j=n+1}^{N} (1 - \beta_j \Delta X_j)^2 \bigg| \mathcal{F}_{n-1}\right]},
\end{equation}

for $n = 1, 2, \ldots, N$. Here, we use the convention that an empty product equals 1, and we set $\beta_n = 0$ if both the numerator and denominator are 0.

In order to calculate the sequence $(\beta_n)$, we define a stochastic process $(D_n)_{n=0,1,\ldots,N}$ as follows.
Definition 3.2.2. Define a stochastic process \( D = \{ D_n : n = 0, 1, \ldots, N \} \) on the probability space \( (\Omega, \mathcal{F}, P) \) by setting \( D_N \equiv 1 \) and,

\[
D_n = E \left[ \prod_{k=n+1}^{N} (1 - \beta_k \Delta X_k) \bigg| \mathcal{F}_n \right],
\]

for \( n = 0, 1, \ldots, N - 1 \).

We also use the following notation for ease of description.

Notation 3.2.3. For \( i, j = 1, 2, \ldots, M \) we define,

\[
B_{ij} = A_{ij} (e^{\mu_j \tau} - e^{\mu_i \tau}),
\]

\[
C_{ij} = A_{ij} (e^{(2\mu_j + \sigma_j^2)\tau} - 2e^{\mu_i \tau} e^{\mu_j \tau} + e^{2\mu_j \tau}).
\]

The following theorem shows how the sequence \( (\beta_n) \) can be calculated using the stochastic process \( (D_n) \).

Theorem 3.2.4 (Calculating the \( \beta \)s). Consider a Switching Black–Scholes market as described in Chapter 2. For \( n = 1, 2, \ldots, N \) we have,

\[
D_{n-1} = E[D_n | \mathcal{F}_{n-1}] - \beta_n E[\Delta X_n D_n | \mathcal{F}_{n-1}]; \quad (3.12)
\]

\[
\beta_n = \frac{E[\Delta X_n D_n | \mathcal{F}_{n-1}]}{E[\Delta X_n^2 D_n | \mathcal{F}_{n-1}]}; \quad (3.13)
\]

\[
E[\Delta X_n D_n | S_{n-1} = s, Z_{n-1} = e_i] = \frac{s \text{er}_{n\tau}}{\text{er}_{n\tau}} \sum_{j=1}^{M} D_n(e_j) B_{ij}; \quad (3.14)
\]

\[
E[\Delta X_n^2 D_n | S_{n-1} = s, Z_{n-1} = e_i] = \left( \frac{s \text{er}_{n\tau}}{\text{er}_{n\tau}} \right)^2 \sum_{j=1}^{M} D_n(e_j) C_{ij}. \quad (3.15)
\]

Also, the random variable \( D_{n-1} \) depends only on \( Z_{n-1} \) and,

\[
D_{n-1}(e_j) := D_{n-1}(s, e_j) = \sum_{i=1}^{M} A_{ij} D_n(e_i) - \frac{\left( \sum_{i=1}^{M} D_n(e_i) B_{ji} \right)^2}{\sum_{i=1}^{M} D_n(e_i) C_{ji}}. \quad (3.16)
\]

Proof. We start by proving Equation (3.12). Conditioning first on the \( \sigma \)-algebra \( \mathcal{F}_n \) in the definition of \( D_{n-1} \) shows that,

\[
D_{n-1} = E[(1 - \beta_n \Delta X_n) D_n | \mathcal{F}_{n-1}].
\]
The $\mathcal{F}_{n-1}$-measurability of $\beta_n$ then gives Equation (3.12).

The remainder of the proof uses backward induction. Start with $n = N$. Equation (3.13) follows immediately from the definition of $\beta_N$ on page 34 and the fact that $D_N \equiv 1$, and Equations (3.14) and (3.15) for $n = N$ are direct consequences of Lemma 2.3.1.

For Equation (3.16) at $n = N$, we have from Equation (3.12),

$$D_{N-1} = E[D_N|\mathcal{F}_{N-1}] - \beta_N E[\Delta X_N D_N|\mathcal{F}_{N-1}] = 1 - \beta_N E[\Delta X_N|\mathcal{F}_{N-1}].$$

Using Equations (3.13), (3.14) and (3.15), and the relation $\sum_{i=1}^{M} A_{ij} = 1$, we see that $D_{N-1}$ depends only on $Z_{N-1}$, and,

$$D_{N-1}(e_j) = \sum_{i=1}^{M} A_{ij} D_N(e_i) - \frac{(\sum_{i=1}^{M} D_N(e_i) B_{ji})^2}{\sum_{i=1}^{M} D_N(e_i) C_{ji}},$$

as required.

Now, assume that Equations (3.13) - (3.16) are true for $n \geq k + 1$. We need to prove that the results are true for $n = k$. Conditioning first on the $\sigma$-algebra $\mathcal{F}_k$ in the following expression and using the definition of $D_k$, we see that,

$$E\left[\Delta X_k \prod_{j=k+1}^{N} (1 - \beta_j \Delta X_j) \bigg| \mathcal{F}_{k-1}\right] = E[\Delta X_k D_k|\mathcal{F}_{k-1}].$$

Using the result given in Equation (2.6) of Proposition 2.3 in Schweizer’s paper [70], which states that,

$$E\left[\prod_{j=k+1}^{N} (1 - \beta_j \Delta X_j)^2 \bigg| \mathcal{F}_k\right] = E\left[\prod_{j=k+1}^{N} (1 - \beta_j \Delta X_j) \bigg| \mathcal{F}_k\right],$$

we also obtain,

$$E\left[\Delta X_k^2 \prod_{j=k+1}^{N} (1 - \beta_j \Delta X_j)^2 \bigg| \mathcal{F}_{k-1}\right] = E[\Delta X_k^2 D_k|\mathcal{F}_{k-1}].$$

From the definition of $\beta_k$ on page 34, this proves Equation (3.13) for $n = k.$
For Equation (3.14), we must compute the conditional expectation. Equation (3.16) at \( n = k + 1 \) tells us that \( D_k \) depends only on \( Z_k \), so using the conditional density function in Theorem 2.2.1, we obtain,

\[
E[\Delta X_k D_k | S_{k-1} = s, Z_{k-1} = e_i] \\
= \sum_{j=1}^{M} \int_0^{\infty} \frac{x}{e^{rkr} - \frac{s}{e^{r(k-1)r}}} \, D_k(e_j) \\
\times \frac{A_{ji}}{\sqrt{2\pi} \sigma_j \sqrt{r}} \exp \left[ -\frac{1}{2} \frac{(\log(x/s) - (\mu_j - \frac{1}{2}\sigma_j^2)r)}{\sigma_j \sqrt{r}} \right]^2 dx.
\]

Changing variables to \( z = \frac{\log(x/s) - (\mu_j - \frac{1}{2}\sigma_j^2)r}{\sigma_j \sqrt{r}} \) gives,

\[
\sum_{j=1}^{M} D_k(e_j) \frac{A_{ji}}{\sqrt{2\pi} \sigma_j \sqrt{r}} \int_{-\infty}^{\infty} \left( e^{(\mu_j - \frac{1}{2}\sigma_j^2)r + \sigma_j \sqrt{r}z} - e^{r \sigma_j^2} \right) e^{-\frac{1}{2}z^2} \, dz,
\]

and applying Lemma A.1.1 of Appendix 1 to calculate the integral gives,

\[
\sum_{j=1}^{M} D_k(e_j) A_{ji} \frac{s}{e^{rkr}} (e^{\mu_j r} - e^{r \sigma_j^2}) = \frac{s}{e^{rkr}} \sum_{j=1}^{M} D_k(e_j) B_{ij},
\]

which proves Equation (3.14) for \( n = k \).

Similarly, for Equation (3.15) we have,

\[
E[\Delta X_k^2 D_k | S_{k-1} = s, Z_{k-1} = e_i] \\
= \sum_{j=1}^{M} D_k(e_j) \frac{A_{ji}}{\sqrt{2\pi} e^{2rkr}} \int_{-\infty}^{\infty} \left( e^{(\mu_j + \frac{1}{2}\sigma_j^2)r + \sigma_j \sqrt{r}z} - e^{2r \sigma_j^2} \right)^2 e^{-\frac{1}{2}z^2} \, dz.
\]

Expanding the quadratic and using Lemma A.1.1, the integral is equal to,

\[
\sqrt{2\pi} \left[ e^{(2\mu_j + \sigma_j^2)r} - 2e^{r \sigma_j r} e^{\mu_j r} + e^{2r \sigma_j^2} \right],
\]

and so,

\[
E[\Delta X_k^2 D_k | S_{k-1} = s, Z_{k-1} = e_i] = \frac{s^2}{e^{2rkr}} \sum_{j=1}^{M} D_k(e_j) C_{ij},
\]

as required.

Finally, we prove Equation (3.16) at \( n = k \). Equation (3.12) gives,

\[
D_{k-1} = E[D_k | \mathcal{F}_{k-1}] - \beta_k E[\Delta X_k D_k | \mathcal{F}_{k-1}].
\]
The form of $\beta$ given in Equation (3.13) and the independence of $D_k$ and $S_{k-1}$ allows us to write,

$$D_{k-1} = E[D_k | Z_{k-1}] = \frac{E[\Delta X_k D_k | \mathcal{F}_{k-1}]}{E[\Delta X_k^2 D_k | \mathcal{F}_{k-1}]}.$$

Equations (3.14) and (3.15) at $n = k$, and the density function calculated in Theorem 2.2.3 then give the result. \hfill \Box

This theorem shows how to calculate the sequence $(D_n)_{n=0,1,\ldots,N}$ recursively by setting $D_N \equiv 1$ and,

$$D_{n-1}(e_j) = \sum_{i=1}^{M} A_{ij} D_n(e_i) - \frac{\left(\sum_{i=1}^{M} D_n(e_i) B_{ji}\right)^2}{\sum_{i=1}^{M} D_n(e_i) C_{ji}},$$

for $j = 1, 2, \ldots, M$. We can then compute $(\beta_n)$ using Equations (3.13)–(3.15), which enables evaluation of the Radon-Nikodym derivative,

$$\frac{dQ}{dP} = \frac{1}{D_0} \prod_{n=1}^{N} (1 - \beta_n \Delta X_n).$$

From Equation (3.8), the mean-variance price of a contingent claim, $H$, is then,

$$H_0 = \frac{e^{-rT}}{D_0} E\left[H \prod_{n=1}^{N} (1 - \beta_n \Delta X_n) \right].$$

The following definition allows us to calculate the mean-variance price of a call option using a backward recursive procedure, obtained by conditioning repeatedly on $\sigma$-algebras of the filtration.

**Definition 3.2.5.** For $K \geq 0$ define $F_N = (S_N - K)^+$, and for $n = 1, 2, \ldots, N$, let,

$$F_{n-1} = E\left[F_n (1 - \beta_n \Delta X_n) \left| \mathcal{F}_{n-1} \right. \right],$$

where $(\beta_n)$ is the adjustment process of the discounted asset price, $X$, defined in 3.2.1.

Using backward induction it can be shown that,

$$F_{n-1} = E\left[F_N \prod_{j=n}^{N} (1 - \beta_j \Delta X_j) \left| \mathcal{F}_{n-1} \right. \right].$$
Therefore, if $H$ is a call option with strike price, $K$, on the asset, $S$, the mean-variance price of $H$ is,

$$H_0 = \frac{e^{-rT}}{D_0} F_0,$$

and we need to evaluate $F_0$ to calculate the price. As explained at the beginning of the chapter, we will show that,

$$F_0 = \sum_{j_1, \ldots, j_N = 1}^{M} f(j_1, \ldots, j_N) \left( S_0 \mathcal{R}(j_{1,N}) - K \mathcal{Q}(j_{1,N}) \right),$$

where $f(j_1, \ldots, j_N)$ is a constant, and $\mathcal{R}$ and $\mathcal{Q}$ are sums of terms $c N (d)$, with,

$$d = \frac{\log(S_0/K) + b}{a} + \xi,$$

for constants $a, b, c, \xi$. In the next section, we define these $d$ values and prove they can be written in this way, as well as showing how they can be computed, and in Section 3.4 we prove Equation (3.18).

### 3.3 The $d$ Values

To define appropriate $d$ values, we use the functions $g^j_n$, given in the following definition. Recall that $j_{i,k}$ is used to denote the vector $(j_i, j_{i+1}, j_{i+2}, \ldots, j_k)$ of $\mathbb{R}^{k-i+1}$.

**Definition 3.3.1.** For $m \in \mathbb{N}$, $j_1 \in \{1, 2, \ldots, M\}$ and $a, b \in \mathbb{R}$, define functions $g^j_n : \mathbb{R} \to \mathbb{R}$ by,

$$g^j_n \left( \frac{\log S_n + b}{a} \right) = \frac{\log S_{n-1} + (\mu_{j_i} - \frac{1}{2}\sigma_{j_i}^2)\tau + b + (m - 1)\sigma_{j_i}^2 \tau}{\sqrt{\sigma_{j_i}^2 \tau + a^2}},$$

for $n \in \{1, 2, \ldots, N\}$. Here, $S$ is our asset, which follows the Switching Black–Scholes model, and $\mu_{j_i}$, $\sigma_{j_i}$, and $\tau$ are parameters of this model.

The $d$ values we require make use of specific values of $a$ and $b$. These are defined recursively in Definition 3.3.2. The recursive relationship is motivated by the definition of $g^j_n$, given above.
Definition 3.3.2. For \( n = 0, 1, \ldots, N - 1 \) and \( K \in (0, \infty) \), define the functions \( a_{N-n}, b_{N-n} : \{1, 2, \ldots, M\}^{n+1} \to \mathbb{R} \) by the recursions,

\[
a_{N-n}(j_{1,n+1}) = \sqrt{\sigma_{j_1}^2 \tau + a_{N-n+1}(j_{2,n+1})^2},
\]

\[
b_{N-n}(j_{1,n+1}) = (\mu_{j_1} - \frac{1}{2} \sigma_{j_1}^2) \tau + b_{N-n+1}(j_{2,n+1}),
\]

with initial functions \( a_N(j) = \sigma_j \sqrt{\tau} \) and \( b_N(j) = (\mu_j + \frac{1}{2} \sigma_j^2) \tau - \log K \).

The following lemma reveals the form of each \( a_{N-n} \) and \( b_{N-n} \).

Lemma 3.3.3. For \( n = 0, 1, \ldots, N - 1 \), the functions \( a_{N-n} \) and \( b_{N-n} \) defined in Equations (3.20) and (3.21) are given by,

\[
a_{N-n}(j_{1,n+1}) = (\sigma_{j_1}^2 + \sigma_{j_2}^2 + \ldots + \sigma_{j_{n+1}}^2)^{1/2} \sqrt{\tau},
\]

\[
b_{N-n}(j_{1,n+1}) = \sum_{i=1}^{n} (\mu_{j_i} - \frac{1}{2} \sigma_{j_i}^2) \tau + (\mu_{j_{n+1}} + \frac{1}{2} \sigma_{j_{n+1}}^2) \tau - \log K.
\]

Proof. We use induction on \( n \). The results hold for \( n = 0 \), by definition. Suppose they are true for arbitrary \( n \). The recursive definition for \( a \) gives,

\[
a_{N-(n+1)}(j_{1,n+2}) = \sqrt{\sigma_{j_1}^2 \tau + a_{N-n}(j_{2,n+2})^2}.
\]

As the induction hypothesis states \( a_{N-n}(j_{2,n+2}) = (\sigma_{j_2}^2 + \ldots + \sigma_{j_{n+2}}^2)^{1/2} \sqrt{\tau} \), we have,

\[
a_{N-(n+1)}(j_{1,n+2}) = (\sigma_{j_1}^2 + \sigma_{j_2}^2 + \ldots + \sigma_{j_{n+2}}^2)^{1/2} \sqrt{\tau},
\]

as required.

The recursive definition for \( b \) gives,

\[
b_{N-(n+1)}(j_{1,n+2}) = (\mu_{j_1} - \frac{1}{2} \sigma_{j_1}^2) \tau + b_{N-n}(j_{2,n+2}).
\]

Again, the induction hypothesis,

\[
b_{N-n}(j_{2,n+2}) = \sum_{i=2}^{n+1} (\mu_{j_i} - \frac{1}{2} \sigma_{j_i}^2) \tau + (\mu_{j_{n+2}} + \frac{1}{2} \sigma_{j_{n+2}}^2) \tau - \log K,
\]

gives the result. \( \qed \)
We now use the functions $a_{N-n}$, $b_{N-n}$ and $g^j_m$ to define our $d$ values, which will be analogous to $d_1$ and $d_2$ of the Black-Scholes pricing formula.

**Definition 3.3.4.** Consider a Switching Black–Scholes market as discussed in Chapter 2. For $j \in \{1, 2, \ldots, M\}$ set,

\[
    d^1_N(S_{N-1}, j) = \frac{\log S_{N-1} + b_N(j)}{a_N(j)} = \frac{\log \left( \frac{S_{N-1}}{K} \right) + \left( \mu_j + \frac{1}{2} \sigma^2_j \right) \tau}{\sigma_j \sqrt{\tau}},
    
    d^2_N(S_{N-1}, j) = d^1_N(S_{N-1}, j) + \sigma_j \sqrt{\tau},
\]

where the functions $a$ and $b$ are defined in 3.3.2.

Also, for $n = 1, 2, \ldots, N-1$, make the following recursive definitions, where $j_{1,n+1} \in \{1, 2, \ldots, M\}^{n+1}$ and the functions $g^j_m$, $m = 1, 2, 3$, are given in Definition 3.3.1.

\[
    d^{p+1}_{N-n}(S_{N-n-1}, j_{1,n+1}) = \left\{ \begin{array}{ll}
        g^1_{j_1} \left( d^{p}_N(S_{N-n-1}, j_{2,n+1}) \right), & 1 \leq i \leq 2^{n-1}, \\
        g^2_{j_2} \left( d^{p}_N(S_{N-n-1}, j_{2,n+1}) \right), & 2^{n-1} < i \leq 2^n,
    \end{array} \right.
\]

\[
    d^{p+1}_{N-n}(S_{N-n-1}, j_{1,n+1}) = \left\{ \begin{array}{ll}
        g^2_{j_2} \left( d^{p}_N(S_{N-n-1}, j_{2,n+1}) \right), & 1 \leq i \leq 2^{n-1}, \\
        g^3_{j_3} \left( d^{p}_N(S_{N-n-1}, j_{2,n+1}) \right), & 2^{n-1} < i \leq 2^n.
    \end{array} \right.
\]

Note from Equation (3.19) that the functions $g^j_m$ preserve the form of their argument, so the expressions above are well-defined. We will see in Section 3.4 that the $d^p_N$ values ($p = 1, 2, \ldots, 2^N$) correspond to $d_1$ and $d_2$ of the Black-Scholes pricing formula.

The remainder of this section proves three results which will be useful in calculating these $d$ values. In particular, Theorems 3.3.5 and 3.3.7 show how to calculate $d^p_{N-n}$ from the values of $a$, $b$ and $\sigma_j \sqrt{\tau}$. The first of these results expresses $d^p_{N-n}$ in the form $\frac{\log S_{N-n-1} + b_{N-n}}{a_{N-n}} + \xi$, for constants $a$, $b$ and $\xi$.

**Theorem 3.3.5.** For $n = 0, 1, 2, \ldots, N-1$ and $p = 1, 2, \ldots, 2^{n+1}$, the function $d^p_{N-n}$ defined in 3.3.4 has the property,

\[
    d^p_{N-n}(S_{N-n-1}, j_{1,n+1}) = \frac{\log S_{N-n-1} + b_{N-n}(j_{1,n+1}) + \xi(j_{1,n+1})}{a_{N-n}(j_{1,n+1})},
    \tag{3.22}
\]

for some function $\xi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$. In particular, for $p = 1$ we have $\xi \equiv 0$. 

Proof. The proof is by induction on $n$. By definition, the result holds for $n = 0$ with $\xi(j_1) = 0$ for $p = 1$ and $\xi(j_1) = \sigma_j^2 \tau$ for $p = 2$.

Next, suppose that Equation (3.22) holds for arbitrary $n$, and let,

$$d_{N-n}^1(S_{N-n-1}, j_{2,n+2}) = \frac{\log S_{N-n-1} + b_{N-n}(j_{2,n+2}) + \eta(j_{2,n+2})}{a_{N-n}(j_{2,n+2})},$$

for some function $\eta : \mathbb{R}^{n+1} \to \mathbb{R}$ and $i \in \{1, 2, \ldots, 2^{n+1}\}$. From Definition 3.3.4, $d_{N-n-1}^p$ is calculated by applying one of the functions $g_{m}^{j_1}$, $m \in \{1, 2, 3\}$, to $d_{N-n}^1(S_{N-n-1}, j_{2,n+2})$, for a particular $i$. The assumption (3.23) and the definition of the function $g_{m}^{j_1}$ (3.3.1) gives,

$$g_{m}^{j_1}(d_{N-n}^1(S_{N-n-1}, j_{2,n+2}))$$

$$= \frac{\log S_{N-n-2} + (\mu_{j_1} - \frac{1}{2} \sigma_j^2 \tau \sigma_j^2 \tau + b_{N-n}(j_{2,n+2}) + \eta(j_{2,n+2}) + (m - 1) \sigma_j^2 \tau}{\sqrt{\sigma_j^2 \tau + a_{N-n}(j_{2,n+2})^2}}.$$

Putting $\xi(j_{1,n+2}) = \eta(j_{2,n+2}) + (m - 1) \sigma_j^2 \tau$ and using the recursive definitions (3.3.2) of the functions $a$ and $b$ proves Equation (3.22).

For the case $p = 1$, we have $\eta = 0$ and $m = 1$, so that $\xi = 0$. This completes the proof. \qed

In Theorem 3.3.7, we express $d_{N-n}^p$ in terms of $d_{N-n}^1$. To prove this theorem, we first need to know the precise effect the function $g_{m}^{j_1}$ has on arguments of the form, $d_{N-n}^p(S_{N-n-1}, j_{2,n+2}) + \eta(j_{2,n+2})$. In the next lemma, we examine this effect.

**Lemma 3.3.6.** Let $n \in \{0, 1, \ldots, N - 1\}$ and $j_1 \in \{1, 2, \ldots, M\}$. Then for any $p \in \{1, 2, \ldots, 2^{n+1}\}$, $m \in \{1, 2, 3\}$ and any function $\eta : \mathbb{R}^{n+1} \to \mathbb{R}$ we have,

$$g_{m}^{j_1}(d_{N-n}^p(S_{N-n-1}, j_{2,n+2}) + \eta(j_{2,n+2}))$$

$$= g_{m}^{j_1}(d_{N-n}^p(S_{N-n-1}, j_{2,n+2}) + \eta(j_{2,n+2}) \frac{a_{N-n}(j_{2,n+2})}{a_{N-n-1}(j_{1,n+2})}. \quad (3.24)$$

**Proof.** From Lemma 3.3.5, the left hand side is equal to,

$$g_{m}^{j_1}\left(\frac{\log S_{N-n-1} + b_{N-n}(j_{2,n+2}) + \xi(j_{2,n+2}) + \eta(j_{2,n+2})a_{N-n}(j_{2,n+2})}{a_{N-n}(j_{2,n+2})}\right),$$
for some function $\xi : \mathbb{R}^{n+1} \to \mathbb{R}$. Applying the definition of $g^i_m$ gives,

$$
\frac{\log S_{N-n-2} + (\mu_{j_1} - \frac{1}{2} \sigma_{j_1}^2) \tau + b_{N-n}(j_{2n+2}) + (m-1)\sigma_{j_1}^2 \tau + \xi(j_{2n+2})}{\sqrt{\sigma_{j_1}^2 \tau + a_{N-n}(j_{2n+2})^2}}
$$

$$
+ \frac{\eta(j_{2n+2})a_{N-n}(j_{2n+2})}{\sqrt{\sigma_{j_1}^2 \tau + a_{N-n}(j_{2n+2})^2}}
$$

which is equal to,

$$
g^i_m \left( d^{p}_{N-n}(S_{N-n-1}, j_{2n+2}) \right) + \eta(j_{2n+2}) \frac{a_{N-n}(j_{2n+2})}{a_{N-n-1}(j_{1,n+1})},
$$

as required. \qed

The final theorem shows how the $d^{p}_{N-n}$ values, for $p = 1, 2, \ldots, 2^{n+1}$, can be computed.

**Theorem 3.3.7.** For $n = 0, 1, \ldots, N-1$ and $1 \leq p \leq 2^n$, the functions $d^{p}_{N-n}$ and $d^{p+2^n}_{N-n}$ defined in 3.3.4 can be computed from the equations,

$$
d^{p}_{N-n}(S_{N-n-1}, j_{1,n+1}) = d^{1}_{N-n}(S_{N-n-1}, j_{1,n+1})
$$

$$
+ \frac{\tau}{a_{N-n}(j_{1,n+1})} \sum_{i=1}^{n} p_{i-1} \sigma_{j_i}^2, \quad (3.25)
$$

$$
d^{p+2^n}_{N-n}(S_{N-n-1}, j_{1,n+1}) = d^{p}_{N-n}(S_{N-n-1}, j_{1,n+1}) + a_{N-n}(j_{1,n+1}), \quad (3.26)
$$

where the constants $p_i \in \{0, 1\}$ are defined by the binary representation of $p - 1$:

$$
p - 1 = \sum_{i=0}^{n-1} p_i 2^i.
$$

**Proof.** First we prove Equation (3.25), using induction on $n$. For $n = 0$, the result is immediate. For $n = 1$, $p$ is either 1 or 2. For $p = 1$, all the coefficients in the binary expansion of $p - 1$ are 0, so the equation holds trivially. For $p = 2$, the only non-zero coefficient is $p_0 = 1$, so the right hand side of Equation (3.25) is,

$$
d^{1}_{N-1}(S_{N-2}, j_1, j_2) + \frac{\tau}{a_{N-1}(j_1, j_2)} \sigma_{j_1}^2.
$$

From the definition of $d^{1}_{N-1}(S_{N-2}, j_1, j_2)$, this is equal to,

$$
g^1_1 \left( d^{1}_{N}(S_{N-1}, j_2) \right) + \frac{\sigma_{j_1}^2 \tau}{a_{N-1}(j_1, j_2)},
$$

for some function $\xi : \mathbb{R}^{n+1} \to \mathbb{R}$. Applying the definition of $g^i_m$ gives,

$$
\frac{\log S_{N-n-2} + (\mu_{j_1} - \frac{1}{2} \sigma_{j_1}^2) \tau + b_{N-n}(j_{2n+2}) + (m-1)\sigma_{j_1}^2 \tau + \xi(j_{2n+2})}{\sqrt{\sigma_{j_1}^2 \tau + a_{N-n}(j_{2n+2})^2}}
$$

$$
+ \frac{\eta(j_{2n+2})a_{N-n}(j_{2n+2})}{\sqrt{\sigma_{j_1}^2 \tau + a_{N-n}(j_{2n+2})^2}}
$$

which is equal to,

$$
g^i_m \left( d^{p}_{N-n}(S_{N-n-1}, j_{2n+2}) \right) + \eta(j_{2n+2}) \frac{a_{N-n}(j_{2n+2})}{a_{N-n-1}(j_{1,n+1})},
$$

as required. \qed

The final theorem shows how the $d^{p}_{N-n}$ values, for $p = 1, 2, \ldots, 2^{n+1}$, can be computed.

**Theorem 3.3.7.** For $n = 0, 1, \ldots, N-1$ and $1 \leq p \leq 2^n$, the functions $d^{p}_{N-n}$ and $d^{p+2^n}_{N-n}$ defined in 3.3.4 can be computed from the equations,

$$
d^{p}_{N-n}(S_{N-n-1}, j_{1,n+1}) = d^{1}_{N-n}(S_{N-n-1}, j_{1,n+1})
$$

$$
+ \frac{\tau}{a_{N-n}(j_{1,n+1})} \sum_{i=1}^{n} p_{i-1} \sigma_{j_i}^2, \quad (3.25)
$$

$$
d^{p+2^n}_{N-n}(S_{N-n-1}, j_{1,n+1}) = d^{p}_{N-n}(S_{N-n-1}, j_{1,n+1}) + a_{N-n}(j_{1,n+1}), \quad (3.26)
$$

where the constants $p_i \in \{0, 1\}$ are defined by the binary representation of $p - 1$:

$$
p - 1 = \sum_{i=0}^{n-1} p_i 2^i.
$$

**Proof.** First we prove Equation (3.25), using induction on $n$. For $n = 0$, the result is immediate. For $n = 1$, $p$ is either 1 or 2. For $p = 1$, all the coefficients in the binary expansion of $p - 1$ are 0, so the equation holds trivially. For $p = 2$, the only non-zero coefficient is $p_0 = 1$, so the right hand side of Equation (3.25) is,

$$
d^{1}_{N-1}(S_{N-2}, j_1, j_2) + \frac{\tau}{a_{N-1}(j_1, j_2)} \sigma_{j_1}^2.
$$

From the definition of $d^{1}_{N-1}(S_{N-2}, j_1, j_2)$, this is equal to,

$$
g^1_1 \left( d^{1}_{N}(S_{N-1}, j_2) \right) + \frac{\sigma_{j_1}^2 \tau}{a_{N-1}(j_1, j_2)},
$$

for some function $\xi : \mathbb{R}^{n+1} \to \mathbb{R}$. Applying the definition of $g^i_m$ gives,
which is \( g_{2j}^i \left( d_N^i(S_{N-1}, j_2) \right) = d_{N-1}^2(S_{N-2}, j_1, j_2) \), as required.

Now suppose Equation (3.25) holds for \( n - 1 \), so that for \( 1 \leq i \leq 2^{n-1} \),

\[
d_{N-n+1}^i(S_{N-n}, j_{2,n+1}) = d_{N-n+1}^1(S_{N-n}, j_{2,n+1}) + \frac{\tau}{a_{N-n+1}(j_{2,n+1})} \sum_{m=1}^{n-1} w_{m-1}\sigma_{j_{m+1}}^2,
\]

(3.27)

where the constants \( w_{m-1} \) come from the representation, \( i - 1 = \sum_{m=0}^{n-2} w_m 2^m \).

We need to show that \( d_{N-n}^p \) has the form of Equation (3.25) for \( p \in \{1, 2, \ldots, 2^n\} \).

We see in Definition 3.3.4 that \( d_{N-n}^p \) is computed from \( d_{N-n+1}^i \), but the definition varies according to whether \( p \) is even or odd. First, consider \( p \) odd, so that \( p = 2i - 1 \) for some \( i \in \{1, 2, \ldots, 2^{n-1}\} \). Using (3.27), the definition,

\[ d_{N-n}^{2i-1}(S_{N-n-1}, j_{1,n+1}) = g_{1i}^1 \left( d_{N-n+1}^i(S_{N-n}, j_{2,n+1}) \right), \]

and Lemma 3.3.6, we obtain,

\[ d_{N-n}^{2i-1}(S_{N-n-1}, j_{1,n+1}) = g_{1i}^1 \left( d_{N-n+1}^1(S_{N-n}, j_{2,n+1}) \right) + \frac{\tau}{a_{N-n}(j_{1,n+1})} \sum_{m=2}^{n} w_m \sigma_{j_m}^2. \]

Next, let \( (2i - 1) - 1 \) have the binary representation \( \sum_{m=0}^{n-1} w_m' 2^m \). From Lemma A.1.4 of Appendix 1, \( w_0' = 0 \) and \( w_m' = w_{m-1} \) for \( m \geq 1 \). Also, using the definition of \( d_{N-n}^1 \) we have,

\[ d_{N-n}^{2i-1}(S_{N-n-1}, j_{1,n+1}) = d_{N-n}^1(S_{N-n-1}, j_{1,n+1}) + \frac{\tau}{a_{N-n}(j_{1,n+1})} \sum_{m=1}^{n} w_m' \sigma_{j_m}^2, \]

(3.28)

as required. Thus Equation (3.25) has been shown for \( p \) odd.

Now consider \( p \) even, so \( p = 2i \) for some \( i \in \{1, 2, \ldots, 2^{n-1}\} \). Definitions give,

\[ d_{N-n}^{2i}(S_{N-n-1}, j_{1,n+1}) = g_{2i}^1 \left( d_{N-n+1}^i(S_{N-n}, j_{2,n+1}) \right) = g_{1i}^1 \left( d_{N-n+1}^i(S_{N-n}, j_{2,n+1}) \right) + \frac{\sigma_{j_i}^2 \tau}{a_{N-n}(j_{1,n+1})}. \]

Also by definition, \( g_{1i}^1 \left( d_{N-n+1}^i(S_{N-n}, j_{2,n+1}) \right) = d_{N-n}^{2i-1}(S_{N-n-1}, j_{1,n+1}) \), so we
may use Equation (3.28) to obtain,

\[
d^2_{N-n} = d^1_{N-n}(S_{N-n-1}, j_{1,n+1}) + \frac{\tau}{a_{N-n}(j_{1,n+1})} \sum_{m=1}^{n} w'_{m-1}\sigma^2_{j_m} + \frac{\sigma^2_{j_1}\tau}{a_{N-n}(j_{1,n+1})} \\
= d^1_{N-n}(S_{N-n-1}, j_{1,n+1}) + \frac{\tau}{a_{N-n}(j_{1,n+1})} \sum_{m=1}^{n} w''_{m-1}\sigma^2_{j_m},
\]

where \(w''_0 = w'_0 + 1 = 1\) and \(w''_m = w'_{m-1} = w_{m-1}\) for \(m \geq 1\). As Lemma A.1.4 of Appendix 1 shows that \(2i - 1 = \sum_{m=0}^{n-1} w''_m 2^m\), this proves Equation (3.25) for \(p = 2i\). Hence it is true for all \(n\) and \(p\), by the principle of mathematical induction.

Now consider Equation (3.26), which holds trivially for \(n = 0\). When \(n = 1\), from Definition 3.3.4 we have,

\[
d^3_{N-1}(S_{N-2}, j_1, j_2) = g^1_{j_2} \left( d^2_N(S_{N-1}, j_2) \right) \\
= g^1_{j_2} \left( d^1_N(S_{N-1}, j_2) + \sigma_j \sqrt{\tau} \right).
\]

Applying Lemma 3.3.6 gives,

\[
d^3_{N-1}(S_{N-2}, j_1, j_2) = g^1_{j_2} \left( d^1_N(S_{N-1}, j_2) + \frac{\sigma^2_{j_2}\tau}{a_{N-1}(j_1, j_2)} \right). \tag{3.29}
\]

But \(g^1_{j_2} \left( d^1_N(S_{N-1}, j_2) \right) = d^2_{N-1}(S_{N-2}, j_1, j_2)\), so from Equation (3.25) with \(n = 1\) and \(p = 2\) we obtain,

\[
d^3_{N-1}(S_{N-2}, j_1, j_2) = d^1_{N-1}(S_{N-2}, j_1, j_2) + \frac{(\sigma^2_{j_1} + \sigma^2_{j_2})\tau}{a_{N-1}(j_1, j_2)}.
\]

As \(a_{N-1}(j_1, j_2) = \sqrt{(\sigma^2_{j_1} + \sigma^2_{j_2})\tau}\) we have,

\[
d^3_{N-1}(S_{N-2}, j_1, j_2) = d^1_{N-1}(S_{N-2}, j_1, j_2) + a_{N-1}(j_1, j_2),
\]

as required.

For \(n = 1\) and \(p = 2\) we have,

\[
d^4_{N-1}(S_{N-2}, j_1, j_2) = g^1_{j_2} \left( d^3_N(S_{N-1}, j_2) \right) \\
= g^1_{j_2} \left( d^2_N(S_{N-1}, j_2) + \frac{\sigma^2_{j_2}\tau}{a_{N-1}(j_1, j_2)} \right).
\]
By definition, \( g_2^{j_1} (d_2^{j_2}(S_{N-1}, j_2)) = d_{N-1}^j(S_{N-2}, j_1, j_2) \), so Equation (3.29) and the subsequent comment give,

\[
d_{N-1}^j(S_{N-2}, j_1, j_2) = d_{N-1}^j(S_{N-2}, j_1, j_2) + \frac{\tau}{a_{N-1}(j_1, j_2)} (\sigma_{j_1}^2 + \sigma_{j_2}^2)
\]

\[
= d_{N-1}^j(S_{N-2}, j_1, j_2) + a_{N-1}(j_1, j_2).
\]

Thus Equation (3.26) holds for \( n = 1 \).

Now suppose it is true for arbitrary \( n - 1 \). Again, we must consider the cases of \( p \) even and odd separately. For \( p = 2i - 1 \), the induction hypothesis implies that,

\[
d_2^{2i-1+2^i} = g_2^{j_1} \left( d_{N-n+1}^{2^i}(S_{N-n}, j_{2,n+1}) \right)
\]

\[
= g_2^{j_1} \left( d_{N-n+1}^{2^i}(S_{N-n}, j_{2,n+1}) + a_{N-n+1}(j_{2,n+1}) \right).
\]

Applying Lemma 3.3.6 gives,

\[
d_2^{2i-1+2^i} = g_2^{j_1} \left( d_{N-n+1}^{2^i}(S_{N-n}, j_{2,n+1}) + \frac{a_{N-n+1}(j_{2,n+1})^2}{a_{N-n}(j_{1,n+1})} \right).
\]

As,

\[
g_2^{j_1} \left( d_{N-n+1}^{2^i}(S_{N-n}, j_{2,n+1}) \right) = g_1^{j_1} \left( d_{N-n+1}^{2^i}(S_{N-n}, j_{2,n+1}) \right) + \frac{\sigma_{j_1}^2 \tau}{a_{N-n}(j_{1,n+1})}
\]

\[
= d_2^{2i-1}(S_{N-n-1}, j_{1,n+1}) + \frac{\sigma_{j_1}^2 \tau}{a_{N-n}(j_{1,n+1})},
\]

the recursive definition of \( a_{N-n} \) gives Equation (3.26) for \( p \) odd.

Following the same procedure for \( p = 2i \) completes the proof. \( \square \)

### 3.4 Option Pricing

Recall from Section 3.2 that the mean-variance price of a call option, \( H \), with strike price, \( K \), on an asset following the Switching Black–Scholes model is,

\[
H_0 = \frac{e^{-rT}}{D_0} F_0,
\]
where \( F_N = (S_N - K)^+ \) and \( F_{n-1} = \mathbb{E}[F_n(1 - \beta_n \Delta X_n) | \mathcal{F}_{n-1}] \). In this section, we define two functions, \( Q_{N-n,m}^k \) and \( R_{N-n,m}^k \), and show that,

\[
F_0 = \sum_{j_1, \ldots, j_N = 1}^M f(j_1, \ldots, j_N) \left( S_0 R_{1,N-1}^1(S_0, Z_0, j_1, N) - K Q_{1,N-1}^1(S_0, Z_0, j_1, N) \right),
\]

for a constant \( f(j_1, \ldots, j_N) \). Here, the functions \( Q_{1,N-1}^1 \) and \( R_{1,N-1}^1 \) correspond to \( N(d_1) \) and \( N(d_2) \) of the Black-Scholes pricing formula. We will see that,

\[
Q_{1,N-1}^1(S_0, Z_0, j_1, N) = \sum_{p=1}^{2^N} c_p N \left( d_1^p(S_0, j_1, N) - x_p \frac{\sigma_{j_1}^2 \tau}{a_1(j_1, N)} \right),
\]

(3.30)

and,

\[
R_{1,N-1}^1(S_0, Z_0, j_1, N) = \sum_{p=1}^{2^N} c'_p N \left( d_1^{p+2^{N-1}}(S_0, j_1, N) - x_p \frac{\sigma_{j_1}^2 \tau}{a_1(j_1, N)} \right),
\]

(3.31)

where the \( d_1^p \) and \( a_1 \) values are as defined in Section 3.3, \( x_p \in \{0, 1\} \), and \( c_p, c'_p \) are constants.

We begin with some notation.

**Notation 3.4.1.** Let \( i \in \{1, 2, \ldots, M\} \) and \( n \in \{1, 2, \ldots, N\} \). Define exponentials,

\[
J_i = e^{\mu_i \tau}, \quad L_i = e^{(2\mu_i + \sigma_i^2) \tau}.
\]

As the adjustment process \( \beta_n \) is predictable, we may write \( \beta_n = \beta_n(S_{n-1}, Z_{n-1}) \).

From Theorem 3.2.4 we have,

\[
\beta_n(s, e_i) = e^{\mu_n \tau} s f_n(e_i),
\]

where,

\[
f_n(e_i) = \frac{G_n(e_i)}{H_n(e_i)},
\]

\[
G_n(e_i) = \sum_{j=1}^M D_n(e_j) B_{ij} \quad \text{and} \quad H_n(e_i) = \sum_{j=1}^M D_n(e_j) C_{ij}.
\]
Also define,

\[ M_n(e_i) = 1 + e^{\gamma} f_n(e_i). \]

With this notation, we see that for \( s_{n-1} = s \) and \( z_{n-1} = e_i \) we have,

\[ 1 - \beta_n(s, e_i) \Delta X_n = M_n(e_i) - \frac{s_n}{s} f_n(e_i), \]

(3.32)

a fact which is used throughout both this and the next chapter.

We are now in a position to define the functions \( Q_{N-n,m}^k \) and \( R_{N-n,m}^k \) recursively.

**Definition 3.4.2.** Let \( n \in \{0, 1, \ldots, N-1\} \). For \( k = 1, 2, \ldots, 2^n \) and \( j_0, j_1, \ldots, j_{n+1} \in \{1, 2, \ldots, M\} \), define,

\[
Q_{N-n,0}^k(S_{N-n-1}, e_{j_0}, j_{1,n+1}) = M_N(e_{j_n}) N \left( d_{N-n}^k(S_{N-n-1}, j_{1,n+1}) - \frac{\sigma_{j_{n+1}}^2}{a_{N-n}(j_{1,n+1})} \right) - f_N(e_{j_n}) J_{j_{n+1}} N \left( d_{N-n}^k(S_{N-n-1}, j_{1,n+1}) \right),
\]

and,

\[
R_{N-n,0}^k(S_{N-n-1}, e_{j_0}, j_{1,n+1}) = M_N(e_{j_n}) J_{j_{n+1}} N \left( d_{N-n}^{k+2^n}(S_{N-n-1}, j_{1,n+1}) - \frac{\sigma_{j_{n+1}}^2}{a_{N-n}(j_{1,n+1})} \right) - f_N(e_{j_n}) L_{j_{n+1}} N \left( d_{N-n}^{k+2^n}(S_{N-n-1}, j_{1,n+1}) \right).
\]

Here, the \( d_{N-n}^k \) and \( a_{N-n} \) values are as defined in Section 3.3.

Also, for \( m = 1, 2, \ldots, n \), \( k = 1, 2, \ldots, 2^n-m \), define,

\[
Q_{N-n,m}^k(S_{N-n-1}, e_{j_0}, j_{1,n+1}) = M_{N-m}(e_{j_{m-1}}) Q_{N-n,m-1}^k(S_{N-n-1}, e_{j_0}, j_{1,n+1}) - f_{N-m}(e_{j_{m-1}}) J_{j_{n+1}} Q_{N-n,m}^{k+2^n-m}(S_{N-n-1}, e_{j_0}, j_{1,n+1}),
\]
and,

\[
R_{n-m}^k(S_{N-n-1}, e_{j_n}, j_{1,n+1})
= M_{n-m}(e_{j_{n-m}})J_{j_{n-m}+1}R_{n-m-1}^k(S_{N-n-1}, e_{j_0}, j_{1,n+1})
- f_{n-m}(e_{j_{n-m}})L_{j_{n-m}+1}R_{n-m-1}^{k+2m-n}(S_{N-n-1}, e_{j_0}, j_{1,n+1}).
\]

Note that \(Q_{N-n}^1\) and \(R_{N-n}^k\) are \(\mathcal{F}_{n-1}\)-measurable random variables.

In order to find the mean-variance price of a call option in a Switching Black–Scholes market, we show that,

\[
F_{n-n} = \sum_{j_1, \ldots, j_n=1}^{M} f(j_1, \ldots, j_n) \left( S_{N-n}R_{n+1,n-1}^1 - K Q_{n+1,n-1}^1 \right),
\]

where \(f(j_1, \ldots, j_n) = A_{j_1} z_{N-n} A_{j_2} \cdots A_{j_n} j_{n-1} \). The proof is by induction on \(n\).

As we defined,

\[
F_{n-1} = E[F_n(1 - \beta_n \Delta X_n) | \mathcal{F}_{n-1}],
\]

we will need to evaluate conditional expectations of the form,

\[
E[A_j z_{N-n} (1 - \beta_{N-n} \Delta X_{N-n}) Q_{n+1,n-1}^1(S_{N-n}, Z_{N-n}, j_2, j_{n+1}) | \mathcal{F}_{n-1}],
\]

and

\[
E[A_j z_{N-n} (1 - \beta_{N-n} \Delta X_{N-n}) S_{N-n} R_{n+1,n-1}^1(S_{N-n}, Z_{N-n}, j_2, j_{n+1}) | \mathcal{F}_{n-1}].
\]

We concentrate in this section on the \(Q\) component of the expression. Analogous calculations involving the \(R\) component are given in Appendix 2.

To evaluate the expectation (3.33), we first write \(Q_{n+1,n-1}^1\) as a sum of the \(Q_{n+1,n-1,0}^1\)'s. The following definition gives the coefficients needed.

**Definition 3.4.3.** For \(n \in \{1, 2, \ldots, N - 1\}\), we recursively define functions \(q_p^n : \{1, 2, \ldots, M\}^{n+1} \to \mathbb{R}\) for \(p = 1, 2, \ldots, 2^n\) by setting,

\[
q_1^1(j_0, j_1) = M_{n-1}(e_{j_0}),
q_2^1(j_0, j_1) = -f_{n-1}(e_{j_0}) J_{j_1},
\]
and,

\[
 q_p^n(j_0, j_1, \ldots, j_n) = \begin{cases} 
 M_{N-n}(e_{j_0}) q_{p+1}^{n-1}(j_1, j_2, \ldots, j_n), & \text{if } p \text{ odd,} \\
 -f_{N-n}(e_{j_0}) J_{j_1} q_{p/2}^{n-1}(j_1, j_2, \ldots, j_n), & \text{if } p \text{ even.} 
\end{cases}
\]

Here, \( M, f \) and \( J \) are as defined in 3.4.1.

With this definition, we are able to write \( Q_{N-n+1,n-1}^k \) as a sum of the \( Q_{N-n+1,0}^k \)'s, as shown in the following theorem.

**Theorem 3.4.4.** Let \( n \in \{2, 3, \ldots, N\} \). For \( m = 1, 2, \ldots, n-1 \) and \( k = 1, 2, \ldots, 2^{n-1}-m \), the function \( Q_{N-n+1,m}^k \) defined in 3.4.2 satisfies,

\[
 Q_{N-n+1,m}(S_{N-n}, e_{j_0}, j_{1,n}) = \sum_{p=1}^{2^m} q_p^m(j_{n-1-m}, \ldots, j_{n-1}) Q_{N-n+1,0}^{k+\sum_{i=0}^{m-1} p_i 2^{n-1-m+i}}(S_{N-n}, e_{j_0}, j_{1,n}),
\]  
where the constants \( p_i \in \{0, 1\} \) come from the binary representation of \( p-1 \),

\[
p - 1 = \sum_{i=0}^{m-1} p_i 2^i.
\]

**Proof.** Let \( n \) be fixed. The proof is by induction on \( m \). Taking \( m = 1 \), the definition gives,

\[
 Q_{N-n+1,1}(S_{N-n}, e_{j_0}, j_{1,n}) = M_{N-1}(e_{j_{n-1}}) Q_{N-n+1,0}^k(S_{N-n}, e_{j_0}, j_{1,n})
\]

\[
 - f_{N-1}(e_{j_{n-1}}) J_{j_{n-1}} Q_{N-n+1,0}^{k+2^{n-2}}(S_{N-n}, e_{j_0}, j_{1,n}).
\]

From Definition 3.4.3, this is equal to the sum (3.34) when \( m = 1 \).

Now suppose the result is true for arbitrary \( m \). The definition gives,

\[
 Q_{N-n+1,m+1}(S_{N-n}, e_{j_0}, j_{1,n}) = M_{N-m-1}(e_{j_{n-1-m-1}}) Q_{N-n+1,m}^k(S_{N-n}, e_{j_0}, j_{1,n})
\]

\[
 - f_{N-m-1}(e_{j_{n-1-m-1}}) J_{j_{n-1-m-1}} Q_{N-n+1,m}^{k+2^{n-m-2}}(S_{N-n}, e_{j_0}, j_{1,n}).
\]

By the induction hypothesis, this is equal to,

\[
 \sum_{p=1}^{2^m} q_p^m(j_{n-1-m}, \ldots, j_{n-1}) \times
\]

\[
 \left\{ M_{N-m-1}(e_{j_{n-2-m}}) Q_{N-n+1,0}^{k+\sum_{i=0}^{m-1} p_i 2^{n-1-m+i}}(S_{N-n}, e_{j_0}, j_{1,n})
\]

\[
 - f_{N-m-1}(e_{j_{n-2-m}}) J_{j_{n-2-m}} Q_{N-n+1,0}^{k+2^{n-m-2}+\sum_{i=0}^{m-1} p_i 2^{n-1-m+i}}(S_{N-n}, e_{j_0}, j_{1,n}) \right\},
\]

Theorem 3.4.4.
where \( p - 1 = \sum_{i=0}^{m-1} p_i 2^i \). Using Definition 3.4.3 we have,

\[
\sum_{p=1}^{2^m} \left\{ \sum_{i=0}^{m-1} p_i 2^{n-1-m+i} \right\} (S_{N-n}, e_{j_0}, j_{1,n}) \\
+ \sum_{p=1}^{2^m} (j_{n-2-m}, \ldots, j_{n-1}) \sum_{i=0}^{m-1} p_i 2^{n-1-m+i} (S_{N-n}, e_{j_0}, j_{1,n}) \right\}.
\]

To obtain the representation (3.34), we now need to find \( p'_i \) and \( p''_i \),

\[
i = 0, 1, \ldots, m, \text{ satisfying,}
\]

\[
(2p - 1) - 1 = \sum_{i=0}^{m} p'_i 2^i,
\]

\[
2p - 1 = \sum_{i=0}^{m} p''_i 2^i.
\]

Lemma A.1.4 of Appendix 1 shows that these values are \( p'_0 = 0, p''_0 = 1 \) and \( p''_i = p'_i = p_{i-1} \) for \( i \geq 1 \). We then have,

\[
\sum_{i=0}^{m} p_i 2^{n-1-m+i} = \sum_{i=1}^{m} p_{i-1} 2^{n-2-m+i} = \sum_{i=0}^{m} p'_i 2^{n-2-m+i},
\]

and,

\[
2^{n-2-m} + \sum_{i=0}^{m-1} p_i 2^{n-1-m+i} = \sum_{i=0}^{m} p''_i 2^{n-2-m+i}.
\]

Using these calculations, we obtain,

\[
\mathcal{Q}_{N-n+1,m+1}^k (S_{N-n}, e_{j_0}, j_{1,n}) \\
= \sum_{p=1}^{2^{m+1}} q_{p}^{m+1} (j_{n-2-m}, \ldots, j_{n-1}) \mathcal{Q}_{N-n+1,0}^{k+\sum_{i=0}^{m} p_i 2^{n-2-m+i}} (S_{N-n}, e_{j_0}, j_{1,n}),
\]

where, as required, \( p - 1 = \sum_{i=0}^{m} p_i 2^i \).

In particular, when \( m = n - 1 \) and \( k = 1 \), this theorem gives,

\[
\mathcal{Q}_{N-n+1,n-1}^1 (S_{N-n}, Z_{N-n}, j_{1,n}) \\
= \sum_{p=1}^{2^{n-1}} q_{p}^{n-1} (Z_{N-n}, j_1, \ldots, j_{n-1}) \mathcal{Q}_{N-n+1,0}^{p}(S_{N-n}, Z_{N-n}, j_{1,n}), \quad (3.35)
\]
for \( n = 2, 3, \ldots, N \), as \( 1 + \sum_{i=0}^{n-2} \kappa_i 2^i = 1 + (p - 1) = p \). Thus to calculate the expectation (3.33), we need to compute conditional expectations of the form,

\[
E\left[ g(\hat{Z}_{N-n}^{p})(1 - \beta_{N-n} \Delta X_{N-n}) Q_{N-n+1,0}^p(S_{N-n}, \hat{Z}_{N-n}, \hat{j}_{2n+1}) \bigg| \mathcal{F}_{N-n-1} \right],
\]

for functions \( g : \mathbb{R}^M \to \mathbb{R} \). This is achieved using the following lemma and corollary.

**Lemma 3.4.5.** Let \( n \in \{1, 2, \ldots, N\} \) and \( \beta_n \) be as defined in 3.2.1. For any function \( g : \mathbb{R}^M \to \mathbb{R} \) and any real numbers \( \alpha, \delta \) and \( \gamma \), we have,

\[
E\left[ g(Z_n) (1 - \beta_n \Delta X_n) S_{n-1}^N \left( \frac{\log S_n + \delta}{\alpha} \right) \bigg| S_{n-1} = s, Z_{n-1} = e_i \right] = \sum_{j=1}^{M} A_{ji} g(e_j) s^\gamma \left\{ \exp \left( \gamma (\mu_j - \frac{1}{2} \sigma_j^2) + \frac{1}{2} \gamma^2 \sigma_j^2 \tau \right) M_n(e_j) N(d_1) \right. \\
\left. - \exp \left( (\gamma + 1)(\mu_j - \frac{1}{2} \sigma_j^2) + \frac{1}{2} (\gamma + 1)^2 \sigma_j^2 \tau \right) f_n(e_j) N(d_2) \right\},
\]

where,

\[
d = \frac{\log s + (\mu_j - \frac{1}{2} \sigma_j^2) \tau + \delta}{\sqrt{\sigma_j^2 \tau + \alpha^2}},
\]

\[
d_1 = d + \frac{\gamma \sigma_j^2 \tau}{\sqrt{\sigma_j^2 \tau + \alpha^2}},
\]

\[
d_2 = d + \frac{(\gamma + 1) \sigma_j^2 \tau}{\sqrt{\sigma_j^2 \tau + \alpha^2}},
\]

and \( N \) is the Normal distribution function.

**Proof.** From Equation (3.32) on page 48, \( 1 - \beta_n \Delta X_n = M_n(e_i) - \frac{S_n}{s} f_n(e_i) \), so the conditional expectation becomes,

\[
M_n(e_i) E\left[ g(Z_n) S_{n-1}^N \left( \frac{\log S_n + \delta}{\alpha} \right) \bigg| S_{n-1} = s, Z_{n-1} = e_i \right] - \frac{f_n(e_i)}{s} E\left[ g(Z_n) S_{n+1}^N \left( \frac{\log S_n + \delta}{\alpha} \right) \bigg| S_{n-1} = s, Z_{n-1} = e_i \right].
\]

Using Lemma 2.3.2, the first conditional expectation is,

\[
\sum_{j=1}^{M} A_{ji} g(e_j) s^\gamma \exp \left( \gamma (\mu_j - \frac{1}{2} \sigma_j^2) + \frac{1}{2} \gamma^2 \sigma_j^2 \tau \right) N\left( d + \frac{\gamma \sigma_j^2 \tau}{\sqrt{\sigma_j^2 \tau + \alpha^2}} \right),
\]
where,
\[
d = \log s + (\mu_j - \frac{1}{2}\sigma_j^2\tau + \delta \sqrt{\sigma_j^2\tau + \alpha^2}),
\]
and the second expectation is,
\[
\sum_{j=1}^{M} A_{ji}g(e_j)s^{(\gamma+1)} \exp \left( (\gamma+1)(\mu_j - \frac{1}{2}\sigma_j^2\tau + \frac{1}{2}(\gamma+1)^2\sigma_j^2\tau) \right) 
\times N \left( d + \frac{(\gamma+1)\sigma_j^2\tau}{\sqrt{\sigma_j^2\tau + \alpha^2}} \right).
\]

Combining the two sums gives,
\[
\sum_{j=1}^{M} A_{ji}g(e_j)s^{\gamma} \left\{ \exp \left( \gamma(\mu_j - \frac{1}{2}\sigma_j^2\tau + \frac{1}{2}\gamma^2 \sigma_j^2\tau) \right) M_n(e_i)N(d_1) 
- \exp \left( (\gamma+1)(\mu_j - \frac{1}{2}\sigma_j^2\tau + \frac{1}{2}(\gamma+1)^2\sigma_j^2\tau) \right) f_n(e_i)N(d_2) \right\},
\]
as required. \qed

In particular, two special cases of this lemma are needed.

**Corollary 3.4.6.** Let \( n \in \{1, 2, \ldots, N\} \) and \( d = \frac{\log S_0 + \delta}{\alpha} \) for some \( \alpha, \delta \in \mathbb{R} \).

For any function \( g : \mathbb{R}^M \to \mathbb{R} \) we have,
\[
E[g(Z_n) (1 - \beta_n \Delta X_n) N(d) \mid S_{n-1} = s, Z_{n-1} = e_i] 
= \sum_{j=1}^{M} A_{ji}g(e_j) \left[ M_n(e_i)N(g^j_1(d)) - J_j f_n(e_i)N(g^j_2(d)) \right],
\]
and,
\[
E[g(Z_n) (1 - \beta_n \Delta X_n) S_n N(d) \mid S_{n-1} = s, Z_{n-1} = e_i] 
= \sum_{j=1}^{M} A_{ji}g(e_j)s \left[ J_j M_n(e_i)N(g^j_3(d)) - L_j f_n(e_i)N(g^j_4(d)) \right],
\]
where the functions \( g^j_k \) are defined in 3.3.1, \( \beta_n \) is defined in 3.2.1, and \( N \) is the Normal distribution function.
Proof. Put $\gamma = 0$ and $\gamma = 1$ in Lemma 3.4.5, and consult the definitions of the $g_{m}^{n}$ in Equation (3.19).

We can now evaluate the conditional expectation (3.33), needed to calculate $F_{N-n}$, for $n = 1, 2, \ldots, N$, and hence mean-variance call option prices in a Switching Black–Scholes market.

**Theorem 3.4.7.** For $n = 1, 2, \ldots, N - 1$ and $\beta_{N-n}$ as defined in 3.2.1, the function $Q_{N-n+1,n-1}^{1}$ defined in 3.4.2 satisfies the following equation,

$$E[A_{j_{2}Z_{N-n}}(1 - \beta_{N-n}\Delta X_{N-n})Q_{N-n+1,n-1}^{1}(S_{N-n}, Z_{N-n}, j_{2,n+1})|F_{N-n-1}] = \sum_{j_{1}=1}^{M} A_{j_{1}Z_{N-n-1}}A_{j_{2}j_{1}} Q_{N-n,n}^{1}(S_{N-n-1}, Z_{N-n-1}, j_{1,n+1}).$$

**Proof.** For $n \geq 2$, we use the representation of $Q_{N-n+1,n-1}^{1}$ given in Equation (3.35) to express the conditional expectation as,

$$\sum_{p=1}^{2^{n-1}} E\left[A_{j_{2}Z_{N-n}}(1 - \beta_{N-n}\Delta X_{N-n}) q_{p}^{n-1}(Z_{N-n}, j_{2}, \ldots, j_{n}) \times Q_{N-n+1,0}^{p}(S_{N-n}, Z_{N-n}, j_{2,n+1})|F_{N-n-1}\right].$$

For the case $n = 1$, the result is obtained using the same method as below up until step (3.38), from which the result is immediate. We do not include it separately, as the steps are analogous.

Continuing for the case $n \geq 2$, the definition of $Q_{N-n+1,0}^{p}(S_{N-n}, Z_{N-n}, j_{2,n+1})$ allows us to write (3.36) as,

$$\sum_{p=1}^{2^{n-1}} E\left[A_{j_{2}Z_{N-n}}(1 - \beta_{N-n}\Delta X_{N-n}) q_{p}^{n-1}(Z_{N-n}, j_{2}, \ldots, j_{n}) \times \left\{ M_{N}(\epsilon_{j_{n}})N\left(d_{N-n+1}^{p}(S_{N-n}, j_{2,n+1}) - \frac{\sigma_{j_{n+1}}^{2} \tau}{a_{N-n+1}(j_{2,n+1})}\right) - f_{N}(\epsilon_{j_{n}})J_{j_{n+1}}N\left(d_{N-n+1}^{p}(S_{N-n}, j_{2,n+1})\right)\right\}|F_{N-n-1}\right].$$
From Lemma 3.3.5, $d_{N-n+1}^p$ can be written as $\frac{\log S_{N-n} + \delta}{\alpha}$ for some $\alpha, \delta \in \mathbb{R}$, so we may apply Corollary 3.4.6 to give,

$$
\sum_{p=1}^{2^n-1} \sum_{j_1=1}^{M} A_{j_1} z_{N-n-1}^j A_{j_2, j_1} q_{n-1}^j (j_1, j_2, \ldots, j_n) \left[ M_N(e_{j_n}) \left\{ M_N(z_{N-n-1}) N \left( g_1^{j_1} \left( d_{N-n+1}^p - \frac{\sigma_{j_{n+1}}^2 \tau}{a_{N-n+1}(j_{2,n+1})} \right) \right) 
- f_{N-n}(z_{N-n-1}) J_{j_{n+1}} N \left( g_2^{j_1} \left( d_{N-n+1}^p - \frac{\sigma_{j_{n+1}}^2 \tau}{a_{N-n+1}(j_{2,n+1})} \right) \right) \right\} 
- f_{N-n}(z_{N-n-1}) J_{j_{n+1}} N \left( g_2^{j_1} \left( d_{N-n+1}^p (S_{N-n}, j_{2,n+1}) \right) \right) 
- f_{N-n}(z_{N-n-1}) J_{j_{n+1}} N \left( g_2^{j_1} \left( d_{N-n+1}^p (S_{N-n}, j_{2,n+1}) \right) \right) \right] \right]. \tag{3.37}
$$

From Definition 3.3.4,

$$
g_1^{j_1} \left( d_{N-n+1}^p (S_{N-n}, j_{2,n+1}) \right) = d_{N-n}^{p-1}(S_{N-n-1}, j_{1,n+1}),
$$

and,

$$
g_2^{j_1} \left( d_{N-n+1}^p (S_{N-n}, j_{2,n+1}) \right) = d_{N-n}^{p}(S_{N-n-1}, j_{1,n+1}),
$$

so rearranging (3.37) and using Lemma 3.3.6 gives,

$$
\sum_{j_1=1}^{M} \sum_{p=1}^{2^n-1} A_{j_1} z_{N-n-1}^j A_{j_2, j_1} q_{n-1}^j (j_1, j_2, \ldots, j_n) \left\{ M_N(z_{N-n-1}) \left[ M_N(e_{j_n}) N \left( d_{N-n}^{p-1} - \frac{\sigma_{j_{n+1}}^2 \tau}{a_{N-n}(j_{1,n+1})} \right) 
- f_{N}(e_{j_n}) J_{j_{n+1}} N \left( d_{N-n}^{p-1}(S_{N-n-1}, j_{1,n+1}) \right) \right] 
- f_{N-n}(z_{N-n-1}) J_{j_{n+1}} N \left( d_{N-n}^{p} - \frac{\sigma_{j_{n+1}}^2 \tau}{a_{N-n}(j_{1,n+1})} \right) \right] 
- f_{N-n}(z_{N-n-1}) J_{j_{n+1}} N \left( d_{N-n}^{p} (S_{N-n-1}, j_{1,n+1}) \right) \right\}. 
$$
By the definition of $Q_{N-n,0}^k(S_{N-n-1}, Z_{N-n-1}, j_{1,n+1})$, this is,

$$
\sum_{j_1=1}^{M} A_{j_1} z_{N-n-1} \sum_{p=1}^{2^{n-1}} q_{p}^{n-1}(j_1, j_2, \ldots, j_n) \prod_{j=2}^{n} A_{j} z_{N-n-1} \sum_{p=1}^{2^{n-1}} q_{p}^{n-1}(j_1, j_2, \ldots, j_n) \\
\times \left\{ M_{N-n}(Z_{N-n-1}) Q_{N-n,0}^{2^{n-1}}(S_{N-n-1}, Z_{N-n-1}, j_{1,n+1}) \\
- f_{N-n}(Z_{N-n-1}) J_{j_1} Q_{N-n,0}^{2^{n-1}}(S_{N-n-1}, Z_{N-n-1}, j_{1,n+1}) \right\}. \ (3.38)
$$

Also, the definition,

$$
q_{p}^{n}(Z_{N-n-1}, j_1, \ldots, j_n) = \begin{cases} 
M_{N-n}(Z_{N-n-1}) q_{[p+1]/2}^{n-1}(j_1, j_2, \ldots, j_n), & p \text{ odd,} \\
- f_{N-n}(Z_{N-n-1}) J_{j_1} q_{[p/2]}^{n-1}(j_1, j_2, \ldots, j_n), & p \text{ even,}
\end{cases}
$$

gives,

$$
\sum_{j_1=1}^{M} A_{j_1} z_{N-n-1} \sum_{p=1}^{2^{n}} q_{p}^{n}(Z_{N-n-1}, j_1, \ldots, j_n) Q_{N-n,0}^{p}(S_{N-n-1}, Z_{N-n-1}, j_{1,n+1})
$$

and from Equation (3.35) this is equal to,

$$
\sum_{j_1=1}^{M} A_{j_1} z_{N-n-1} A_{j_2,j_1} Q_{N-n,n}^{1}(S_{N-n-1}, Z_{N-n-1}, j_{1,n+1}),
$$

as required. \(\square\)

In an analogous way, it is shown in Theorem A.2.3 of Appendix 2 that,

$$
E[A_{j} z_{N-n} (1 - \beta_{N-n} \Delta X_{N-n}) S_{N-n} R_{N-n,0,1}(S_{N-n}, Z_{N-n}, j_{2,n+1}) | F_{N-n-1}]
= S_{N-n-1} \sum_{j_1=1}^{M} A_{j_1} z_{N-n-1} A_{j_2,j_1} R_{N-n,n,0}^{1}(S_{N-n-1}, Z_{N-n-1}, j_{1,n+1}).
$$

These two results enable us to prove the key theorem of this section.

**Theorem 3.4.8.** For $n = 0, 1, 2, \ldots, N - 1$, the $F_{N-n-1}$-measurable random variable $F_{N-n-1}$ defined in 3.2.5 can be written as a sum of the random variables $Q_{N-n,n}^{1}$ and $R_{N-n,n}^{1}$ (defined in 3.4.2) as follows,

$$
F_{N-n-1}(S_{N-n-1}, Z_{N-n-1}) = \sum_{j_1, \ldots, j_{n+1}=1}^{M} A_{j_1} z_{N-n-1} A_{j_2,j_1} \ldots A_{j_{n+1},j_{n}} \\
\times \left( S_{N-n-1} R_{N-n,n}^{1}(S_{N-n-1}, Z_{N-n-1}, j_{1,n+1}) \\
- K Q_{N-n,n}^{1}(S_{N-n-1}, Z_{N-n-1}, j_{1,n+1}) \right).
$$
Proof. The proof is by induction on $n$. First, we prove the equation for $n = 0$. The value of $F_{N-1}$ is defined to be,

$$F_{N-1}(s, e_i) = E[F_N(1 - \beta N \Delta X_N)|S_{N-1} = s, Z_{N-1} = e_i],$$

where $F_N = (S_N - K)^+$. Using the relation,

$$1 - \beta N(s, e_i) \Delta X_n = M_n(e_i) - \frac{S_n}{s} f_n(e_i),$$

from Equation (3.32) on page 48, and Theorem 2.2.2, which gives the conditional density function of $S_N$ given $S_{N-1} = s$, $Z_{N-1} = e_i$ under the real world probability measure, we have,

$$F_{N-1}(s, e_i) = \sum_{j=1}^{M} \int_{K}^{\infty} (x - K) \left( M_n(e_i) - \frac{x}{s} f_N(e_i) \right) \times

\frac{A_{ji}}{\sqrt{2\pi} \sigma_j \sqrt{\tau}} \exp \left[ \frac{-1}{2} \left( \log \frac{(x/s) - (\mu_j - \frac{1}{2} \sigma_j^2 \tau)}{\sigma_j \sqrt{\tau}} \right)^2 \right] dx.\]$$

Expanding the term $(x - K) \left( M_n(e_i) - \frac{x}{s} f_N(e_i) \right)$ and making the change of variables $z = \frac{\log(x/s) - (\mu_j - \frac{1}{2} \sigma_j^2 \tau)}{\sigma_j \sqrt{\tau}}$, which is equivalent to $x = s e^{(\mu_j - \frac{1}{2} \sigma_j^2 \tau) \gamma + \sigma_j \sqrt{\tau} z}$, gives,

$$\sum_{j=1}^{M} \frac{A_{ji}}{\sqrt{2\pi}} \int_{d}^{\infty} e^{-\frac{1}{2} z^2} \left\{ -KM_N(e_i) + \left( M_N(e_i) + \frac{K}{s} f_N(e_i) \right) s e^{(\mu_j - \frac{1}{2} \sigma_j^2 \tau) \gamma + \sigma_j \sqrt{\tau} z} \right.$$

$$\left. - \frac{1}{s} f_N(e_i) s^2 e^{2(\mu_j - \frac{1}{2} \sigma_j^2 \tau) \gamma + 2\sigma_j \sqrt{\tau} z} \right\} dz, \quad (3.39)$$

where,

$$d = \frac{\log(K/s) - (\mu_j - \frac{1}{2} \sigma_j^2 \tau)}{\sigma_j \sqrt{\tau}} = -d_N(s, j) + \sigma_j \sqrt{\tau}.$$

This integral has three components. The first is equal to $-KM_N(e_i) \sqrt{2\pi} N(-d)$, and the other parts can be calculated using Lemma A.1.3 of Appendix 1, which states that,

$$\int_{p}^{\infty} e^{-\frac{1}{2} y^2 + \beta y} dy = \sqrt{2\pi} e^{\frac{1}{2} \beta^2} N(\beta - p).$$
We then have,

\[
F_{N-1}(s, e_i) = \sum_{j=1}^{M} A_{ji} \left\{ -KM_N(e_i) N \left( d_{k}^1(s, j) - \sigma_j \sqrt{T} \right) + \left( M_N(e_i) + \frac{K}{s} f_N(e_i) \right) se^{\mu_1 T} N \left( d_{k}^1(s, j) \right) - sf_N(e_i)e^{(2\mu_1 + \sigma_j^2)T} N \left( d_{k}^1(s, j) + \sigma_j \sqrt{T} \right) \right\}.
\]

As \( d_{k}^2(s, j) = d_{k}^1(s, j) + \sigma_j \sqrt{T} \) we obtain,

\[
F_{N-1}(s, e_i) = \sum_{j=1}^{M} A_{ji} \left\{ s \left[ M_N(e_i)J_N(d_{k}^2(s, j) - \sigma_j \sqrt{T}) - f_N(e_i)J_N(d_{k}^2(s, j)) \right] - K \left[ M_N(e_i)N(d_{k}^1(s, j) - \sigma_j \sqrt{T}) - f_N(e_i)J_N(d_{k}^1(s, j)) \right] \right\}
\]

\[
= \sum_{j=1}^{M} A_{ji} \left[ sR_{N,0}^1(s, e_i, j) - KQ_{N,0}^1(s, e_i, j) \right],
\]

as required.

Now assume the result is true for arbitrary \( n - 1 \). That is,

\[
F_{N-n}(S_{N-n}, Z_{N-n}) = \sum_{j_1, \ldots, j_{n-1}=1}^{M} A_{j_1}Z_{N-n}A_{j_2,j_1} \cdots A_{j_n,j_{n-1}} \times \left( S_{N-n}R_{N-n+1,n-1}^1(S_{N-n}, Z_{N-n}, j_1,n) - KQ_{N-n+1,n-1}^1(S_{N-n}, Z_{N-n}, j_1,n) \right).
\]

The definition,

\[
F_{N-n-1}(S_{N-n-1}, Z_{N-n-1}) = \mathbb{E}[F_{N-n}(1 - \beta_{N-n}\Delta X_{N-n}) | \mathcal{F}_{N-n-1}],
\]

and the induction hypothesis imply that,

\[
F_{N-n-1}(S_{N-n-1}, Z_{N-n-1}) = \sum_{j_2, \ldots, j_{n+1}=1}^{M} A_{j_2} \cdots A_{j_{n+1},j_{n}} \times \left( E \left[ S_{N-n}A_{j_2}Z_{N-n}R_{N-n+1,n-1}^1(S_{N-n}, Z_{N-n}, j_2,n+1)(1 - \beta_{N-n}\Delta X_{N-n}) | \mathcal{F}_{N-n-1} \right] \right)
\]

\[
- KE \left[ A_{j_2}Z_{N-n}Q_{N-n+1,n-1}^1(S_{N-n}, Z_{N-n}, j_2,n+1)(1 - \beta_{N-n}\Delta X_{N-n}) | \mathcal{F}_{N-n-1} \right].
\]
Using Theorem 3.4.7 for the second expectation, and the analogous result A.2.3 of Appendix 2 for the first expectation, we obtain,

\[
\sum_{j_1, \ldots, j_{n+1}=1}^M A_{j_1} z_{N-n-1} A_{j_2} \cdots A_{j_{n+1}} \times \left( S_{N-n-1}^1 R_{N-n,n}^1(S_{N-n-1}, Z_{N-n-1}, j_{1,n+1}) \right.
\]

\[
- K Q^1_{N-n,n}(S_{N-n-1}, Z_{N-n-1}, j_{1,n+1}),
\]

which completes the proof.

The concluding theorem expresses the mean-variance price of a call option with strike price, \( K \), on an asset following the Switching Black–Scholes model in a way analogous to the Black-Scholes pricing formula.

**Theorem 3.4.9.** The mean-variance price of a call option expiring at time \( T \) with strike price, \( K \), on an asset, \( S \), following the Switching Black–Scholes model, with initial asset price, \( S_0 \), and initial market state, \( e_{j_0} \), is,

\[
H_0(S_0, e_{j_0}) = \frac{e^{-rT}}{D_0(e_{j_0})} \sum_{j_1, \ldots, j_{n+1}=1}^M A_{j_1} A_{j_2} \cdots A_{j_{n+1}} \times \left( S_0 R_{1,N-1}^1(S_0, e_{j_0}, j_{1,n}) - K Q^1_{1,N-1}(S_0, e_{j_0}, j_{1,n}) \right), \tag{3.40}
\]

where \( D_0 \) can be computed using Theorem 3.2.4,

\[
Q^1_{1,N-1}(S_0, e_{j_0}, j_{1,n}) = \sum_{p=1}^{2^{N-1}} q^N_p(j_0, j_1, \ldots, j_{N-1}) Q^p_{1,0}(S_0, e_{j_0}, j_{1,n}),
\]

\[
R^1_{1,N-1}(S_0, e_{j_0}, j_{1,n}) = \sum_{p=1}^{2^{N-1}} r^N_p(j_0, j_1, \ldots, j_{N-1}) R^p_{1,0}(S_0, e_{j_0}, j_{1,n}),
\]

\( q^N_p \) and \( r^N_p \) are defined in 3.4.3 and A.2.1 respectively, and \( Q^p_{1,0}(S_0, e_{j_0}, j_{1,n}) \) and \( R^p_{1,0}(S_0, e_{j_0}, j_{1,n}) \) are defined in 3.4.2.

**Proof.** From Equation (3.17) on page 39, the mean-variance price of a call option in a Switching Black–Scholes market is,

\[
H_0(S_0, e_{j_0}) = \frac{e^{-rT}}{D_0(e_{j_0})} F_0(S_0, e_{j_0}).
\]
The form of \( F_0 \) is given in Theorem 3.4.8, which proves (3.40). The expressions for \( R^1_{1,N-1}(S_0, e_{j_0}, j_{1,N}) \) and \( Q^1_{1,N-1}(S_0, e_{j_0}, j_{1,N}) \) follow from Equations (3.35) and (A2-5).

\[ \square \]

### 3.5 Call Option Price with No Switching

In this section, we consider the special case of the Switching Black–Scholes model when \( M = 1 \), so that the Markov chain, \( Z \), is constant, and the asset, \( S \), evolves according to the Black-Scholes model,

\[
dS(t) = \mu S(t)dt + \sigma S(t)dW(t).
\]

We saw in the last section that the price of a call option with strike price, \( K \), on an asset, \( S \), following the Switching Black–Scholes model, with initial asset price, \( S_0 \), and initial market state, \( e_{j_0} \), is,

\[
H_0(S_0, e_{j_0}) = \frac{e^{-rT}}{D_0(e_{j_0})} \sum_{j_{1,N-1}} A_{j_1}A_{j_2} \cdots A_{j_{N-1}}
\]

\(
\times \left( S_0 R^1_{1,N-1}(S_0, e_{j_0}, j_{1,N}) - K Q^1_{1,N-1}(S_0, e_{j_0}, j_{1,N}) \right),
\tag{3.41}
\)

where,

\[
Q^1_{1,N-1}(S_0, e_{j_0}, j_{1,N}) = \sum_{p=1}^{2^{N-1}} q^N_{p}^{N-1}(j_0, j_1, \ldots, j_{N-1}) Q^p_{1,0}(S_0, e_{j_0}, j_{1,N}),
\]

and,

\[
R^1_{1,N-1}(S_0, e_{j_0}, j_{1,N}) = \sum_{p=1}^{2^{N-1}} r^N_{p}^{N-1}(j_0, j_1, \ldots, j_{N-1}) R^p_{1,0}(S_0, e_{j_0}, j_{1,N}).
\]

We proceed in this section to simplify this formula when there is no switching. Of particular importance is contraction of the sum over \( 2^{N-1} \) terms to a sum over \( N + 1 \) terms, which makes computation of the price much more efficient. Specifically, we will see that the mean-variance price of a call option with strike
price, $K$, on an asset, $S$, following the Switching Black–Scholes model with $M = 1$ is,

$$H_0(S_0) = \sum_{k=0}^{N} \binom{N}{k} a^{N-k} (1-a)^k \left\{ S_0 e^{(\mu-r)T} e^{\sigma^2 r k} N \left( d^+ + k \frac{\sigma \sqrt{T}}{N} \right) - K e^{-rT} N \left( d^- + k \frac{\sigma \sqrt{T}}{N} \right) \right\},$$

where,

$$a = \frac{e^{\sigma^2 r} - e^{(r-\mu)T}}{e^{\sigma^2 r} - 1},$$

$$d^+ = \frac{\log(S_0/K) + (\mu + \frac{1}{2}\sigma^2)T}{\sigma \sqrt{T}},$$

and $d^- = d^+ - \sigma\sqrt{T}$.

We begin by simplifying all the notation which has been introduced during this chapter. In doing so, we use the identities $N\tau = T$ and $\frac{T}{N} = \frac{\sqrt{T}}{\tau}$.

### 3.5.1 The $d$ Values

From Lemma 3.3.3, which gives the form of $a_n$ and $b_n$, we have,

$$a_1 := a_1(j_{1,N}) = \sigma \sqrt{T},$$

and $b_1 := b_1(j_{1,N}) = (N-1)(\mu - \frac{1}{2}\sigma^2)T + (\mu + \frac{1}{2}\sigma^2)T - \log K$

$$= (\mu - \frac{1}{2}\sigma^2)T + \sigma^2 \tau + \log K.$$

The values of $d_1^p$ and $d_1^{p+2^{N-1}}$, for $p = 1, 2, \ldots, 2^{N-1}$, can be computed using Theorem 3.3.7, which gives,

$$d_1^p = d_1^1 + \frac{\sigma \sqrt{T}}{N} \sum_{i=0}^{N-2} p_i,$$

$$d_1^{p+2^{N-1}} = d_1^1 + \sigma \sqrt{T}. \quad (3.42)$$

Here, the constants $p_i \in \{0, 1\}$ come from the representation $p-1 = \sum_{i=0}^{N-2} p_i2^i$.

The value of $d_1^1$, from Theorem 3.3.5, is,

$$d_1^1(S_0) = \frac{\log S_0 + b_1}{a_1}$$

$$= \frac{\log(S_0/K) + (\mu - \frac{1}{2}\sigma^2)T}{\sigma \sqrt{T}} + \frac{\sigma \sqrt{T}}{N}. \quad (3.44)$$
The values of $d^p$, $p = 1, 2, \ldots, 2^N$, given in Equations (3.42) – (3.44) will feature in the price of a call option on an asset following the Switching Black–Scholes model with $M = 1$, as will be seen in Section 3.5.3.

### 3.5.2 The $Q$ and $R$

We now consider simplifications of the formulae (from Theorem 3.4.9),

\[ Q^1_{1,N-1}(S_0) = \sum_{p=1}^{2^{N-1}} q^p_{N-1} Q^p_{1,0}(S_0), \]  

(3.45)

and,

\[ R^1_{1,N-1}(S_0) = \sum_{p=1}^{2^{N-1}} r^p_{N-1} R^p_{1,0}(S_0), \]

when there is no switching. We will see that these sums over $2^{N-1}$ terms can be condensed to sums over $N + 1$ terms, by simplification of the expressions for $q^p_{N-1}$, $r^p_{N-1}$, $Q^p_{1,0}$ and $R^p_{1,0}$. We begin by recalling notation previously introduced.

From Notation 3.2.3 we have,

\[ B := B_{ij} = e^{\mu \tau} - e^{r \tau}, \]

\[ C := C_{ij} = e^{(2\mu + \sigma^2)\tau} - 2e^{(r + \mu)\tau} + e^{2r \tau}. \]

The Notation 3.4.1 at the beginning of Section 3.4 gives,

\[ J := J_j = e^{\mu \tau}, \]

\[ L := L_j = e^{(2\mu + \sigma^2)\tau}, \]

\[ f := f_n(e_j) = \frac{B}{C}, \]

\[ M := M_n(e_j) = 1 + e^{r \tau} f_n(e_j) = 1 + e^{r \tau} \frac{B}{C}. \]

With this notation, the coefficients $q^p_n$ for $n \in \{1, 2, \ldots, N - 1\}$ and for $p = 1, 2, \ldots, 2^n$ are defined recursively (from Definition 3.4.3) by setting

$q^1_1 = M$, $q^1_2 = -f J$, and,

\[ q^p_n = \begin{cases} 
M q^{n-1}_{(p+1)/2}, & p \text{ odd}, \\
-f J q^{n-1}_{p/2}, & p \text{ even}. 
\end{cases} \]
We now prove a lemma which expresses \( q^n_p \) absolutely, and not as a recursion.

**Lemma 3.5.1.** Let \( n \in \{1, 2, \ldots, N - 1\} \). When there is no switching (that is, \( M = 1 \)), the coefficients \( q^n_p \) defined in 3.4.3, for \( p = 1, 2, \ldots, 2^n \), are given by,

\[
q^n_p = M^{n-\sum_{i=0}^{n-1} p_i} (-fJ)^{\sum_{i=0}^{n-1} p_i},
\]

where the constants \( p_i \in \{0, 1\} \) come from the binary representation,

\[
p - 1 = \sum_{i=0}^{n-1} p_i 2^i.
\]

**Proof.** The proof is by induction on \( n \). For \( n = 1 \) we have,

\[
q^1_1 = M = M^{1-0} (-fJ)^0, \\
q^1_2 = -fJ = M^{1-1} (-fJ)^1,
\]

as required.

Suppose the result is true for arbitrary \( n - 1 \), so that for \( k = 1, 2, \ldots, 2^{n-1} \),

\[
q^{n-1}_k = M^{n-1-\sum_{i=0}^{n-2} k_i} (-fJ)^{\sum_{i=0}^{n-2} k_i},
\]

where the constants \( k_i \in \{0, 1\} \) come from the representation \( k - 1 = \sum_{i=0}^{n-2} k_i 2^i \).

As the definition of \( q^n_p \) varies according to whether \( p \) is even or odd, we must consider the two cases individually. First suppose that \( p \) is odd, so that \( p = 2k - 1 \) for some \( k \). Then,

\[
q^n_p = M q^{n-1}_{2k-1} = M^{n-\sum_{i=0}^{n-2} k_i} (-fJ)^{\sum_{i=0}^{n-2} k_i},
\]

and from Lemma A.1.4 of Appendix 1,

\[
p - 1 = (2k - 1) - 1 = \sum_{i=0}^{n-1} p_i 2^i,
\]

where \( p_0 = 0 \) and \( p_i = k_{i-1} \) for \( i \geq 1 \). We see that \( \sum_{i=0}^{n-1} k_i = \sum_{i=0}^{n-1} p_i \), so we have,

\[
q^n_p = M^{n-\sum_{i=0}^{n-1} p_i} (-fJ)^{\sum_{i=0}^{n-1} p_i},
\]

as required.
Similarly, if \( p = 2k \) is even,
\[
d^n_p = -fJd^{n-1}_k = M^{n-1-\sum_{i=0}^{n-2}k_i}(-fJ)^{1+\sum_{i=0}^{n-2}k_i}.
\]

Again, Lemma A.1.4 of Appendix 1 gives,
\[
p - 1 = 2k - 1 = \sum_{i=0}^{n-1} p_i 2^i,
\]
where \( p_0 = 1 \) and \( p_i = k_{i-1} \) for \( i \geq 1 \). Then \( \sum_{i=0}^{n-1} p_i = 1 + \sum_{i=0}^{n-2} k_i \), so we have,
\[
d^n_p = M^{n-\sum_{i=0}^{n-1} p_i} (-fJ)^{\sum_{i=0}^{n-1} p_i},
\]
as required.

With this lemma, we now show how to write \( Q^1_{1,N-1}(S_0) \), given in (3.45) as a sum over \( 2^{N-1} \) terms, as a sum over \( N + 1 \) terms.

**Theorem 3.5.2.** When \( M = 1 \) in the Switching Black–Scholes model, the function \( Q^1_{1,N-1} \) defined in 3.4.2 is given by,
\[
Q^1_{1,N-1}(S_0) = \sum_{k=0}^{N} \binom{N}{k} M^{N-k} (-fJ)^k N \left( d^-(S_0) + k \frac{\sigma \sqrt{T}}{N} \right),
\]
where,
\[
d^-(S_0) = \frac{\log(S_0/K) + (\mu - \frac{1}{2} \sigma^2)T}{\sigma \sqrt{T}}.
\]

**Proof.** From Equation (3.45) we have,
\[
Q^1_{1,N-1}(S_0) = \sum_{p=1}^{2^{N-1}} d^{N-1}_p Q^p_{1,0}(S_0).
\]
The definition of \( Q^p_{1,0} \) given in 3.4.2 and the previous lemma allow us to write this as,
\[
Q^1_{1,N-1}(S_0) = \sum_{p=1}^{2^{N-1}} M^{N-1-\sum_{i=0}^{N-2} p_i} (-fJ)^{\sum_{i=0}^{N-2} p_i} \left( \left\{ MN \left( d^p_1(S_0) - \frac{\sigma \sqrt{T}}{N} \right) \right) - fJN \left( d^p_1(S_0) \right) \right),
\]
(3.46)
where the constants \( p_i \in \{0, 1\} \) come from the representation \( p - 1 = \sum_{i=0}^{N-2} p_i 2^i \).

Also, from Equation (3.42),
\[
d^p_1 = d^1_1 + \frac{\sigma \sqrt{T}}{N} \sum_{i=0}^{N-2} p_i.
\]

Thus all the parts dependent on \( p \) in (3.46) are actually dependent on the sum \( \sum_{i=0}^{N-2} p_i \). As \( p_i \in \{0, 1\} \), and every possible sequence \( (p_0, p_1, \ldots, p_{N-2}) \) is attained, \( \sum_{i=0}^{N-2} p_i \) can take the values 0, 1, 2, \ldots, \( N - 1 \), and \( \sum_{i=0}^{N-2} p_i = k \) is attained in \( \binom{N-1}{k} \) different ways. Therefore we have,
\[
Q_{1,N-1}^1(S_0) = \sum_{k=0}^{N-1} \binom{N-1}{k} M^{N-k}(-f J)^k \times \left\{ MN \left( d^1_1(S_0) + k \frac{\sigma \sqrt{T}}{N} \right) - f J N \left( d^1_1(S_0) + k \frac{\sigma \sqrt{T}}{N} \right) \right\}.
\]

From Equation (3.44),
\[
d^1_1(S_0) = \frac{\log(S_0/K) + (\mu - \frac{1}{2} \sigma^2)T}{\sigma \sqrt{T}} + \frac{\sigma \sqrt{T}}{N},
\]
so we have \( d^1_1(S_0) = d^-(S_0) + \frac{\sigma \sqrt{T}}{N} \) and therefore,
\[
Q_{1,N-1}^1(S_0) = \sum_{k=0}^{N-1} \binom{N-1}{k} M^{N-k}(-f J)^k \times \left\{ M N \left( d^-(S_0) + k \frac{\sigma \sqrt{T}}{N} \right) - f J N \left( d^-(S_0) + k \frac{\sigma \sqrt{T}}{N} \right) \right\}.
\]

As \( \binom{N-1}{k} \) and \( \binom{N-1}{k-1} \) for \( k = 1, 2, \ldots, N - 1 \), we have,
\[
Q_{1,N-1}^1(S_0) = \sum_{k=0}^{N-1} \binom{N}{k} M^{N-k}(-f J)^k \times \left\{ M N \left( d^-(S_0) + k \frac{\sigma \sqrt{T}}{N} \right) - f J N \left( d^-(S_0) + k \frac{\sigma \sqrt{T}}{N} \right) \right\}.
\]

This completes the proof. \( \square \)

Analogous calculations for \( R_{1,N-1}^1 \) are given in Appendix 2. Theorem A.2.5 proves that,
\[
R_{1,N-1}^1(S_0) = \sum_{k=0}^{N} \binom{N}{k} (M J)^{N-k}(-f L)^k \times \left\{ d^+(S_0) + k \frac{\sigma \sqrt{T}}{N} \right\},
\]
where,
\[
d^+(S_0) = \frac{\log(S_0/K) + (\mu + \frac{1}{2} \sigma^2)T}{\sigma\sqrt{T}}.
\]

We are now able to find a simplified expression for the mean-variance price of a call option on an asset following the Switching Black–Scholes model with \( M = 1 \).

### 3.5.3 Call Option Price

From Equation (3.41), the mean-variance price of a call option with strike price, \( K \), on an asset, \( S \), following the Switching Black–Scholes model with \( M = 1 \) is,

\[
H_0(S_0) = \frac{e^{-rT}}{D_0} \left( S_0 R_{1,N-1}^1(S_0) - K Q_{1,N-1}^1(S_0) \right).
\]

The following theorem simplifies this expression, using Theorems 3.5.2 and A.2.5.

**Theorem 3.5.3.** The mean-variance price of a call option with strike price, \( K \), on an asset, \( S \), following the Switching Black–Scholes model with \( M = 1 \) is,

\[
H_0(S_0) = \sum_{k=0}^{N} \binom{N}{k} a^{N-k} (1-a)^k \left\{ S_0 e^{(\mu-r)T} e^{\sigma^2 r k N} \left( d^+ + k \frac{\sigma \sqrt{T}}{N} \right) 

- K e^{-rT} N \left( d^- + k \frac{\sigma \sqrt{T}}{N} \right) \right\},
\]

where,
\[
a = \frac{e^{\sigma^2 \tau} - e^{(r-\mu)\tau}}{e^{\sigma^2 \tau} - 1},
\]
\[
d^+ = \frac{\log(S_0/K) + (\mu + \frac{1}{2} \sigma^2)T}{\sigma\sqrt{T}},
\]
and \( d^- = d^+ - \sigma\sqrt{T} \).

**Proof.** First, we find \( D_0 \) when \( M = 1 \). Equation (3.16) on page 35 gives,

\[
D_{n-1} = D_n - \frac{(D_n B)^2}{D_n C} = D_n \left( 1 - \frac{B^2}{C} \right),
\]
for \( n = 1, 2, \ldots, N \), which implies, using \( D_N = 1 \), that,

\[
D_0 = \left( 1 - \frac{\mu^2}{C} \right)^N.
\]

With the expressions \( B = e^{\mu r} - e^{rT} \) and \( C = e^{(2\mu + \sigma^2) r} - 2e^{(\mu + \mu) r} + e^{2rT} \), we obtain,

\[
D_0 = \frac{e^{2\mu T}}{C^N} \left( e^{\sigma^2 T} - 1 \right)^N.
\]

Next, we have from Theorem 3.5.2,

\[
Q_{1,N-1}^1(S_0) = \sum_{k=0}^{N} \binom{N}{k} M^{N-k} (-fJ)^k N \left( d^- + k \frac{\sigma \sqrt{T}}{N} \right).
\]

Writing \( M \) and \(-fJ\) in terms of basic parameters, from page 62 we have,

\[
-fJ = -\frac{B}{C} e^{\mu r} = \frac{\mu}{C} \left( e^{rT} - e^{\mu r} \right),
\]

and,

\[
M = 1 + e^{\mu r} \frac{B}{C} = \frac{\mu}{C} \left( e^{(\mu + \sigma^2) r} - e^{\mu r} \right).
\]

Also, from Theorem A.2.5,

\[
R_{1,N-1}^1(S_0) = \sum_{k=0}^{N} \binom{N}{k} (MJ)^{N-k} (-fL)^k N \left( d^+ + k \frac{\sigma \sqrt{T}}{N} \right).
\]

Using the identities \( L = e^{(\mu + \sigma^2) r} J \) and \( J = e^{\mu r} \), the price,

\[
H_0(S_0) = \frac{e^{-rT}}{D_0} \left( S_0 R_{1,N-1}^1(S_0) - K Q_{1,N-1}^1(S_0) \right),
\]

then becomes,

\[
H_0(S_0) = e^{-(r + \mu) T} \times
\left\{ S_0 \sum_{k=0}^{N} \binom{N}{k} \left( e^{(\mu + \sigma^2) r} - e^{rT} \right)^{N-k} \frac{e^{rT} - e^{\mu r}}{e^{\sigma^2 T} - 1} \frac{e^{\mu T} e^{\sigma^2 T} k N \left( d^+ + k \frac{\sigma \sqrt{T}}{N} \right)}{e^{\sigma^2 T} - 1} \right.
\]

\[
- K \sum_{k=0}^{N} \binom{N}{k} \left( e^{(\mu + \sigma^2) r} - e^{rT} \right)^{N-k} \frac{e^{rT} - e^{\mu r}}{e^{\sigma^2 T} - 1} \frac{e^{\mu T} e^{\sigma^2 T} k N \left( d^- + k \frac{\sigma \sqrt{T}}{N} \right)}{e^{\sigma^2 T} - 1} \}.
\]

Setting,

\[
a = \frac{e^{\sigma^2 T} - e^{(\tau - \mu) r}}{e^{\sigma^2 T} - 1},
\]

completes the proof. \( \square \)
We would expect that as the number of switches, $N$, tends to infinity, the mean-variance call option price would approximate the Black-Scholes call price. This would be an interesting topic for further research.
Chapter 4

Option Pricing via Mean–Variance Hedging - Part 2

This chapter, a continuation of Chapter 3, further investigates the use of mean-variance hedging for pricing call options on assets following the Switching Black–Scholes model. In Chapter 3, we found the mean-variance price of such call options using a backwardisation procedure. This was developed by expressing the price of a contingent claim, $H$, as,

$$H_0 = \mathbb{E}^Q[H e^{-rT} | \mathcal{F}_0],$$

where $Q$ is the $b$-variance-optimal signed martingale measure, and repeatedly conditioning backwards in time on $\sigma$-algebras of the filtration. In this chapter, rather than using this direct method to find the price, we instead examine the joint conditional density functions of $(S_n, Z_n)$ given $\mathcal{F}_m$ under the martingale measure, $Q$.

We begin in Section 4.1 by recursively defining some functions, $E_n$ and $K_n$ (for $n = 1, 2, \ldots, N$), and proving some results concerning these values. These results are applied in Section 4.2, where we find the joint conditional density function, $\phi_{n,m}$, of $(S_n, Z_n)$ given $S_m = s$ and $Z_m = c$, under the measure, $Q$, for all $n > m$. The density function $\phi_{n,m}$ is found using a forward recursive procedure, with initial density function $\phi_{m+1,m}$. Bayes rule and the real world
joint conditional density function of \((S_{m+1}, Z_{m+1})\) given \(\mathcal{F}_m\) are used to find the initial function, \(\phi_{m+1,m}\). To ensure validity of Bayes rule, throughout this chapter we assume that \(Q\) is a probability measure, rather than only a signed measure. Finding conditions under which this is true is left as a topic for further investigation.

Given these density functions, the mean-variance price at time \(m\) of an \(\mathcal{F}_m\)-measurable contingent claim, \(H_n\), is,

\[
H_m(s, e_i) = e^{-r(n-m)} \sum_{j=1}^{M} \int_{-\infty}^{\infty} \phi_{n,m}(s, e_i, x, e_j) H_n(x, e_j) \, dx,
\]

where the initial values \(S_m = s\) and \(Z_m = e_i\) are known. In Section 4.3, we use this relationship to find the mean-variance price of a call option on an asset following the Switching Black–Scholes model.

We begin by giving some definitions which will be needed.

### 4.1 Preliminary Definitions and Results

In this section, we define two real-valued functions, \(E^p_n\) and \(K_n\), on the domain \(\{1, 2, \ldots, M\}^n\), for \(n = 1, 2, \ldots, N\) and \(p = 1, 2, \ldots, 2^n\). We will see in Theorem 4.2.3 of Section 4.2 that \(\phi_n\), the joint conditional density function of \((S_n, Z_n)\) given \(S_0 = s\) and \(Z_0 = e_i\) under the martingale measure, \(Q\), can be written as a sum of terms of the form,

\[
\frac{1}{K_n(j_{1,n})} \exp \left\{ -\frac{1}{2} \left( \frac{\log(x/s) - E^p_n(j_{1,n})}{K_n(j_{1,n})} \right)^2 \right\},
\]

which is why these definitions are given. We also give three preliminary results required to prove Theorem 4.2.3.

**Definition 4.1.1.** For \(n = 1, 2, \ldots, N - 1\), define the functions
\(E_{n+1}, K_{n+1}: \{1, 2, \ldots, M\}^{n+1} \rightarrow \mathbb{R}\) by the recursions,

\[
E_{n+1}(j_{1,n+1}) = E_n(j_{1,n}) + \left( \mu_{j_{n+1}} - \frac{1}{2} \sigma^2_{j_{n+1}} \right) \tau, \tag{4.1}
\]

\[
K_{n+1}(j_{1,n+1}) = \sqrt{\sigma^2_{j_{1}} \tau + K_n(j_{2,n+1})^2}, \tag{4.2}
\]
where the initial functions are \( E_1(j_1) = (\mu_j - \frac{1}{2}\sigma^2_{j_1})\tau \) and \( K_1(j_1) = \sigma_{j_1}\sqrt{\tau} \). Also, for \( p = 1, 2, \ldots, 2^n \) let,

\[
E_n^p(j_{2,n+1}) = E_n(j_{2,n+1}) + \sum_{i=1}^{n} \sigma^2_{j_{i+1}} \tau p_{n-i}, \tag{4.3}
\]

where the constants \( p_i \in \{0, 1\} \) come from the binary representation,

\[
p - 1 = \sum_{i=0}^{n-1} p_i 2^i.
\]

Here, \((\mu_1, \ldots, \mu_M)\) and \((\sigma_1, \ldots, \sigma_M)\) are the vectors of drifts and volatilities in the Switching Black–Scholes model, and \( \tau \) is the time between switches.

It is straightforward to find the general form of \( E_n \) and \( K_n \).

**Lemma 4.1.2.** For \( n = 1, 2, \ldots, N \), the functions \( E_n \) and \( K_n \) defined in Equations (4.1) and (4.2) are given by,

\[
E_n(j_{1,n}) = \sum_{i=1}^{n} (\mu_{j_i} - \frac{1}{2}\sigma^2_{j_i})\tau, \\
K_n(j_{1,n}) = a_{N-n+1}(j_{1,n}) = (\sigma^2_{j_1} + \sigma^2_{j_2} + \ldots + \sigma^2_{j_n})^{1/2} \sqrt{\tau},
\]

where the function \( a_{N-n+1} \) is defined in 3.3.2.

**Proof.** We use induction on \( n \). From Definition 4.1.1, the results hold for \( n = 1 \). Suppose they are true for arbitrary \( n \). Then the definition of \( E_{n+1} \),

\[
E_{n+1}(j_{1,n+1}) = E_n(j_{1,n}) + (\mu_{j_{n+1}} - \frac{1}{2}\sigma^2_{j_{n+1}})\tau,
\]

and the induction hypothesis,

\[
E_n(j_{1,n}) = \sum_{i=1}^{n} (\mu_{j_i} - \frac{1}{2}\sigma^2_{j_i})\tau,
\]

prove the equation for \( E_{n+1} \).

For \( K_{n+1} \), we use the representation of \( a_{N-n} \), given in Definition 3.3.2, and the induction hypothesis to obtain,

\[
K_{n+1}(j_{1,n+1}) = \sqrt{\sigma^2_{j_1}\tau + K_n(j_{2,n+1})^2} \\
= \sqrt{\sigma^2_{j_1}\tau + a_{N-n+1}(j_{2,n+1})^2} \\
= a_{N-n}(j_{1,n+1}).
\]
From Lemma 3.3.3 we have,

\[ a_{N-n}(j_{1,n+1}) = \left( \sigma_{j_1}^2 + \sigma_{j_2}^2 + \ldots + \sigma_{j_{n+1}}^2 \right)^{1/2} \sqrt{\tau}, \]

which completes the proof. \( \square \)

The following two lemmas show how expressions involving \( E^p_n \) can be converted into expressions involving \( E^{2p-1}_{n+1} \) and \( E^{2p}_{n+1} \). The first lemma is needed to prove the second.

**Lemma 4.1.3.** Let \( n \in \{1, 2, \ldots, N-1\} \) and \( j_{n+2} \in \{1, 2, \ldots, M\} \). Then for \( p = 1, 2, \ldots, 2^n \) we have,

\[
\begin{align*}
(\mu_{j_{n+2}} - \frac{1}{2} \sigma_{j_{n+2}}^2) \tau + E^p_n(j_{2,n+1}) &= E^{2p-1}_{n+1}(j_{2,n+2}), \\
(\mu_{j_{n+2}} + \frac{1}{2} \sigma_{j_{n+2}}^2) \tau + E^p_n(j_{2,n+1}) &= E^{2p}_{n+1}(j_{2,n+2}).
\end{align*}
\]  

(4.4)  

(4.5)

**Proof.** The definition of \( E^p_n \) in Equation (4.3) gives,

\[
E^p_n(j_{2,n+1}) = E_n(j_{2,n+1}) + \sum_{i=1}^{n} \sigma_{j_{i+1}}^2 \tau p_{n-i},
\]

(4.6)

where the constants \( p_i \) come from the binary representation, \( p - 1 = \sum_{i=0}^{n-1} p_i 2^i \).

The sum in this equation can also be written as,

\[
\sum_{i=1}^{n+1} \sigma_{j_{i+1}}^2 \tau p'_{n+1-i} \quad \text{where} \quad p'_i = \begin{cases} 
0, & k = 0, \\
p_{k-1}, & k = 1, 2, \ldots, n,
\end{cases}
\]

and from Lemma A.1.4 of Appendix 1,

\[
(2p - 1) - 1 = \sum_{i=0}^{n} p'_i 2^i.
\]

Adding \( (\mu_{j_{n+2}} - \frac{1}{2} \sigma_{j_{n+2}}^2) \tau \) to Equation (4.6) and using this representation of the sum, we have,

\[
\begin{align*}
(\mu_{j_{n+2}} - \frac{1}{2} \sigma_{j_{n+2}}^2) \tau + E^p_n(j_{2,n+1}) &= E_{n+1}(j_{2,n+2}) + \sum_{i=1}^{n+1} \sigma_{j_{i+1}}^2 \tau p'_{n+1-i} \\
&= E^{2p-1}_{n+1}(j_{2,n+2}),
\end{align*}
\]

(4.7)
from the definitions of $E_{n+1}$ and $E_{n+1}^{2p-1}$. This proves Equation (4.4).

For Equation (4.5), we add $\sigma^2_{j,n+2} \tau$ to both sides of Equation (4.7) to obtain,

\[
(\mu_{j,n+2} + \frac{1}{2} \sigma^2_{j,n+2}) \tau + E_n^p(j_{n+1}) = E_{n+1}(j_{n+2}) + \sum_{i=1}^{n+1} \sigma^2_{j,i+1} \tau p_{n+1-i} + \sigma^2_{j,n+2} \tau
\]

\[
= E_{n+1}(j_{n+2}) + \sum_{i=1}^{n+1} \sigma^2_{j,i+1} \tau p''_{n+1-i},
\]

where $p''_0 = 1$ and $p''_k = p'_k = p_{k-1}$ for $k \geq 1$. From Lemma A.1.4 of Appendix 1,

\[
2p - 1 = \sum_{i=0}^{n} p''_i^2,
\]

so the definition of $E_{n+1}^{2p}$ gives Equation (4.5). □

The final lemma of this section calculates an integral which is used to find the density function, $\phi_n$, explicitly in Section 4.2.

**Lemma 4.1.4.** For $n \in \{1, 2, \ldots, N - 1\}$ and $p \in \{1, 2, \ldots, 2^n\}$, let $E_n^p$ and $K_n$ be the functions of Definition 4.1.1. Then for $j_{2,n+2} \in \{1, 2, \ldots, M\}^{n+1}$ and $x, s \in (0, \infty)$ we have,

\[
\int_0^\infty \frac{1}{K_n(j_{2,n+2})^2} \frac{1}{\sigma_{j,n+2} \sqrt{\pi x}} e^{-\frac{1}{2} \left( \frac{\log(x/s) - (\mu_{j,n+2} - \frac{1}{2} \sigma^2_{j,n+2}) \tau}{\sigma_{j,n+2} \sqrt{x}} \right)^2} e^{-\frac{1}{2} \left( \frac{\log(x/s) - E_n^p(j_{2,n+2})}{K_n(j_{2,n+2})} \right)^2} dz
\]

\[
= \sqrt{\frac{2\pi}{K_{n+1}(j_{2,n+2}) x}} e^{-\frac{1}{2} \left( \frac{\log(x/s) - E_n^p(j_{2,n+2})}{K_{n+1}(j_{2,n+2})} \right)^2}.
\]

If we replace $(\mu_{j,n+2} - \frac{1}{2} \sigma^2_{j,n+2}) \tau$ by $(\mu_{j,n+2} + \frac{1}{2} \sigma^2_{j,n+2}) \tau$, the result is analogous, but with $2p - 1$ replaced by $2p$.

**Proof.** Let $A = \log x - (\mu_{j,n+2} - \frac{1}{2} \sigma^2_{j,n+2}) \tau$ and $B = \log s + E_n^p(j_{2,n+1})$. Making the substitution $y = \log z$, the integral is equal to,

\[
\int_{-\infty}^\infty \frac{1}{K_n(j_{2,n+1}) \sigma_{j,n+2} \sqrt{\pi x}} e^{-\frac{1}{2} \left( \frac{(y-A)(y-B)}{\sigma_{j,n+2} \sqrt{x}} \right)^2} dy.
\]

Expanding the quadratics and using a common denominator, the inner part of the exponent becomes,

\[
\frac{1}{\sigma^2_{j,n+2} \tau K_n(j_{2,n+1})^2} \left\{ K_n(j_{2,n+1})^2(y^2 - 2Ay + A^2) + \sigma^2_{j,n+2} \tau (y^2 - 2By + B^2) \right\},
\]
and using the identity \( K_n(j_{2,n+1})^2 + \sigma_{j_{n+2}}^2 \tau = K_{n+1}(j_{2,n+2})^2 \), from Definition 4.1.1, this equals,

\[
\frac{K_{n+1}(j_{2,n+2})^2}{\sigma_{j_{n+2}}^2 \tau K_n(j_{2,n+1})^2} \left\{ y^2 - 2Sy + T \right\},
\]

where,

\[
S = \frac{K_n(j_{2,n+1})^2 A + \sigma_{j_{n+2}}^2 \tau B}{K_{n+1}(j_{2,n+2})^2},
\]

and \( T = \frac{K_n(j_{2,n+1})^2 A^2 + \sigma_{j_{n+2}}^2 \tau B^2}{K_{n+1}(j_{2,n+2})^2} \).

With these manipulations, the integral becomes,

\[
\int_{-\infty}^{\infty} \frac{1}{K_n(j_{2,n+1})\sigma_{j_{n+2}} \sqrt{\pi} x} e^{-\frac{1}{2} \frac{K_{n+1}(j_{2,n+2})^2}{\sigma_{j_{n+2}}^2 \tau K_n(j_{2,n+1})^2} \left\{ (y-S)^2 + T-S^2 \right\}} dy.
\]

Next, change variables to \( z = \frac{K_{n+1}(j_{2,n+2})}{\sigma_{j_{n+2}} \sqrt{\pi} K_n(j_{2,n+1})} (y - S) \) to obtain,

\[
\int_{-\infty}^{\infty} \frac{1}{K_{n+1}(j_{2,n+2}) \sqrt{\pi} x} e^{-\frac{1}{2} \frac{K_{n+1}(j_{2,n+2})^2}{\sigma_{j_{n+2}}^2 \tau K_n(j_{2,n+1})^2} \left\{ T - S^2 \right\}} e^{-\frac{1}{2} z^2} dz,
\]

which can be evaluated to give,

\[
\frac{\sqrt{2\pi}}{K_{n+1}(j_{2,n+2}) x} e^{-\frac{1}{2} \frac{K_{n+1}(j_{2,n+2})^2}{\sigma_{j_{n+2}}^2 \tau K_n(j_{2,n+1})^2} \left\{ T - S^2 \right\}}.
\]

(4.8)

It remains to simplify the exponent in this expression. Using the definitions of \( S \) and \( T \) we have,

\[
K_{n+1}(j_{2,n+2})^2 (T - S^2)
= K_n(j_{2,n+1})^2 A^2 + \sigma_{j_{n+2}}^2 \tau B^2
- \frac{K_n(j_{2,n+1})^4 A^2 + 2\sigma_{j_{n+2}}^2 \tau K_n(j_{2,n+1})^2 AB + \sigma_{j_{n+2}}^2 \tau B^2}{K_{n+1}(j_{2,n+2})^2}
\]

\[
= K_n(j_{2,n+1})^2 A^2 \left( 1 - \frac{K_n(j_{2,n+1})^2}{K_{n+1}(j_{2,n+2})^2} \right) + \sigma_{j_{n+2}}^2 \tau B^2 \left( 1 - \frac{\sigma_{j_{n+2}}^2 \tau}{K_{n+1}(j_{2,n+2})^2} \right)
- 2 K_n(j_{2,n+1})^2 \sigma_{j_{n+2}}^2 \tau AB
\]

As \( K_{n+1} \) satisfies the equation \( K_{n+1}(j_{2,n+2})^2 = \sigma_{j_{n+2}}^2 \tau + K_n(j_{2,n+1})^2 \), this gives,

\[
\frac{K_{n+1}(j_{2,n+2})^2}{\sigma_{j_{n+2}}^2 \tau K_n(j_{2,n+1})^2} [T - S^2] = \left( \frac{A - B}{K_{n+1}(j_{2,n+2})} \right)^2.
\]
From (4.8), the integral is equal to,

$$\sqrt{2\pi} \frac{1}{K_{n+1}(j_{2,n+2})x} e^{-\frac{1}{2}(\frac{A-B}{K_{n+1}(j_{2,n+2})})^2}.$$ 

The definitions of $A$ and $B$ give,

$$A - B = \log(x/s) - (\mu_{j_{n+2}} - \frac{1}{2}\sigma^2_{j_{n+2}})\tau - E_n^p(j_{2,n+1}),$$

and from Lemma 4.1.3,

$$A - B = \log(x/s) - E_{n+1}^{2p-1}(j_{2,n+2}).$$

The result follows.

Finally, if $(\mu_{j_{n+2}} - \frac{1}{2}\sigma^2_{j_{n+2}})\tau$ is replaced by $(\mu_{j_{n+2}} + \frac{1}{2}\sigma^2_{j_{n+2}})\tau$, the calculations above are unchanged, except that we set $A = \log x - (\mu_{j_{n+2}} + \frac{1}{2}\sigma^2_{j_{n+2}})\tau$. Then, again using Lemma 4.1.3,

$$A - B = \log(x/s) - (\mu_{j_{n+2}} + \frac{1}{2}\sigma^2_{j_{n+2}})\tau - E_n^p(j_{2,n+1})$$

$$= \log(x/s) - E_{n+1}^{2p}(j_{2,n+2}),$$

which completes the proof. $\square$

## 4.2 Density Functions

With the definitions and results of Section 4.1, we now compute the required density functions. Recall from Definition 2.2.4 of Chapter 2 that $\phi_n$ denotes the joint conditional density function of $(S_n, Z_n)$ given $\mathcal{F}_0$, and $\psi_n$ denotes the joint conditional density function of $(S_n, Z_n)$ given $\mathcal{F}_{n-1}$, under a martingale measure $Q$. In this chapter, we take $Q$ to be the $b$-variance-optimal signed martingale measure, assumed to be non-negative.

Also, recall from Theorem 2.2.5 how the density function $\phi_n$ can be calculated from $\psi_n$ and $\phi_{n-1}$ using the forward induction relationship,

$$\phi_n(s, e_i, x, e_j) = \sum_{l=1}^{M} \int_{-\infty}^{\infty} \psi_n(z, e_l, x, e_j) \phi_{n-1}(s, e_i, z, e_l) \, dz.$$
This recursion is used in Theorem 4.2.3 to find the general form of the density function, \( \phi_n \). In order to use this relationship, we begin in Theorem 4.2.1 by finding the general form of \( \psi_n \) in terms of previously defined processes.

**Theorem 4.2.1.** Let \( n \in \{1, 2, \ldots, N\} \) and let \( Q \) be the \( b \)-variance-optimal martingale measure for the Switching Black–Scholes model. Then for any \( F_n \)-measurable random variable, \( H_n \), we have,

\[
E_Q[H_n | F_{n-1}] = \frac{1}{D_{n-1}} E[D_n (1 - \beta_n \Delta X_n) H_n | F_{n-1}],
\]

where \( (\beta_k)_k = 1, 2, \ldots, N \) and \( (D_k)_{k=0,1,\ldots,N} \) are the stochastic processes defined in 3.2.1 and 3.2.2 respectively.

Also, the joint conditional density function of \( (S_n, Z_n) \) given \( S_{n-1} = s \) and \( Z_{n-1} = e_i \) under the martingale measure, \( Q \), is,

\[
\psi_n(s, e_i, x, e_j) = \begin{cases} 
\frac{D_n(e_j)}{D_{n-1}(e_i)} \frac{A_{ji}}{\sqrt{2\pi}} \left(M_n(e_i) - \frac{x}{s} f_n(e_i)\right) e^{-\frac{1}{2} \left(\frac{\log(x/s) - (\mu_j - \frac{\sigma_j^2}{2})}{\sigma_j \sqrt{\tau}}\right)^2}, & x > 0 \\
0, & x \leq 0,
\end{cases}
\]

for \( e_j \in \mathcal{H} \) and \( x \in \mathbb{R} \). (Recall the definitions of \( f_n \) and \( M_n \) from Notation 3.4.1 on page 47).

**Proof.** Bayes rule gives,

\[
E_Q[H_n | F_{n-1}] = \frac{E \left[ \frac{dQ}{dP} H_n \right] | F_{n-1}}{E \left[ \frac{dQ}{dP} \right] | F_{n-1}}.
\]

From Equation (3.10) on page 33 we have \( \frac{dQ}{dP} = \frac{1}{\tilde{c}} \prod_{j=1}^N (1 - \beta_j \Delta X_j) \), where \( \tilde{c} \) is a constant. The first \( n - 1 \) terms of this expression, \( \frac{1}{\tilde{c}} \prod_{j=1}^{n-1} (1 - \beta_j \Delta X_j) \), are \( F_{n-1} \)-measurable, so Equation (4.10) becomes,

\[
E_Q[H_n | F_{n-1}] = \frac{E \left[ \prod_{j=n}^N (1 - \beta_j \Delta X_j) H_n \right] | F_{n-1}}{E \left[ \prod_{j=n}^N (1 - \beta_j \Delta X_j) \right] | F_{n-1}}.
\]

By Definition 3.2.2 on page 34, the denominator is just \( D_{n-1} \), and conditioning first on \( F_n \) in the numerator gives,

\[
E_Q[H_n | F_{n-1}] = \frac{1}{D_{n-1}} E[D_n (1 - \beta_n \Delta X_n) H_n | F_{n-1}],
\]
which verifies Equation (4.9).

Next, we need to find \( \psi_n \), the joint conditional density function of \((S_n, Z_n)\)
given \( S_{n-1} = s \) and \( Z_{n-1} = \epsilon_i \) under the martingale measure, \( Q \). We do this by computing the above expectation. Recall from Equation (3.32) on page 48 that,

\[
1 - \beta_n \Delta X_n = M_n(Z_{n-1}) - \frac{S_n}{S_{n-1}} f_n(Z_{n-1}).
\]

Using this and Theorem 2.2.1, which gives the joint conditional density function of \((S_n, Z_n)\) given \( S_{n-1} = s \), \( Z_{n-1} = \epsilon_i \) under the real world probability measure, Equation (4.9) implies that,

\[
E^Q[H_n | S_{n-1} = s, Z_{n-1} = \epsilon_i] = \sum_{j=1}^{M} \int_0^\infty \frac{D_n(e_j)}{D_{n-1}(e_i)} \frac{A_{ji}}{\sqrt{2\pi}} \left( \frac{M_n(e_i) - x}{s f_n(e_i)} \right)
\times \frac{1}{\sigma_j \sqrt{\pi x}} e^{-\frac{1}{2} \left( \frac{\log(s/x) - (\mu_j + \frac{1}{2} \sigma_j^2) x}{\sigma_j \sqrt{x}} \right)^2} H_n(x, e_j) dx.
\]

Choosing \( H_n(S_n, Z_n) = I(S_n \leq s', Z_n = e_{j'}) \) gives the result. \( \Box \)

For future calculations, it is more convenient to write the density function, \( \psi_n \), in an alternative way, given in the following lemma.

**Lemma 4.2.2.** Let \( n \in \{1, 2, \ldots, N\} \) and let \( Q \) be the b-variance-optimal martingale measure for the Switching Black–Scholes model. The joint conditional density function of \((S_n, Z_n)\) given \( S_{n-1} = s \) and \( Z_{n-1} = \epsilon_i \) under the martingale measure, \( Q \), is,

\[
\psi_n(s, \epsilon_i, x, e_j) = \frac{D_n(e_j)}{D_{n-1}(e_i)} \frac{A_{ji}}{\sqrt{2\pi}} \left\{ \frac{M_n(e_i)}{\sigma_j \sqrt{x}} e^{-\frac{1}{2} \left( \frac{\log(s/x) - (\mu_j + \frac{1}{2} \sigma_j^2) x}{\sigma_j \sqrt{x}} \right)^2} f_n(e_i) J_j \frac{1}{\sigma_j \sqrt{\pi x}} e^{-\frac{1}{2} \left( \frac{\log(s/x) - (\mu_j + \frac{1}{2} \sigma_j^2) x}{\sigma_j \sqrt{x}} \right)^2} \right\},
\]

for \( e_j \in H \) and \( x > 0 \), and \( \psi_n(s, \epsilon_i, x, e_j) = 0 \) for \( x \leq 0 \). Here, the stochastic process \((D_k)_{k=0,1,\ldots,N}\) is defined in 3.2.2 and the coefficients \( M_n, f_n \) and \( J_j \) are defined in 3.4.1.
Proof. Using \( x/s = e^{\log(x/s)} \) and Theorem 4.2.1, \( \psi_n(s, e_i, x, e_j) \) is equal to,

\[
\frac{D_n(e_j)}{D_{n-1}(e_i)} \frac{A_{ji}}{\sqrt{2\pi}} \left\{ M_n(e_i) \frac{1}{\sigma_j \sqrt{T} x} e^{-\frac{1}{2} \left( \frac{\log(x/s) - (\mu_j - \frac{1}{2} \sigma_j^2) \tau}{\sigma_j \sqrt{T}} \right)^2} - f_n(e_i) \frac{1}{\sigma_j \sqrt{T} x} e^{-\frac{1}{2} \left( \frac{\log(x/s) - (\mu_j - \frac{1}{2} \sigma_j^2) \tau}{\sigma_j \sqrt{T}} \right)^2} \right\}.
\]

For the remainder of the proof, we consider the exponent in the second term, which is,

\[
- \frac{1}{2} \left\{ \frac{\log^2(x/s) - 2 \log(x/s) (\mu_j + \frac{1}{2} \sigma_j^2) \tau + (\mu_j - \frac{1}{2} \sigma_j^2)^2 \tau^2}{\sigma_j^2 \tau} \right\}.
\]

Expanding the square gives,

\[
- \frac{1}{2} \left\{ \frac{\log^2(x/s) - 2 \log(x/s) (\mu_j + \frac{1}{2} \sigma_j^2) \tau + (\mu_j - \frac{1}{2} \sigma_j^2)^2 \tau^2}{\sigma_j^2 \tau} \right\},
\]

or equivalently,

\[
- \frac{1}{2} \left\{ \left( \frac{\log(x/s) - (\mu_j + \frac{1}{2} \sigma_j^2) \tau}{\sigma_j \sqrt{T}} \right)^2 + \frac{(\mu_j - \frac{1}{2} \sigma_j^2)^2 \tau^2 - (\mu_j + \frac{1}{2} \sigma_j^2)^2 \tau^2}{\sigma_j^2 \tau} \right\}.
\]

The second term here simplifies to \(-2\mu_j \tau\), so the second exponential in (4.11) is,

\[
e^{\mu_j \tau} e^{-\frac{1}{2} \left( \frac{\log(x/s) - (\mu_j + \frac{1}{2} \sigma_j^2) \tau}{\sigma_j \sqrt{T}} \right)^2}
\]

As \( J_j = e^{\mu_j \tau} \), this completes the proof. \( \square \)

Having found the form of \( \psi_n \), we now use the forward induction relationship,

\[
\phi_n(s, e_i, x, e_j) = \sum_{l=1}^{M} \int_{-\infty}^{\infty} \psi_n(z, e_l, x, e_j) \phi_{n-1}(s, e_i, z, e_l) \, dz,
\]

to find the density function, \( \phi_n \), for \( n = 1, 2, \ldots, N \). This, in turn, will be used in Section 4.3 to find the mean-variance price of a call option, \( H \), on an asset following the Switching Black–Scholes model using the formula,

\[
H_0(S_0, e_{j_1}) = e^{-rT} \sum_{j=1}^{M} \int_{-\infty}^{\infty} \phi_N(S_0, e_{j_1}, x, e_j) H(x, e_j) \, dx,
\]

where \( H \) is a contingent claim and the asset has initial price, \( S_0 \), and initial market state, \( e_{j_1} \).

The next theorem gives the general form of the density function, \( \phi_n \).
Theorem 4.2.3. For $n \in \{1, 2, \ldots, N\}$, the joint conditional density function of $(S_n, Z_n)$ given $S_0 = s$ and $Z_0 = e_{j_1}$ under the $b$-variance optimal martingale measure, $Q$, is,

$$
\phi_n(s, e_{j_1}, x, e_{j_{n+1}}) = \sum_{j_2,j_3,\ldots,j_n=1}^{M_n} \frac{D_n(e_{j_{n+1}}) A_{j_{n+1}j_n} A_{j_nj_{n-1}} \cdots A_{j_2j_1}}{D_0(e_{j_1})} \sqrt{\frac{2\pi}{K_n(j_2,j_n+1,x)}}
\times \frac{1}{K_n(j_2,n+1,x)} \sum_{p=1}^{2^n} c^n_p(j_{1,n+1}) e^{-\frac{1}{2} \left( \frac{\log(s/x) - E_n^p(j_{2,n+1})}{\sqrt{K_n(j_2,n+1,x)}} \right)^2},
$$

for $e_{j_{n+1}} \in \mathcal{H}$ and $x > 0$, and $\phi_n(s, e_{j_1}, x, e_{j_{n+1}}) = 0$ for $x \leq 0$. Here, the stochastic process $(D_n)_{k=0,1,\ldots,N}$ is defined in 3.2.2, $K_n$ and $E_n^p$ are as defined in 4.1.1, and the constants $c^n_p$ are defined recursively by,

$$
c^{n+1}_p(j_{1,n+2}) = \begin{cases} 
M_n+1(e_{j_{n+1}}) c^{n+1}_{p+1/2}(j_{1,n+1}), & p \text{ odd,} \\
-f_{n+1}(e_{j_{n+1}}) J_{j_{n+2}} c^{n+1}_{p+1/2}(j_{1,n+1}), & p \text{ even,}
\end{cases}
$$

for $p = 1, 2, \ldots, 2^{n+1}$. The initial values are,

$$
c^1(j_1, j_2) = M_1(e_{j_1}),
$$

$$
c^1_2(j_1, j_2) = -f_1(e_{j_1}) J_{j_2},
$$

and $M_n$, $f_n$ and $J_{j_2}$ are defined in 3.4.1.

Note that for the case $n = 1$, the sum disappears, but the same expression remains with the $A$ part being $A_{j_2j_1}$.

Proof. The proof is by induction on $n$. For $n = 1$, the definitions of $\phi_1$ and $\psi_1$ immediately give $\phi_1(s, e_{j_1}, x, e_{j_2}) = \psi_1(s, e_{j_1}, x, e_{j_2})$. Using Lemma 4.2.2, we then see that $\phi_1(s, e_{j_1}, x, e_{j_2}) = 0$ for $x \leq 0$ and,

$$
\phi_1(s, e_{j_1}, x, e_{j_2}) = \frac{D_1(e_{j_1}) A_{j_2j_1}}{D_0(e_{j_1})} \sqrt{\frac{2\pi}{K_1(j_2, x)}} \sum_{p=1}^{2} c^1_p(j_1, j_2) e^{-\frac{1}{2} \left( \frac{\log(s/x) - E^1_p(j_2)}{\sqrt{K_1(j_2)}} \right)^2},
$$

for $e_{j_2} \in \mathcal{H}$ and $x > 0$. For $n > 1$, the induction hypothesis is applied to $\phi_{n-1}$ and $\psi_{n-1}$, which yields $\phi_n(s, e_{j_1}, x, e_{j_{n+1}}) = \psi_n(s, e_{j_1}, x, e_{j_{n+1}})$, and the result follows.
for \( x > 0 \), as required.

Next, suppose that the result is true for arbitrary \( n \). From Theorem 2.2.5, \( \phi_{n+1} \) can be calculated from \( \phi_n \) and \( \psi_{n+1} \) using the formula,

\[
\phi_{n+1}(s, e_j, x, e_{j,n+2}) = \sum_{j_{n+1} = 1}^{M} \int_{-\infty}^{\infty} \psi_{n+1}(z, e_{j,n+1}, x, e_{j,n+2}) \phi_n(s, e_j, z, e_{j,n+1}) \, dz.
\]

Applying the induction hypothesis and using the expression for \( \psi_{n+1} \) given in Lemma 4.2.2, this equals,

\[
\sum_{j_{n+1} = 1}^{M} \int_{0}^{\infty} \frac{D_{n+1}(e_{j,n+2})}{D_n(e_{j_{n+1}})} \frac{A_{j_{n+2}j_{n+1}}}{\sqrt{2\pi}} \left\{ M_{n+1}(e_{j_{n+1}}) \frac{1}{\sigma_{j_{n+2}} \sqrt{\pi \tau}} e^{-\frac{1}{2} \left( \frac{\log(x/z) - (\mu_{j_{n+2}} + \frac{\lambda^2}{2\sigma_{j_{n+2}}^2})^2}{\sigma_{j_{n+2}}^2} \right)^2} \right\} \times \sum_{j_2, j_3, \ldots, j_n = 1}^{M} \frac{D_n(e_{j_{n+1}})}{D_0(e_{j_1})} \frac{A_{j_{n+1}j_1}A_{j_{n+1}j_2} \cdots A_{j_{n+1}j_{n-1}}A_{j_{n+1}j_n}A_{j_{n+1}j_{n+2}}}{\sqrt{2\pi}} \frac{1}{K_n(j_{2,n+1})} \sum_{p=1}^{2^n} c_p^n(j_{1,n+1}) e^{-\frac{1}{2} \left( \frac{\log(x/z) - \mu_{j_{2,n+1}}}{K_n(j_{2,n+1})} \right)^2} \, dz.
\]

We can simplify and rearrange this expression to obtain,

\[
\sum_{j_2, j_3, \ldots, j_{n+1} = 1}^{M} \frac{D_{n+1}(e_{j_{n+2}})}{D_0(e_{j_1})} \frac{A_{j_{n+2}j_{n+1}}A_{j_{n+1}j_{n-1}} \cdots A_{j_{n+1}j_{n+2}}}{2\pi} \sum_{p=1}^{2^n} c_p^n(j_{1,n+1}) \times \left\{ M_{n+1}(e_{j_{n+1}}) \int_{0}^{\infty} \left[ \frac{1}{\sigma_{j_{n+2}} \sqrt{\pi \tau}} e^{-\frac{1}{2} \left( \frac{\log(x/z) - (\mu_{j_{n+2}} + \frac{\lambda^2}{2\sigma_{j_{n+2}}^2})^2}{\sigma_{j_{n+2}}^2} \right)^2} \right]\right. \\
\left. \times \frac{1}{K_n(j_{2,n+1})} e^{-\frac{1}{2} \left( \frac{\log(x/z) - \mu_{j_{2,n+1}}}{K_n(j_{2,n+1})} \right)^2} \right. \\
\left. \frac{1}{K_n(j_{2,n+1})} e^{-\frac{1}{2} \left( \frac{\log(x/z) - \mu_{j_{2,n+1}}}{K_n(j_{2,n+1})} \right)^2} \right) \, dz \right. \\
- \sum_{p=1}^{2^n} c_p^n(j_{1,n+1}) \int_{0}^{\infty} \left[ \frac{1}{\sigma_{j_{n+2}} \sqrt{\pi \tau}} e^{-\frac{1}{2} \left( \frac{\log(x/z) - (\mu_{j_{n+2}} + \frac{\lambda^2}{2\sigma_{j_{n+2}}^2})^2}{\sigma_{j_{n+2}}^2} \right)^2} \right] \\
\left. \times \frac{1}{K_n(j_{2,n+1})} e^{-\frac{1}{2} \left( \frac{\log(x/z) - \mu_{j_{2,n+1}}}{K_n(j_{2,n+1})} \right)^2} \right) \, dz \right\},
\]
and invoking Lemma 4.1.4 for these two integrals gives,

\[
\sum_{j_2,j_3,\ldots,j_{n+1}=1}^{M} \frac{D_{n+1}(e_{j_{n+2}})}{D_0(e_{j_1})}\frac{A_{j_{n+2}j_{n+1}}A_{j_{n+1}j_n}\cdots A_{j_{2}j_1}}{\sqrt{2\pi}} \sum_{p=1}^{2^n} e_p^{n+1}(j_{1,n+1})
\times \frac{1}{K_{n+1}(j_{2,n+2})x} \left\{ M_{n+1}(e_{j_{n+1}}) - \frac{1}{2} \left( \frac{\log(x/j_{n+2}) - s_n^{n+1}(j_{2,n+2})}{K_{n+1}(j_{2,n+2})} \right)^2 
- f_{n+1}(e_{j_{n+1}})J_{j_{n+2}}e - \frac{1}{2} \left( \frac{\log(x/j_{n+2}) - s_n^{n+1}(j_{2,n+2})}{K_{n+1}(j_{2,n+2})} \right)^2 \right\}.
\]

Finally, the definition of \( e_p^{n+1} \) allows us to write this as,

\[
\sum_{j_2,j_3,\ldots,j_{n+1}=1}^{M} \frac{D_{n+1}(e_{j_{n+2}})}{D_0(e_{j_1})}\frac{A_{j_{n+2}j_{n+1}}A_{j_{n+1}j_n}\cdots A_{j_{2}j_1}}{\sqrt{2\pi}} \frac{1}{K_{n+1}(j_{2,n+2})x} \times \sum_{p=1}^{2^n+1} e_p^{n+1}(j_{1,n+2}) - \frac{1}{2} \left( \frac{\log(x/j_{n+2}) - s_n^{n+1}(j_{2,n+2})}{K_{n+1}(j_{2,n+2})} \right)^2.
\]

This shows that \( \phi_{n+1} \) is of the required form, so by the principle of mathematical induction, the proof is complete.

(Note: to prove the equation for \( n = 2 \), we follow precisely the same procedure as for the induction step). \( \square \)

So that we can price securities at any time \( m \), rather than just at time 0, Theorem 4.2.3 can be modified to give the joint conditional density function of \((S_n, Z_n)\) given \( S_m = s \) and \( Z_m = e_{j_1} \) under the measure \( Q \), for all \( n > m \). This density function is computed in Appendix 2.

We now use the density function, \( \phi_n \), given in Theorem 4.2.3, to price a call option on an asset following the Switching Black–Scholes model.

### 4.3 Call Option Price

Suppose \( H_n \) is an \( \mathcal{F}_n \)-measurable contingent claim. The mean-variance price at time 0 of \( H_n \), given \( S_0 \) and \( Z_0 = e_{j_1} \), is,

\[
H_0(S_0, e_{j_1}) = e^{-rn} \mathbb{E}^{Q}[H_n | S_0, Z_0 = e_{j_1}].
\]
Using the density function, \( \phi_n \), of \((S_n, Z_n)\) given \(S_0\) and \(Z_0 = e_{j_1}\) under the martingale measure, \(Q\), we can write this as,

\[
H_0(S_0, e_{j_1}) = e^{-rT} \sum_{j_{n+1}=1}^{M} \phi_n(S_0, e_{j_1}, x, e_{j_{n+1}}) H_n(x, e_{j_{n+1}}) dx. \tag{4.12}
\]

We now use this formula to compute the mean-variance price of a call option on an asset following the Switching Black–Scholes model.

**Theorem 4.3.1.** Let \(S\) be an asset following the Switching Black–Scholes model, with initial asset price, \(S_0\), and initial market state, \(e_{j_1}\). The mean-variance price of a call option expiring at time \(T\) with strike price, \(K\), on the asset, \(S\), is,

\[
H_0(S_0, e_{j_1}) = \frac{e^{-rT}}{D_0(e_{j_1})} \sum_{j_2, \ldots, j_{n+1}=1}^{M} A_{j_{n+1}j_n} \cdots A_{j_{2}j_1} \sum_{p=1}^{2^N} c_p^N(j_1, N+1)
\times \left\{S_0 e^{\frac{\sum_{i=2}^{N+1} \mu_{j_i} \tau}{2 \sigma_{j_i}^2} + \sum_{i=1}^{N} \sigma_{j_{i+1}}^2 \tau p_{N-i}} - K N(d_p^-) \right\},
\]

where \(D_0\) is defined in 3.2.2, \(c_p^N\) is defined in 4.2.3,

\[
d_p^\pm = \log(S_0/K) + \sum_{i=2}^{N+1} (\mu_{j_i} \pm \frac{1}{2} \sigma_{j_i}^2) \tau + \sum_{i=1}^{N} \sigma_{j_{i+1}}^2 \tau p_{N-i},
\]

and the constants \(p_i \in \{0, 1\}\) come from the binary representation of \(p - 1\),

\[
p - 1 = \sum_{i=0}^{N-1} p_i 2^i.
\]

**Proof.** From Equation (4.12), the price is,

\[
H_0(S_0, e_{j_1}) = e^{-rT} \sum_{j_{n+1}=1}^{M} \phi_N(S_0, e_{j_1}, x, e_{j_{n+1}}) H_N(x, e_{j_{n+1}}) dx,
\]

where \(H_N(x, e_{j_{n+1}}) = \max(x - K, 0)\), and using Theorem 4.2.3, which gives the density \(\phi_N\), we have,

\[
H_0(S_0, e_{j_1}) = \frac{e^{-rT}}{D_0(e_{j_1})} \sum_{j_2, \ldots, j_{n+1}=1}^{M} \frac{D_N(e_{j_{n+1}})}{D_0(e_{j_1})} A_{j_{n+1}j_n} A_{j_{n}j_{n-1}} \cdots A_{j_{2}j_1} A_{j_{1}j_1}
\times \sum_{p=1}^{2^N} c_p^N(j_1, N+1) \int_{K}^{\infty} \frac{1}{K_N(j_2, N+1)x} e^{-\frac{1}{2} \left(\frac{\log(x/K_0) - K_{N}^2(j_{n+1})}{K_{N}(j_{n+1})x}\right)^2} (x - K) dx. \tag{4.13}
\]
Changing variables to $y = \frac{\log(x/S_0) - E^p_N (j_{2,N+1})}{K_N(j_{2,N+1})}$, the integral becomes,

$$
\int_{-d^-_p}^\infty e^{-\frac{1}{2}y^2} \left( S_0 e^{E^p_N (j_{2,N+1}) + K_N(j_{2,N+1})y} - K \right) \, dy,
$$

(4.14)

where $d^-_p = d^-_p (j_{2,N+1}, S_0, K) = \frac{\log(S_0/K) + E^p_N (j_{2,N+1})}{K_N(j_{2,N+1})}$. We consider the two parts of this integral individually. The first part is,

$$
\int_{-d^-_p}^\infty S_0 e^{-\frac{1}{2}y^2 + E^p_N (j_{2,N+1}) + K_N(j_{2,N+1})y} \, dy
$$

$$
= \sqrt{2\pi} S_0 e^{E^p_N (j_{2,N+1}) + \frac{1}{2}K_N(j_{2,N+1})^2} N \left( d^-_p + K_N(j_{2,N+1}) \right),
$$

from Lemma A.1.3 of Appendix 1, and the second part is,

$$
\int_{-d^-_p}^\infty K e^{-\frac{1}{2}y^2} \, dy = \sqrt{2\pi} K N(d^-_p).
$$

Hence the integral (4.14) becomes,

$$
\sqrt{2\pi} \left\{ S_0 e^{E^p_N (j_{2,N+1}) + \frac{1}{2}K_N(j_{2,N+1})^2} N \left( d^-_p + K_N(j_{2,N+1}) \right) - K N(d^-_p) \right\}.
$$

Now, from (4.13), the price of the call option is,

$$
H_0(S_0, e_{j_1}) = \frac{e^{-rT}}{D_0(e_{j_1})} \sum_{j_{2,N+1} = 1}^M A_{j_{N+1}j_{N}} \cdots A_{j_{2}j_{1}} \sum_{p=1}^{2^N} c_p^N (j_{1,N+1})
$$

$$
\times \left\{ S_0 e^{E^p_N (j_{2,N+1}) + \frac{1}{2}K_N(j_{2,N+1})^2} N \left( d^-_p + K_N(j_{2,N+1}) \right) - K N(d^-_p) \right\},
$$

as, by definition, $D_N \equiv 1$.

Consider the expressions for $E^p_N (j_{2,N+1}) + \frac{1}{2}K_N(j_{2,N+1})^2$ and $d^-_p$. From Definition 4.1.1 and Lemma 4.1.2, we have,

$$
E^p_N (j_{2,N+1}) + \frac{1}{2}K_N(j_{2,N+1})^2 = \sum_{i=2}^{N+1} \mu_{j_i} \tau + \sum_{i=1}^{N} \sigma^2_{j_{i+1}} \tau p_{N-i},
$$

and,

$$
d^-_p = \frac{\log(S_0/K) + \sum_{i=2}^{N+1} (\mu_{j_i} - \frac{1}{2}\sigma^2_{j_i}) \tau + \sum_{i=1}^{N} \sigma^2_{j_{i+1}} \tau p_{N-i}}{\left( \sum_{i=2}^{N+1} \sigma^2_{j_i} \tau \right)^{1/2}}.
$$

Clearly,

$$
d^+ + K_N(j_{2,N+1}) = d^-_p + \left( \sum_{i=2}^{N+1} \sigma^2_{j_i} \tau \right)^{1/2} = d^+_p,
$$

$\text{assertion}$. 

$\text{proposition}$
so the call price is,

\[
H_0(S_0, e_{j_1}) = \frac{e^{-rT}}{D_0(e_{j_1})} \sum_{j_2, \ldots, j_{N+1} = 1}^{M} A_{j_{N+1}j_N} \cdots A_{j_2j_1} \sum_{p=1}^{2^N} c_p^{N}(j_1, N+1) \\
\times \left\{ S_0 e^{\sum_{i=2}^{N+1} \mu_s \tau} e^{\sum_{i=1}^{N} \sigma_{i+1}^2 \tau_{i+1} \tau_{i+1} - \frac{\mu_s \tau}{2}} N(d^+_{d_p}) - K N(d^-_{d_p}) \right\} .
\]
Chapter 5

Option Pricing via the Esscher Transform

In this chapter, we use Esscher transforms to define a martingale measure under which derivative pricing can be performed. As for the method of mean-variance hedging, described in Chapters 3 and 4, the asset price is assumed to behave according to the Switching Black–Scholes model, and the price of an option on such an asset is found, in this case via the Esscher martingale measure.

We begin in Section 5.1 by defining and giving a brief history of Esscher transforms, which were first developed in 1932, but were not applied to option pricing until 1994. The reasons for choosing the Esscher martingale measure are also discussed. A precise definition of this martingale measure in our context is given in Section 5.2, followed by a theorem proving equations which define the unknown parameters involved in the definition.

The remaining two sections of the chapter follow methods analogous to those of Chapter 4. The Esscher price of an $\mathcal{F}_n$-measurable contingent claim, $H_n$, at time 0 is calculated using the joint conditional density function, $\phi_n$, of $(S_n, Z_n)$ given $S_0 = s$ and $Z_0 = e_i$ under the Esscher martingale measure, according to the equation,

$$H_0(s, e_i) = e^{-r_0 t} \sum_{j=1}^M \int_{-\infty}^{\infty} \phi_n(s, e_i, x, e_j) H_n(x, e_j) \, dx.$$  \hspace{1cm} (5.1)
As in Section 4.1 of Chapter 4, to compute the density function, $\phi_n$, and hence option prices, we first recursively define two functions, $E_n$ and $K_n$ for $n = 1, 2, \ldots, N$, and prove some results concerning these values. These results are given in Section 5.3.

Next, to calculate the density function, $\phi_n$, we use Theorem 2.2.5 of Chapter 2, which states that,

$$\phi_n(s, e_i, x, e_j) = \sum_{l=1}^{M} \int_{-\infty}^{\infty} \psi_n(u, e_l, x, e_j) \phi_{n-1}(s, e_i, u, e_l) \, du.$$ 

Here, $\psi_n$ is the joint conditional density function of $(S_n, Z_n)$ given $S_{n-1} = u$ and $Z_{n-1} = e_l$ under the Esscher martingale measure. The form of $\psi_n$, and hence the form of $\phi_n$, is found in Section 5.4, and the relationship (5.1) is then used to evaluate the Esscher price of a call option in a Switching Black-Scholes market. We will see that the Esscher call price has the form,

$$\sum_{j_2, j_3, \ldots, j_{N+1} = 1}^{M} p(j_2, j_3, \ldots, j_{N+1}) \{ S_0 e^{f(j_2, \ldots, j_{N+1})} N \left( d_j^- \right) - Ke^{-rT} N \left( d_j^+ \right) \}$$

where $p$ is a probability, $f(j_2, \ldots, j_{N+1})$ is a constant, and $d_j^-$ and $d_j^+$ are analogous to $d_1$ and $d_2$ of the Black-Scholes formula. In particular, for the special case of no switching (that is, $M = 1$), the Esscher call price is equal to the Black-Scholes call price.

We begin by giving a brief history of the Esscher transform and its application to option pricing.

### 5.1 History of the Esscher Transform

Esscher transforms were first developed in 1932 by the Swedish actuary, F. Esscher. In his paper [30], he applied the transform to a random variable to give a new distribution centred at a point of interest, enabling more accurate approximations to be made at that point. More generally, Esscher transforms have become widely used in actuarial practices as a method of assigning larger
weights to adverse events and smaller weights to beneficial events. A typical example is in the calculation of life insurance premiums, where companies base their calculations not on the known distribution of life expectancy, but on a distribution shifted in this way.

In option pricing, the Esscher transform is applied to a stochastic process, rather than its original application to a single random variable. The asset is assumed to evolve according to the model,

\[ S_t = S_0 e^{X_t}, \]

for some stochastic process, X, and a new measure is defined by taking the Esscher transform of X,

\[ \frac{dQ}{dP} \big|_{\mathcal{F}_t} = \frac{e^{\alpha X_t}}{E[e^{\alpha X_t}]}, \]

where \((\mathcal{F}_t)\) is the filtration generated by \((S_t)\). The constant \(\alpha\) is chosen so that the discounted asset price is a martingale with respect to the measure \(Q\), and \(Q\) is called the Esscher martingale measure. Option pricing is then carried out using this martingale measure.

Gerber and Shiu, in their 1994 paper [34], were the first to use this method of option pricing. They assumed that \((X_t)_{t \geq 0}\) has stationary and independent increments, and considered in particular the cases where \(X\) follows a Wiener, Poisson, inverse Gaussian and Gamma process. In 1996, they extended this work to allow multiple assets and dividend paying stocks ([35]).

Yao ([81]), Chan ([15]), and Chan and van der Hoek ([16]) also consider the use of Esscher transforms for pricing. Yao examines the case of stochastic interest rates, while Chan and van der Hoek consider a discrete model, \(S_t = S_0 e^{Y_1 + Y_2 + \ldots + Y_t}\), where the random variables \(Y_i\) are independent and identically distributed, and have a finite number of possible values. This discrete model can be regarded as a generalisation of the Binomial asset price model.

The paper by Chan ([15]) discusses Esscher pricing for assets following the model,

\[ dS_t = \sigma_t S_t \, dY_t + b_t S_t \, dt, \]
where \( Y_t \) is a general Lévy process. He uses a more general Esscher transform (also discussed by Kallsen and Shiryaev, \([55]\)) which applies when the asset price is a stochastic exponential. Chan observes that for this model, pricing via Esscher transforms is equivalent to pricing via minimum entropy\(^1\).

When given a martingale measure, it is important to consider the benefits of that particular martingale measure over any other choice. The Esscher martingale measure has two main benefits. Firstly, as commented by Michaud in discussion of Gerber and Shiu’s paper, \([34]\), the new measure remains in the same family of models as the real world probability measure. Secondly, pricing by Esscher transforms corresponds to maximising the expected utility with utility function,

\[
    u(x) = \frac{x^\gamma}{\gamma},
\]

for \( 0 < \gamma < 1 \), (see Gerber and Shiu, \([35]\), or Chan and van der Hoek, \([16]\)).

### 5.2 The Esscher Martingale Measure

In this section, we define, using Esscher transforms, a martingale measure for a Switching Black–Scholes market, and consider how this martingale measure can be found. Recall that an asset price, \((S_n)_{n = 0, 1, \ldots, N}\), which evolves according to the Switching Black–Scholes model satisfies the equation,

\[
    S_n = S_{n-1} \exp \left\{ \left( \mu(Z_n) - \frac{1}{2} \sigma(Z_n)^2 \right) \tau + \sigma(Z_n) \Delta W_n \right\},
\]

for \( n = 1, 2, \ldots, N \). Hence we may write the asset price in a form amenable to pricing via the Esscher transform:

\[
    S_n = S_{n-1} e^{X_n},
\]

where \( X_n = (\mu(Z_n) - \frac{1}{2} \sigma(Z_n)^2) \tau + \sigma(Z_n) \Delta W_n \).

For this model, the Esscher martingale measure, under which pricing via the Esscher transform is performed, is defined by taking the Esscher transform of

\(^1\)Pricing via minimum entropy will be discussed in Chapter 6.
the process $X$ in each time period, and multiplying these terms together. The following definition makes this explicit.

**Definition 5.2.1.** For the Switching Black–Scholes market model, the Esscher martingale measure, $Q$, is defined by the equation,

$$
\frac{dQ}{dP}_{F_n} = \xi_n = \xi_{n-1} Y_n,
$$

(5.2)

for $n = 0, 1, \ldots, N$. Here, $Y_0 = 1$, $\xi_{-1} = 1$, and $Y_n$, for $n = 1, 2, \ldots, N$, is defined by the equation,

$$
Y_n = \frac{e^{\alpha(n-1) X_n}}{E[e^{\alpha(n-1) X_n} | F_{n-1}]},
$$

(5.3)

where the Esscher parameter, $\alpha^{(n-1)}$, is an $F_{n-1}$-measurable random variable chosen so that the discounted asset price is a $(Q, (F_n))$-martingale.

Written as an equation, the random variable $\alpha^{(n-1)}$ must satisfy,

$$
E^Q[S_t e^{-rt} | F_{n-1}] = S_{n-1}.
$$

(5.4)

This condition holds for $n = 1, 2, \ldots, N$, and ensures that the measure, $Q$, is indeed a martingale measure.

We will show in Theorem 5.2.2 that such variables, $\alpha^{(n-1)}$, exist, and explain how they can be found. The proof of this theorem uses the $F_n$-measurability of the random variable $\xi_n$, which follows from writing the Radon-Nikodym derivative as,

$$
\frac{dQ}{dP}igg|_{F_n} = \xi_n = Y_0 Y_1 \cdots Y_n.
$$

(This can be shown by induction).

**Theorem 5.2.2.** Let $Q$ be the Esscher martingale measure of Definition 5.2.1. For each $n = 1, 2, \ldots, N$, the $F_{n-1}$-measurable Esscher parameter, $\alpha^{(n-1)}$, exists, is unique, and can be computed as the inner product,

$$
\alpha^{(n-1)} = \langle \alpha, Z_{n-1} \rangle,
$$

(5.5)
where $\alpha = (\alpha_1, \ldots, \alpha_M)^T$ is a vector of solutions to the equations,
\[
\sum_{j=1}^{M} A_{ji} e^{\alpha_i (\mu_j - \frac{1}{2} \sigma_j^2) r + \frac{1}{2} \sigma_j^2 \alpha_i} \left\{ e^{\mu_j r + \sigma_j^2 r \alpha_i} - e^{r \tau} \right\} = 0,
\]
for $i = 1, 2, \ldots, M$.

Proof. Fix $n \in \{1, 2, \ldots, N\}$ and let $\alpha^{(n-1)} = \alpha$. Using Bayes rule, the martingale condition (5.4) becomes,
\[
\frac{E[\xi_{n-1} Y_n S_n e^{-r \tau} | F_{n-1}]}{E[\xi_{n-1} Y_n | F_{n-1}]} = S_{n-1}.
\]
As $\xi_{n-1}$ is $F_{n-1}$-measurable, $E[Y_n | F_{n-1}] = 1$ and $S_n = S_{n-1} e^{X_n}$, an equivalent expression is,
\[
E \left[ Y_n S_{n-1} e^{X_n} e^{-r \tau} | F_{n-1} \right] = S_{n-1}.
\]
Substituting the definition of $Y_n$ given in Equation (5.3), we see that $\alpha$ must satisfy,
\[
E \left[ e^{(\alpha+1)X_n} | F_{n-1} \right] = e^{r \tau} E \left[ e^{\alpha X_n} | F_{n-1} \right].
\]
Using the Markov property of the joint process, $\{(S_n, Z_n)\}$, and setting $S_{n-1} = s$ and $Z_{n-1} = e_i$ in the $\sigma$-algebra $F_{n-1}$, the conditional expectations in this equation can be calculated using Lemma 2.3.3 of Chapter 2. This gives,
\[
\sum_{j=1}^{M} A_{ji} e^{\alpha (\mu_j - \frac{1}{2} \sigma_j^2) r + \frac{1}{2} \sigma_j^2 \alpha} \left\{ e^{\mu_j r + \sigma_j^2 r \alpha} - e^{r \tau} \right\} = 0.
\]
Thus we see that $\alpha$ depends only on $e_i$, the state of $Z_{n-1}$, which justifies Equation (5.5).

To prove the existence of such an $\alpha$, denote by $\Psi(\alpha)$ the continuous function on the left hand side of Equation (5.7). It is clear that $\Psi(\alpha) > 0$ if $\mu_j + \sigma_j^2 \alpha > r$ for all $j$, or equivalently, $\alpha > \max_j \left( \frac{r - \mu_j}{\sigma_j^2} \right) = \overline{\alpha}$. Similarly, $\Psi(\alpha) < 0$ if $\alpha < \min_j \left( \frac{r - \mu_j}{\sigma_j^2} \right) = \underline{\alpha}$. The intermediate value theorem then proves the existence of an $\alpha \in (\underline{\alpha}, \overline{\alpha})$ such that Equation (5.7) holds.

It remains to show that $\alpha$ is unique. To do so, notice that $Y_n$, defined by Equation (5.3), can be written alternatively as,
\[
Y_n = \frac{e^{\alpha (X_n - r \tau)}}{E \left[ e^{\alpha (X_n - r \tau)} | F_{n-1} \right]}.
\]
Rewriting Equation (5.6) using this form of \( Y_n \), we see that \( \alpha \) must satisfy the equation,

\[
\Psi(\alpha) := \mathbb{E}\left[ e^{\alpha (X_n - rT)} \left( e^{X_n - rT} - 1 \right) \bigg| \mathcal{F}_{n-1} \right] = 0.
\]

The derivative of the function \( \Psi \) is,

\[
\Psi'(\alpha) = \mathbb{E}\left[ (X_n - rT) e^{\alpha (X_n - rT)} \left( e^{X_n - rT} - 1 \right) \bigg| \mathcal{F}_{n-1} \right],
\]

and as both \( X_n - rT \) and \( e^{X_n - rT} - 1 \) have the same sign, and \( e^{\alpha (X_n - rT)} \) is always positive, \( \Psi'(\alpha) \) is non-negative and therefore \( \Psi \) is increasing. This shows that \( \alpha \) is unique, completing the proof. \( \square \)

We see in the above theorem and its proof how each Esscher parameter, \( \alpha^{(n-1)} \), can be computed numerically. First, the vector \( \alpha = (\alpha_1, ..., \alpha_M)^T \) can be found using the method of interval bisection, as each \( \alpha_i \) is the unique point in the interval \((\underline{\alpha}, \overline{\alpha})\),

\[
\underline{\alpha} = \min_j \left( \frac{r - \mu_j}{\sigma_j^2} \right), \quad \overline{\alpha} = \max_j \left( \frac{r - \mu_j}{\sigma_j^2} \right),
\]

which satisfies the equation,

\[
\sum_{j=1}^{M} A_{ji} e^{\alpha_i (\mu_j - \frac{1}{2} \sigma_j^2 T) + \frac{1}{4} \sigma_j^2 T \alpha_i^2} \left\{ e^{\mu_j T + \sigma_j^2 T \alpha_i} - e^{rT} \right\} = 0.
\]

The required Esscher parameter is then computed from the inner product,

\[
\alpha^{(n-1)} = \langle \alpha, Z_{n-1} \rangle.
\]

Summarising, in this section we have given a precise definition of the Esscher martingale measure in a Switching Black–Scholes market, and a method for computing the unknown parameters involved in the definition. Next, we need to find the joint conditional density function of \((S_n, Z_n)\) given \( S_0 = s \) and \( Z_0 = c_i \) under the Esscher martingale measure, and use this to price contingent claims on assets in a Switching Black–Scholes market. The following section gives some preliminaries needed for these calculations.
5.3 Preliminary Definitions and Results

As for the method of mean-variance hedging, the Esscher price of a contingent claim in a Switching Black-Scholes market can be computed either by repeatedly conditioning down on the filtration, or by finding the required density function. We focus on the second method of computation, and in this section give some preliminaries needed to find the required density function. These definitions and results are comparable to those in Section 4.1 of Chapter 4, where preliminaries to finding density functions for the mean-variance hedging method of pricing were given.

We begin by recursively defining two real-valued functions, $E_n$ and $K_n$, which are similar to those given in Definition 4.1.1, and then use induction to find their general form.

Recall the convention that an empty sum is zero and the notation $j_{m,n}$ for the string $j_m, j_{m+1}, \ldots, j_n$.

**Definition 5.3.1.** Let $\alpha = (\alpha_1, \ldots, \alpha_M)^T$ be the vector of Esscher parameters defined in Theorem 5.2.2. For $n = 0, 1, \ldots, N - 1$, define the functions $E_{n+1} : \{1, 2, \ldots, M\}^{n+2} \rightarrow \mathbb{R}$ and $K_{n+1} : \{1, 2, \ldots, M\}^{n+1} \rightarrow \mathbb{R}$ by the recursions,

$$E_{n+1}(j_{1,n+2}) = E_n(j_{1,n+1}) + (\mu_{j_{n+2}} - \frac{1}{2}\sigma_{j_{n+2}}^2)\tau + (\alpha_{j_n} - \alpha_{j_{n+1}})K_n(j_{2,n+1})^2,$$

$$K_{n+1}(j_{1,n+1}) = \sqrt{\sigma_{j_1}^2\tau + K_n(j_{2,n+1})^2},$$

where the initial functions are $E_1(j_{1,2}) = (\mu_{j_2} - \frac{1}{2}\sigma_{j_2}^2)\tau$ and $K_1(j_1) = \sigma_{j_1}\sqrt{\tau}$.

The function $K_n$ defined above is the same as the function $K_n$ given in Definition 4.1.1 of Chapter 4, and the function $E_n$ is similar to the $E_n$ of Chapter 4, except it involves an extra term, $(\alpha_{j_n} - \alpha_{j_{n+1}})K_n(j_{2,n+1})^2$. This additional term appears due to the form of the Esscher martingale measure, as will become evident as the chapter progresses. The following lemma gives the general form of $E_n$ and $K_n$. 
Lemma 5.3.2. For \( n = 1, 2, \ldots, N \), the functions \( E_n \) and \( K_n \) defined in Equations (5.8) and (5.9) are given by,

\[
E_n(j_{1,n+1}) = \sum_{i=2}^{n+1} (\mu_{j_i} - \frac{1}{2}\sigma_{j_i}^2)\tau + \sum_{i=1}^{n-1} \alpha_{j_i}\sigma_{j_{i+1}}^2 \tau - \alpha_{j_n} K_{n-1}(j_{2,n})^2,
\]

\[
K_n(j_{1,n}) = (\sigma_{j_1}^2 + \sigma_{j_2}^2 + \cdots + \sigma_{j_n}^2)^{1/2} \sqrt{\tau}.
\]

Here, we use the convention that an empty sum is zero, and \( K_0 = 0 \).

Proof. The equation for \( K_n \) has already been proven in Lemma 4.1.2.

To prove the equation for \( E_n \), we use induction on \( n \). By definition, the result is true for \( n = 1 \). Suppose that it is true for arbitrary \( n \). The recursive definition allows us to write,

\[
E_{n+1}(j_{1,n+2}) = E_n(j_{1,n+1}) + (\mu_{j_{n+2}} - \frac{1}{2}\sigma_{j_{n+2}}^2)\tau + (\alpha_{j_n} - \alpha_{j_{n+1}}) K_n(j_{2,n+1})^2.
\]

Since the induction hypothesis states that,

\[
E_n(j_{1,n+1}) = \sum_{i=2}^{n+1} (\mu_{j_i} - \frac{1}{2}\sigma_{j_i}^2)\tau + \sum_{i=1}^{n-1} \alpha_{j_i}\sigma_{j_{i+1}}^2 \tau - \alpha_{j_n} K_{n-1}(j_{2,n})^2,
\]

and \( K_n(j_{2,n+1})^2 - K_{n-1}(j_{2,n})^2 = \sigma_{j_{n+1}}^2 \tau \), the result follows.

Using the above definition and lemma, we can now determine the following properties of the functions \( E_n \) and \( K_n \).

Lemma 5.3.3. For \( n = 2, 3, \ldots, N \), the functions \( E_n \) and \( K_n \), defined in 5.3.1, satisfy the equations,

\[
\alpha_{j_n} E_n(j_{1,n+1}) + \frac{1}{2} \alpha_{j_n}^2 K_n(j_{2,n+1})^2 \\
+ (\alpha_{j_{n-1}} - \alpha_{j_n}) E_{n-1}(j_{1,n}) + \frac{1}{2}(\alpha_{j_{n-1}} - \alpha_{j_n})^2 K_{n-1}(j_{2,n})^2 \\
= \alpha_{j_{n-1}} E_{n-1}(j_{1,n}) + \frac{1}{2} \alpha_{j_{n-1}}^2 K_{n-1}(j_{2,n})^2 \\
+ \alpha_{j_n} (\mu_{j_{n+1}} - \frac{1}{2}\sigma_{j_{n+1}}^2) \tau + \frac{1}{2} \sigma_{j_{n+1}}^2 \alpha_{j_{n+1}}^2 \tau,
\]

(5.10)
and,

\[
\sum_{p=n}^{N} \left\{ \left( \alpha_{j_{p-1}} - \alpha_{j_{p}} \right) E_{p-1}(j_{1,p}) + \frac{1}{2} \left( \alpha_{j_{p-1}} - \alpha_{j_{p}} \right)^2 K_{p-1}(j_{2,p})^2 \right\} \\
+ \alpha_{j_{N}} E_{N}(j_{1,N+1}) + \frac{1}{2} \alpha_{j_{N}}^2 K_{N}(j_{2,N+1})^2 \\
= \sum_{i=n+1}^{N+1} \left\{ \alpha_{j_{i-1}} \left( \mu_{j_{i}} - \frac{1}{2} \sigma_{j_{i}}^2 \right) \tau + \frac{1}{2} \sigma_{j_{i}}^2 \alpha_{j_{i-1}}^2 \right\} \\
+ \alpha_{j_{N-1}} E_{n-1}(j_{1,n}) + \frac{1}{2} \alpha_{j_{N-1}}^2 K_{n-1}(j_{2,n})^2,
\]  

(5.11)

where \( \alpha = (\alpha_1, \alpha_2, ..., \alpha_M)^T \) is the vector of Esscher parameters defined in Theorem 5.2.2.

Proof. To prove Equation (5.10), we use the recursive relationship given in Definition 5.3.1 to write \( E_n \) in terms of \( E_{n-1} \):

\[
E_n(j_{1,n+1}) = E_{n-1}(j_{1,n}) + (\mu_{j_{n+1}} - \frac{1}{2} \sigma_{j_{n+1}}^2) \tau + (\alpha_{j_{n-1}} - \alpha_{j_{n}}) K_{n-1}(j_{2,n})^2.
\]

The left hand side of Equation (5.10) then becomes,

\[
\alpha_{j_{n}} \left( \mu_{j_{n+1}} - \frac{1}{2} \sigma_{j_{n+1}}^2 \right) \tau + \alpha_{j_{n-1}} E_{n-1}(j_{1,n}) \\
+ \frac{1}{2} \left( \alpha_{j_{n-1}}^2 - \alpha_{j_{n}}^2 \right) K_{n-1}(j_{2,n})^2 + \frac{1}{2} \alpha_{j_{n}}^2 K_{n}(j_{2,n+1})^2.
\]

Using the identity, \( K_{n}(j_{2,n+1})^2 - K_{n-1}(j_{2,n})^2 = \sigma_{j_{n+1}}^2 \tau \), which follows from Lemma 5.3.2, this equals,

\[
\alpha_{j_{n-1}} E_{n-1}(j_{1,n}) + \frac{1}{2} \alpha_{j_{n-1}}^2 K_{n-1}(j_{2,n})^2 + \alpha_{j_{n}} \left( \mu_{j_{n+1}} - \frac{1}{2} \sigma_{j_{n+1}}^2 \right) \tau + \frac{1}{2} \sigma_{j_{n+1}}^2 \alpha_{j_{n}}^2,
\]

as required.

We now prove Equation (5.11) using backward induction on the integer \( n \). For convenience, let \( f(n) \) denote the left hand side of Equation (5.11). At \( n = N \), we have,

\[
f(N) = (\alpha_{j_{N-1}} - \alpha_{j_{N}}) E_{N-1}(j_{1,N}) + \frac{1}{2} (\alpha_{j_{N-1}} - \alpha_{j_{N}})^2 K_{N-1}(j_{2,N})^2 \\
+ \alpha_{j_{N}} E_{N}(j_{1,N+1}) + \frac{1}{2} \alpha_{j_{N}}^2 K_{N}(j_{2,N+1})^2.
\]

Use of the identity given in Equation (5.10) proves the result for this case.
Next, suppose that the equation holds for an arbitrary \( n + 1 \). Then, the value at \( n \) is,

\[
f(n) = f(n+1) + (\alpha_{j_{n-1}} - \alpha_{j_n})E_{n-1}(j_{1,n}) + \frac{1}{2}(\alpha_{j_{n-1}} - \alpha_{j_n})^2K_{n-1}(j_{2,n})^2.
\]

We can then apply Equation (5.10) at \( n \) and use the induction hypothesis to obtain the result. \( \square \)

Next, we introduce some new notation, which is used in the subsequent lemma and will simplify the expressions involved in finding the density functions in the next section.

**Definition 5.3.4.** Let \( \mathbf{\alpha} = (\alpha_1, ..., \alpha_M)^T \) be the vector of Esscher parameters defined in Theorem 5.2.2, and let \( E_n \) and \( K_n \) be the functions defined in 5.3.1. For \( n = 2, 3, ..., N \), we define the function \( a_n : \{1, 2, ..., M\}^{n-1} \to \mathbb{R} \) by,

\[
a_n(j_{1,n}) = \exp \left\{ (\alpha_{j_{n-1}} - \alpha_{j_n})E_{n-1}(j_{1,n}) + \frac{1}{2}(\alpha_{j_{n-1}} - \alpha_{j_n})^2K_{n-1}(j_{2,n})^2 \right\}.
\]

We conclude this section with a lemma evaluating an integral which appears when computing the density functions required to calculate Esscher contingent claim prices. This lemma is analogous to Lemma 4.1.4 of Chapter 4.

**Lemma 5.3.5.** Let \( \mathbf{\alpha} = (\alpha_1, ..., \alpha_M)^T \) be the vector of Esscher parameters defined in Theorem 5.2.2, \( n \in \{1, 2, ..., N - 1\} \), \( E_n \) and \( K_n \) the functions of Definition 5.3.1, and \( a_{n+1} \) the function of Definition 5.3.4. Then for \( x, s \in (0, \infty) \) and \( j_{1,n+2} \in \{1, 2, ..., M\}^{n+2} \), we have,

\[
\int_0^\infty \frac{1}{K_n(j_{2,n+1})} \frac{1}{\sqrt{2\pi}x^2} \left( \frac{\alpha_{j_n} - \alpha_{j_{n+1}}}{\sigma_{j_{n+2}}} \right)^2 e^{-\frac{1}{2} \left( \frac{\log(x) - \mu_{j_{n+2}} - \frac{1}{2}\sigma_{j_{n+2}}^2}{\sigma_{j_{n+2}}} \right)^2} dz
\]

\[
= \sqrt{2\pi} \frac{a_{n+1}(j_{1,n+1})}{K_{n+1}(j_{2,n+2})} s \left( \frac{\alpha_{j_n} - \alpha_{j_{n+1}}}{e} \right) e^{-\frac{1}{2} \left( \frac{\log(x) - \mu_{j_{n+2}} - \frac{1}{2}\sigma_{j_{n+2}}^2}{K_{n+1}(j_{2,n+2})} \right)^2}.
\]

**Proof.** The proof is analogous to the proof of Lemma 4.1.4. Letting \( A = \log x - (\mu_{j_{n+2}} - \frac{1}{2}\sigma_{j_{n+2}}^2) \tau \), \( B = \log s + E_n(j_{1,n+1}) \), and \( y = \log z \),
the integral becomes,

\[ \int_{-\infty}^{\infty} \frac{1}{K_n(j_{2,n+1})C^{\frac{1}{2}}} \left( \rho_{j_{n+1}} - \rho_{j_{n+2}} \right) \left( \frac{y-B}{\sigma_{j_{n+2}}\sqrt{T}} \right)^2 dy. \]

Consulting the proof of Lemma 4.1.4, with \( j_1 \) replaced by \( j_{n+2} \), we have,

\[ \int_{-\infty}^{\infty} \frac{1}{K_n(j_{2,n+1})C^{\frac{1}{2}}} e^{-\frac{1}{2} \left( \frac{y-A}{\sigma_{j_{n+2}}\sqrt{T}} \right)^2} \left\{ \left( \frac{y-S}{\tau} \right)^2 + T - S^2 \right\} \left( \rho_{j_{n+1}} - \rho_{j_{n+2}} \right) dy, \]

where,

\[ S = \frac{K_n(j_{2,n+1})^2 A + \sigma_{j_{n+2}}^2 \tau B}{K_{n+1}(j_{2,n+1})^2}, \]

and \( T = \frac{K_n(j_{2,n+1})^2 A^2 + \sigma_{j_{n+2}}^2 \tau B^2}{K_{n+1}(j_{2,n+1})^2}. \]

Next, change variables to \( z = \frac{K_{n+1}(j_{2,n+1})^2}{\sigma_{j_{n+2}}\sqrt{T}K_n(j_{2,n+1})^2} y \) to obtain,

\[ \int_{-\infty}^{\infty} \frac{1}{K_{n+1}(j_{2,n+2})C^{\frac{1}{2}}} e^{-\frac{1}{2} \left( \frac{K_{n+1}(j_{2,n+2})^2}{\sigma_{j_{n+2}}\sqrt{T}K_n(j_{2,n+1})^2} \right)^2} \left\{ \frac{K_{n+1}(j_{2,n+2})^2}{\sigma_{j_{n+2}}\sqrt{T}K_n(j_{2,n+1})^2} \left( \rho_{j_{n+1}} - \rho_{j_{n+2}} \right) \right\} \]

Again from the proof of Lemma 4.1.4, with \( j_1 \) replaced by \( j_{n+2} \),

\[ \frac{K_{n+1}(j_{2,n+2})^2}{\sigma_{j_{n+2}}^2\tau K_n(j_{2,n+1})^2} \left[ T - S^2 \right] = \left( \frac{A - B}{K_{n+1}(j_{2,n+2})} \right)^2. \]

The remaining part of the exponent is \(-1/2 \) times,

\[ \left( z - \frac{K_{n+1}(j_{2,n+2})}{\sigma_{j_{n+2}}\sqrt{T}K_n(j_{2,n+1})} S \right)^2 - 2\sigma_{j_{n+2}}\sqrt{T}K_n(j_{2,n+1}) \left( \rho_{j_{n+1}} - \rho_{j_{n+2}} \right) \]

\[ = (z - D)^2 + \frac{K_{n+1}(j_{2,n+2})^2}{\sigma_{j_{n+2}}^2\tau K_n(j_{2,n+1})^2} S^2 - D^2, \]

where,

\[ D = \frac{K_{n+1}(j_{2,n+2})}{\sigma_{j_{n+2}}\sqrt{T}K_n(j_{2,n+1})} S + \frac{\sigma_{j_{n+2}}\sqrt{T}K_n(j_{2,n+1})}{K_{n+1}(j_{2,n+2})} \left( \rho_{j_{n+1}} - \rho_{j_{n+2}} \right). \]

Expanding the \( D^2 \) term, the integral becomes,

\[ \int_{-\infty}^{\infty} \frac{1}{K_{n+1}(j_{2,n+2})C^{\frac{1}{2}}} e^{-\frac{1}{2} \left( z - D \right)^2} e^{-\frac{1}{2} \left( \frac{A - B}{K_{n+1}(j_{2,n+2})} \right)^2} \]

\[ \times e^{-\frac{1}{2} \left( -2(\rho_{j_{n+1}} - \rho_{j_{n+2}}) S - \frac{\sigma_{j_{n+2}}^2\tau K_n(j_{2,n+1})^2}{K_{n+1}(j_{2,n+2})^2} \left( \rho_{j_{n+1}} - \rho_{j_{n+2}} \right)^2 \right)} dz. \]
Next, change variables to \( u = z - D \) and use the expression for \( S \) to obtain,

\[
\frac{\sqrt{2\pi}}{K_{n+1}(j_{2n+2})^x} e^{-\frac{1}{2} \left( \frac{(A - B)^2}{2} - 2(\alpha_{j_n} - \alpha_{j_{n+1}})K_n(j_{2n+1})^2 - 2\alpha_{j_n} \frac{\sigma^2_{j_{n+2}} \tau B}{K_{n+1}(j_{2n+2})} - (\alpha_{j_n} - \alpha_{j_{n+1}})^2 \frac{\sigma^2_{j_{n+2}} \tau K_n(j_{2n+1})}{K_{n+1}(j_{2n+2})^2} \right)}.
\]

The numerator in the exponential here is,

\[
A^2 - 2A \left( B + (\alpha_{j_n} - \alpha_{j_{n+1}})K_n(j_{2n+1})^2 \right) + B^2 - 2(\alpha_{j_n} - \alpha_{j_{n+1}}) \sigma^2_{j_{n+2}} \tau B - (\alpha_{j_n} - \alpha_{j_{n+1}})^2 \sigma^2_{j_{n+2}} \tau K_n(j_{2n+1})^2,
\]

which equals,

\[
(A - B - (\alpha_{j_n} - \alpha_{j_{n+1}})K_n(j_{2n+1})^2)^2 + B^2 - 2(\alpha_{j_n} - \alpha_{j_{n+1}}) \sigma^2_{j_{n+2}} \tau B - (\alpha_{j_n} - \alpha_{j_{n+1}})^2 K_n(j_{2n+1})^2 K_{n+1}(j_{2n+2})^2.
\]

From the recursive definition of \( K_{n+1} \), this becomes,

\[
(A - B - (\alpha_{j_n} - \alpha_{j_{n+1}})K_n(j_{2n+1})^2)^2 - 2B(\alpha_{j_n} - \alpha_{j_{n+1}})K_{n+1}(j_{2n+2})^2 - (\alpha_{j_n} - \alpha_{j_{n+1}})^2 K_n(j_{2n+1})^2 K_{n+1}(j_{2n+2})^2.
\]

Also, the definitions of \( E_{n+1}, A \) and \( B \) give,

\[
A - B - (\alpha_{j_n} - \alpha_{j_{n+1}})K_n(j_{2n+1})^2 = \log(x/s) - (\mu_{j_{n+2}} - \frac{1}{2} \sigma^2_{j_{n+2}}) \tau - E_n(j_{1,n+1}) - (\alpha_{j_n} - \alpha_{j_{n+1}})K_n(j_{2n+1})^2 = \log(x/s) - E_{n+1}(j_{1,n+2}).
\]

Using \( B = \log s + E_n(j_{1,n+1}) \) and the definition of \( a_{n+1} \) given in 5.3.4, the integral equals,

\[
\frac{\sqrt{2\pi}}{K_{n+1}(j_{2n+2})^x} e^{-\frac{1}{2} \left( \frac{\log(x/s) - E_{n+1}(j_{1,n+2})}{K_{n+1}(j_{2n+2})} \right)^2} \\
\times e^{(\alpha_{j_n} - \alpha_{j_{n+1}}) \log(s + E_n(j_{1,n+1}))} e^{\frac{1}{2} (\alpha_{j_n} - \alpha_{j_{n+1}})^2 K_n(j_{2n+1})^2} = \sqrt{2\pi} \frac{a_{n+1}(j_{1,n+1})}{K_{n+1}(j_{2n+2})^x} (\alpha_{j_n} - \alpha_{j_{n+1}}) e^{-\frac{1}{2} \left( \frac{\log(x/s) - E_{n+1}(j_{1,n+2})}{K_{n+1}(j_{2n+2})} \right)^2},
\]

as required. \( \square \)
5.4 Density Functions and Option Pricing

Having given the preliminary definitions and results in the previous section, we now compute the Esscher price of a call option on an asset which evolves according to the Switching Black–Scholes model. This involves three steps. First, the joint conditional density function, \( \psi_n \), of \((S_n, Z_n)\) given \( \mathcal{F}_{n-1} \) under the Esscher martingale measure must be found. This can be computed directly. Secondly, the recursive relationship,

\[
\phi_n(s, e_i, x, e_j) = \sum_{l=1}^{M} \int_{-\infty}^{\infty} \psi_n(u, e_l, x, e_j) \phi_{n-1}(s, e_i, u, e_l) \, du,
\]
given in Theorem 2.2.5 of Chapter 2 is used to find \( \phi_n \), the joint conditional density function of \((S_n, Z_n)\) given \( S_0 = s \) and \( Z_0 = e_i \) under the Esscher martingale measure. The Esscher price of an \( \mathcal{F}_n \)-measurable contingent claim, \( H_n \), at time 0 is then,

\[
H_0(s, e_i) = e^{-\tau n} \sum_{j=1}^{M} \int_{-\infty}^{\infty} \phi_n(s, e_i, x, e_j) H_n(x, e_j) \, dx.
\]

This integral is evaluated in Theorem 5.4.3, in the case where \( H_n \) is a call option on an asset which evolves according to the Switching Black–Scholes model. As a corollary to this theorem, we show that when there is no switching, (that is, \( M = 1 \)), the Esscher call price is equal to the Black-Scholes call price.

We begin by computing the general form of the density function, \( \psi_n \).

**Theorem 5.4.1.** For any \( n \in \{1, 2, \ldots, N\} \), the joint conditional density function of \((S_n, Z_n)\) given \( S_{n-1} = s \) and \( Z_{n-1} = e_i \) under the Esscher martingale measure is,

\[
\psi_n(s, e_i, x, e_j) = \left\{
\begin{array}{ll}
\frac{1}{c(e_i) \sqrt{2\pi}} \left( \frac{x}{s} \right)^{\alpha_i} & \frac{1}{\sigma_j \sqrt{\pi x}} e^{-\frac{1}{2} \left( \frac{\log(s/x) - [\mu_j - \frac{1}{2} \sigma_j^2] x}{\sigma_j \sigma_{ij}} \right)^2}, \ x > 0 \\
0, & \ x \leq 0,
\end{array}
\right.
\]
for $e_j \in \mathcal{H}$ and $x \in \mathbb{R}$. Here, the $\mathcal{F}_{n-1}$-measurable random variable, $c$, is defined by,

$$c(e_i) = \sum_{j=1}^{M} A_{ji} e^{\alpha_i (\mu_j - \frac{1}{2}\sigma_j^2) x + \frac{1}{2}\sigma_j^2 x^2},$$

and $\alpha = (\alpha_1, \ldots, \alpha_M)^T$ is the vector of Esscher parameters defined in Theorem 5.2.2. Notice that $\psi_n$ is independent of $n$.

**Proof.** Let $V_n$ be an arbitrary $\mathcal{F}_n$-measurable random variable, and denote the Esscher martingale measure by $Q$. Using Bayes rule and the definition of $Q$, we have,

$$E^Q[ V_n | \mathcal{F}_{n-1}] = \frac{E \left[ \frac{dQ}{dP} \bigg| \mathcal{F}_n, V_n \right]}{E \left[ \frac{dQ}{dP} \bigg| \mathcal{F}_n \right]} \frac{E[V_n | \mathcal{F}_{n-1}]}{E[V_n | \mathcal{F}_{n-1}]} = \frac{E[Y_n V_n | \mathcal{F}_{n-1}]}{E[Y_n | \mathcal{F}_{n-1}]}.$$  

To simplify this expression, recall that,

$$Y_n = \frac{e^{\alpha (n-1) X_n}}{E[e^{\alpha (n-1) X_n} | \mathcal{F}_{n-1}]} = \frac{e^{\alpha (Z_{n-1}) X_n}}{E[e^{\alpha (Z_{n-1}) X_n} | \mathcal{F}_{n-1}]} = \left( \frac{S_n}{S_{n-1}} \right)^{\alpha (Z_{n-1})},$$

and so $E[Y_n | \mathcal{F}_{n-1}] = 1$. Also, write,

$$e^{\alpha (Z_{n-1}) X_n} = (e^{X_n})^{\alpha (Z_{n-1})} = \left( \frac{S_n}{S_{n-1}} \right)^{\alpha (Z_{n-1})},$$

for the exponential in the numerator. We then have,

$$E^Q[ V_n | \mathcal{F}_{n-1}] = \frac{E \left[ \left( \frac{S_n}{S_{n-1}} \right)^{\alpha (Z_{n-1})} V_n \bigg| \mathcal{F}_{n-1} \right]}{E \left[ \left( \frac{S_n}{S_{n-1}} \right)^{\alpha (Z_{n-1})} X_n \bigg| \mathcal{F}_{n-1} \right]}.$$  

The denominator can be evaluated using Lemma 2.3.3 of Chapter 2:

$$E[e^{\alpha (Z_{n-1}) X_n} | S_{n-1} = s, Z_{n-1} = e_i] = \sum_{j=1}^{M} A_{ji} e^{\alpha_i (\mu_j - \frac{1}{2}\sigma_j^2) x + \frac{1}{2}\sigma_j^2 x^2} = c(e_i).$$

Finally, using the joint conditional density function of $(S_n, Z_n)$ given $S_{n-1} = s$ and $Z_{n-1} = e_i$, from Theorem 2.2.1 of Chapter 2, we obtain,

$$E^Q[ V_n | S_{n-1} = s, Z_{n-1} = e_i] = \frac{1}{c(e_i)} \sum_{j=1}^{M} \int_{0}^{\infty} \frac{A_{ji}}{\sqrt{2\pi \sigma_j^2}} \left( \frac{x}{s} \right)^{\alpha_i e_i} e^{-\frac{1}{2} \left( \frac{\sigma_j^2}{\sigma_j^2} \left( \frac{x}{s} \right)^2 \right)} V_n(x, e_j) \, dx.$$  

Choosing $V_n(S_n, Z_n) = I(S_n \leq s', Z_n = e_j')$ gives the result. $\Box$
Having found the density function, \( \psi_n \), we now use the recursion,

\[
\phi_1(s, e_i, x, e_j) = \psi_1(s, e_i, x, e_j),
\]

\[
\phi_n(s, e_i, x, e_j) = \sum_{k=1}^{M} \int_{-\infty}^{\infty} \phi_{n-1}(s, e_i, u, e_k)\psi_n(u, e_i, x, e_j) \, du,
\]

to find the density function, \( \phi_n \), explicitly. The following theorem expresses \( \phi_n \) in terms of the functions \( E_n, K_n \) and \( a_p \), for \( p = 2, 3, \ldots, n \), which were defined in Section 5.3.

**Theorem 5.4.2.** For \( n \in \{2, 3, \ldots, N\} \), the joint conditional density function of \( (S_n, Z_n) \) given \( S_0 = s \) and \( Z_0 = e_{j_1} \) under the Esscher martingale measure is,

\[
\phi_n(s, e_{j_1}, x, e_{j_{n+1}}) = \sum_{j_2, j_3, \ldots, j_n=1}^{M} \frac{1}{c(e_{j_1})} \frac{A_{j_1j_2}A_{j_2j_3} \cdots A_{j_{n-1}j_n}A_{j_nj_{n+1}}}{\sqrt{2\pi}} \times \frac{1}{K_n(j_2, j_{n+1})} x \alpha_n (\frac{s}{x}) \frac{1}{K_n(j_2, j_{n+1})} \frac{1}{\sigma_{j_1} \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\log(s/x) - E_n(j_1, n+1)}{\sigma_{j_1} \sigma_{j_2}} \right)^2} \prod_{p=2}^{n} \frac{a_p(j_1, p)}{c(e_{j_p})},
\]

for \( e_{j_{n+1}} \in \mathcal{H} \) and \( x > 0 \), and \( \phi_n(s, e_i, x, e_{j_1}) = 0 \) for \( x \leq 0 \). For \( n = 1 \), we have \( \phi_1 = \psi_1 \).

In this theorem, \( \alpha = (\alpha_1, \ldots, \alpha_M)^T \) is the vector of Esscher parameters defined in Theorem 5.2.2, and the functions \( c, a_p, E_n \) and \( K_n \) are defined in 5.4.1, 5.3.4 and 5.3.1, respectively.

**Proof.** The proof is by induction on \( n \). For \( n = 2 \), Theorems 2.2.5 and 5.4.1 give,

\[
\phi_2(s, e_{j_1}, x, e_{j_2}) = \sum_{j_2=1}^{M} \int_{-\infty}^{\infty} \phi_1(s, e_{j_1}, z, e_{j_2})\psi_2(z, e_{j_2}, x, e_{j_2}) \, dz
\]

\[
= \sum_{j_2=1}^{M} \int_{0}^{\infty} \frac{1}{c(e_{j_1})} \frac{A_{j_1j_2}}{\sqrt{2\pi}} \frac{(z)}{s} \alpha_{j_1} \frac{1}{\sigma_{j_2} \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\log(z/s) - E_n(j_1, j_2)}{\sigma_{j_1} \sigma_{j_2}} \right)^2} \times \frac{1}{c(e_{j_2})} \frac{A_{j_2j_3}}{\sqrt{2\pi}} \frac{1}{\sigma_{j_2} \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\log(x/z) - E_n(j_2, j_3)}{\sigma_{j_2} \sigma_{j_3}} \right)^2} \, dz.
\]
Rearranging and using Lemma 5.3.5, we have,

\[
\phi_2(s, e_j, x, e_{j+1}) = \sum_{j_1, j_2 = 1}^M \frac{1}{c(e_{j_1})} A_{j_1j_2} A_{j_2j_1} (x) \alpha_{j_2} \frac{1}{K_2(j_2, x)} e^{-\frac{1}{2} \left( \frac{\log(s/e_j) - E_{\alpha_{j_2} + \alpha_{j_1}}}{K_2(j_2, x)} \right)^2} a_2(j_1, j_2) \frac{c(e_{j_2})}{c(e_{j_1})},
\]

which is as the Theorem states.

Now suppose the result is true for arbitrary \( n \). From Theorem 2.2.5,

\[
\phi_{n+1}(s, e_j, x, e_{j+n+2}) = \sum_{j_{n+1} = 1}^M \int_{-\infty}^{\infty} \phi_n(s, e_j, z, e_{j+n+1}) \psi_{n+1}(z, e_{j+n+1}, x, e_{j+n+2}) \, dz.
\]

Applying the induction hypothesis and Theorem 5.4.1, this equals,

\[
\sum_{j_2, j_3, \ldots, j_{n+1} = 1}^M \int_{0}^{\infty} \frac{1}{c(e_{j_1})} \frac{A_{j_1j_2} A_{j_2j_3} \cdots A_{j_{n-1}j_n} A_{j_{n+1}j_n} A_{j_{n+2}j_1}}{2\pi} \times \left( \frac{z}{s} \right)^{\alpha_{j_n}} \frac{1}{K_n(j_{n+1}, z)} e^{-\frac{1}{2} \left( \frac{\log(s/e_j) - E_{\alpha_{j_{n+1}} - \frac{\alpha_{j_{n+2}}}{\alpha_{j_{n+2}}}}}{K_n(j_{n+1}, z)} \right)^2} \prod_{p=2}^n \frac{a_p(j_1, p)}{c(e_{j_p})} \int_{0}^{\infty} \frac{1}{c(e_{j_1})} \frac{A_{j_1j_2} A_{j_2j_3} \cdots A_{j_{n-1}j_n} A_{j_{n+1}j_n} A_{j_{n+2}j_1}}{2\pi} \times \left( \frac{x}{s} \right)^{\alpha_{j_{n+1}}} \frac{1}{K_n(j_{n+1}, z)} e^{-\frac{1}{2} \left( \frac{\log(s/e_j) - E_{\alpha_{j_{n+1}} - \frac{\alpha_{j_{n+2}}}{\alpha_{j_{n+2}}}}}{K_n(j_{n+1}, z)} \right)^2} \psi_{n+1}(x, e_{j+n+2}) \, dx \, dz.
\]

Simplifying and rearranging, we obtain,

\[
\sum_{j_2, j_3, \ldots, j_{n+1} = 1}^M \frac{1}{c(e_{j_1})} \frac{A_{j_1j_2} A_{j_2j_3} \cdots A_{j_{n-1}j_n} A_{j_{n+1}j_n} A_{j_{n+2}j_1}}{2\pi} \times \left( \frac{x}{s} \right)^{\alpha_{j_{n+1}}} \int_{0}^{\infty} \left[ \frac{1}{\sigma_{j_{n+2}} \sqrt{\pi x}} e^{-\frac{1}{2} \left( \frac{\log(s/e_j) - E_{\alpha_{j_{n+1}} - \frac{\alpha_{j_{n+2}}}{\alpha_{j_{n+2}}}}}{\sigma_{j_{n+2}} \sqrt{\pi x}} \right)^2} \times \frac{1}{K_n(j_{n+1}, z)} e^{-\frac{1}{2} \left( \frac{\log(s/e_j) - E_{\alpha_{j_{n+1}} - \frac{\alpha_{j_{n+2}}}{\alpha_{j_{n+2}}}}}{K_n(j_{n+1}, z)} \right)^2} (\alpha_{j_n} - \alpha_{j_{n+1}}) \right] \, dx \, dz.
\]

We can invoke Lemma 5.3.5 to compute this integral, which gives,

\[
\phi_{n+1}(s, e_j, x, e_{j+n+2}) = \sum_{j_2, j_3, \ldots, j_{n+1} = 1}^M \frac{1}{c(e_{j_1})} \frac{A_{j_1j_2} A_{j_2j_3} \cdots A_{j_{n-1}j_n} A_{j_{n+1}j_n} A_{j_{n+2}j_1}}{2\pi} \times \left( \frac{x}{s} \right)^{\alpha_{j_{n+1}}} \frac{a_n(j_1, n+1)}{K_n(j_{n+1}, z)} e^{-\frac{1}{2} \left( \frac{\log(s/e_j) - E_{\alpha_{j_{n+1}} - \frac{\alpha_{j_{n+2}}}{\alpha_{j_{n+2}}}}}{K_n(j_{n+1}, z)} \right)^2},
\]

as required. \( \square \)
As mentioned in the introduction to this chapter, the Esscher price of an $\mathcal{F}_N$-measurable contingent claim, $H_N$, at time 0, given initial asset price, $S_0$, and initial market state, $e_{j_1}$, is,

$$H_0(S_0, e_{j_1}) = \mathbb{E}^Q \left[ e^{-rT} H_N \right]_{S_0, Z_0 = e_{j_1}} = e^{-rT} \sum_{j_{N+1} = 1}^{M} \int_{-\infty}^{\infty} \phi_N(S_0, e_{j_1}, x, e_{j_{N+1}}) H_N(x, e_{j_{N+1}}) \, dx.$$ 

We can now use the form of the joint conditional density function, $\phi_N$, given in Theorem 5.4.2, to express the price as,

$$H_0(S_0, e_{j_1}) = e^{-rT} \sum_{j_2j_3\ldots j_{N+1} = 1}^{M} \frac{A_{j_{N+1}j_N} A_{j_Nj_{N-1}} \ldots A_{j_2j_1} \prod_{p=2}^{N} a_p(j_1,p)}{\sqrt{2\pi}} \times \int_{0}^{\infty} \left( x \frac{1}{S_0} \right)^{\alpha_{j_N}} \frac{1}{K_N(j_{2,N+1})} x^{-1} \left( \frac{\log(S_0/\mu) + \sigma^2}{K_N(j_{2,N+1})} \right)^2 H_N(x, e_{j_{N+1}}) \, dx.$$  

The following theorem evaluates this integral in the situation where $H_N$ is a call option on the asset, $S$, which expires at time $T$.

**Theorem 5.4.3.** Let $S$ be an asset which evolves according to the Switching Black–Scholes model, with initial asset price, $S_0$, and initial market state, $e_{j_1}$. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_M)^T$ be the vector of Esscher parameters defined in Theorem 5.2.2. The Esscher price of a call option expiring at time $T$ with strike price, $K$, on the asset, $S$, is,

$$H_0(S_0, e_{j_1}) = \sum_{j_2j_3\ldots j_{N+1} = 1}^{M} p(j_1, j_2, \ldots, j_{N+1}) \times \left( S_0 \exp \left\{ \sum_{i=1}^{N} \left[ (\mu_{j_{i+1}} - r)\tau + \alpha_{j_i}\sigma_{j_{i+1}}^2 \tau \right] \right\} N(d_j^+ + \alpha_{j_N} K_N(j_{2,N+1})) - K e^{-rT} N(d_j^- + \alpha_{j_N} K_N(j_{2,N+1})) \right), \quad (5.13)$$

where,

$$p(j_1, j_2, \ldots, j_{N+1}) = \prod_{i=1}^{N} \frac{A_{j_{i+1}j_i} e^{\alpha_{j_i}(\mu_{j_{i+1}} - \frac{1}{2}\sigma_{j_{i+1}}^2)\tau + \frac{1}{2}\alpha_{j_{i+1}}\sigma_{j_{i+1}}^2\tau}}{\sum_{j_{i+1}=1}^{M} A_{j_{i+1}j_i} e^{\alpha_{j_i}(\mu_{j_{i+1}} - \frac{1}{2}\sigma_{j_{i+1}}^2)\tau + \frac{1}{2}\alpha_{j_{i+1}}\sigma_{j_{i+1}}^2\tau}}.$$
is a probability (that is, \( \sum_{j=1, j \neq j_1}^{M} p(j_1, j_2, \ldots, j_{N+1}) = 1 \) and
\[ 0 < p(j_1, j_2, \ldots, j_{N+1}) < 1, \]
\[
d_j^+ = \frac{\log(S_0/K) + \sum_{i=2}^{N+1} \left( \mu_{j_i} \pm \frac{1}{2} \sigma_{j_i}^2 \right) \tau + \sum_{i=1}^{N-1} (\alpha_{j_i} - \alpha_{j_{i+1}}) \sigma_{j_{i+1}}^2 \tau}{K_N(j_2, j_{N+1})}, \quad (5.14)
\]
and
\[
K_N(j_2, j_{N+1}) = \left( \sigma_{j_2}^2 + \sigma_{j_3}^2 + \ldots + \sigma_{j_{N+1}}^2 \right)^{1/2} \sqrt{\tau}.
\]
Proof. The value of a call option with strike price \( K \) at expiry is,
\[
H_N(x, e_{j_{N+1}}) = (x - K)^+, \]
so in this case, the integral in Equation (5.12) becomes,
\[
\int_{-d_j^-}^{\infty} \left( x \right)^{\alpha_{j_N}} \frac{K_N(j_2, j_{N+1})}{x} \left( \frac{\log(x/S_0) - E_N(j_1, j_{N+1})}{K_N(j_2, j_{N+1})} \right)^2 e^{-\frac{1}{2}y^2} dy.
\]
Using the identity,
\[
\left( \frac{x}{S_0} \right)^{\alpha_{j_N}} (x - K) = S_0 \left( \frac{x}{S_0} \right)^{\alpha_{j_N} + 1} - K \left( \frac{x}{S_0} \right)^{\alpha_{j_N}},
\]
and changing variables to \( y = \frac{\log(x/S_0) - E_N(j_1, j_{N+1})}{K_N(j_2, j_{N+1})} \), we have,
\[
\int_{-d_j^-}^{\infty} \left( S_0 \exp \left\{ \left( E_N(j_1, j_{N+1}) + K_N(j_2, j_{N+1})y \right) \left( \alpha_{j_N} + 1 \right) \right\} - K \exp \left\{ \left( E_N(j_1, j_{N+1}) + K_N(j_2, j_{N+1})y \right) \alpha_{j_N} \right\} \right) e^{-\frac{1}{2}y^2} dy,
\]
where
\[
d_j^- = \frac{E_N(j_1, j_{N+1}) + \log(S_0/K)}{K_N(j_2, j_{N+1})}. \]
This integral can be evaluated using Lemma A.1.3 of Appendix 1, which gives,
\[
\sqrt{2\pi} \left[ \left( \frac{E_N(j_1, j_{N+1}) (\alpha_{j_N} + 1)}{2} \right) \right. \times \left. N(d_j^- + K_N(j_2, j_{N+1}) (\alpha_{j_N} + 1)) \right.
\]
\[
- K \exp \left\{ E_N(j_1, j_{N+1}) \alpha_{j_N} + \frac{1}{2} K_N(j_2, j_{N+1})^2 \alpha_{j_N}^2 \right\} N(d_j^- + K_N(j_2, j_{N+1}) \alpha_{j_N}).
\]
Thus the call price is,

\[
H_0(S_0, e_{j_1}) = \frac{e^{-rT}}{c(e_{j_1})} \sum_{j_2, j_3, \ldots, j_{N+1}}^{M} A_{j_N+1,j_N} A_{j_{N-1},j_{N-2}} \cdots A_{j_2,j_1} \prod_{p=2}^{N} \frac{a_p(j_1,p)}{c(e_{j_p})} \\
\times \exp \left\{ \alpha_{j_N} E_N(j_{1,N+1}) + \frac{1}{2} \alpha_{j_N}^2 K_N(j_{2,N+1})^2 \right\} \\
\times \left( S_0 e^{E_N(j_{1,N+1}) + \frac{1}{2} K_N(j_{1,N+1})^2}(2\alpha_{j_N} + 1) N\left( d_j^+ + K_N(j_{2,N+1})\alpha_{j_N} \right) - K_N(d_j^- + K_N(j_{2,N+1})\alpha_{j_N}) \right),
\]

where \( d_j^+ = d_j^- + K_N(j_{2,N+1}) \). Using the equations for \( E_N \) and \( K_N \) given in Lemma 5.3.2, the equations for \( d_j^+ \) and \( d_j^- \) given in (5.14) follow.

Next, consider the product,

\[
\prod_{p=2}^{N} a_p(j_1,p) \exp \left\{ \alpha_{j_N} E_N(j_{1,N+1}) + \frac{1}{2} \alpha_{j_N}^2 K_N(j_{2,N+1})^2 \right\}.
\]

From the definition of \( a_p \) given in 5.3.4, the exponent in this expression is,

\[
\sum_{p=2}^{N} \left\{ (\alpha_{j_{p-1}} - \alpha_{j_p}) E_{p-1}(j_{1,p}) + \frac{1}{2} (\alpha_{j_{p-1}} - \alpha_{j_p})^2 K_{p-1}(j_{2,p})^2 \right\} \\
+ \alpha_{j_N} E_N(j_{1,N+1}) + \frac{1}{2} \alpha_{j_N}^2 K_N(j_{2,N+1})^2,
\]

which, from Lemma 5.3.3, is equal to,

\[
\sum_{i=3}^{N+1} \left\{ \alpha_{j_{i-1}} (\mu_{j_i} - \frac{1}{2} \sigma_{j_i}^2) \tau + \frac{1}{2} \sigma_{j_i}^2 \tau \alpha_{j_{i-1}}^2 \right\} + \alpha_{j_1} E_1(j_{1,2}) + \frac{1}{2} \alpha_{j_1}^2 K_1(j_2)^2 \\
= \sum_{i=1}^{N} \left\{ \alpha_{j_i} (\mu_{j_{i+1}} - \frac{1}{2} \sigma_{j_{i+1}}^2) \tau + \frac{1}{2} \sigma_{j_{i+1}}^2 \tau \alpha_{j_i}^2 \right\}.
\]

Also, using Lemma 5.3.2,

\[
E_N(j_{1,N+1}) + \frac{1}{2} K_N(j_{2,N+1})^2(2\alpha_{j_N} + 1) \\
= \sum_{i=2}^{N+1} (\mu_{j_i} - \frac{1}{2} \sigma_{j_i}^2) \tau + \sum_{i=1}^{N-1} \alpha_{j_i} \sigma_{j_{i+1}}^2 \tau + \frac{1}{2} K_N(j_{2,N+1})^2 \\
- \alpha_{j_N} K_{N-1}(j_{2,N})^2 + \alpha_{j_N} K_N(j_{2,N+1})^2.
\]
As $K_n(j_{2,n+1})^2 = (\sigma_{j_2}^2 + \ldots + \sigma_{j_{n+1}}^2)\tau$, we have,

$$E_N(j_{1,N+1}) + \frac{1}{2}K_n(j_{2,N+1})^2(2\alpha_{j_N} + 1) = \sum_{i=2}^{N+1}{\mu_{j_i}\tau} + \sum_{i=1}^{N}{\alpha_{j_i}\sigma_{j_{i+1}}^2\tau},$$

and subtracting $rT = \sum_{i=1}^{N}r\tau$ gives,

$$-rT + E_N(j_{1,N+1}) + \frac{1}{2}K_n(j_{2,N+1})^2(2\alpha_{j_N} + 1) = \sum_{i=1}^{N}\left\{ (\mu_{j_{i+1}} - r)\tau + \alpha_{j_i}\sigma_{j_{i+1}}^2\tau \right\}.$$

Hence the call price is,

$$H_0(S_0, e_{j_1}) = \sum_{j_2,j_3,\ldots,j_{N+1}=1}^{M} \prod_{i=1}^{N}A_{j_{i+1},j_i}e^{\alpha_{j_i}(\mu_{j_{i+1}} - \frac{1}{2}\sigma_{j_{i+1}}^2)\tau + \frac{1}{2}\sigma_{j_{i+1}}^2\tau\alpha_{j_i}^2}c(e_{j_i})$$

$$\times \left( S_0 \exp\left\{ \sum_{i=1}^{N}\left[ (\mu_{j_{i+1}} - r)\tau + \alpha_{j_i}\sigma_{j_{i+1}}^2\tau \right]\right\} \right. N\left( d_1^{j_i} + K_n(j_{2,N+1})\alpha_{j_N} \right)$$

$$- \left. K e^{-rT}N\left( d_2^{j_i} + K_n(j_{2,N+1})\alpha_{j_N} \right) \right).$$

The result is then immediate, as from Theorem 5.4.1,

$$c(e_{j_i}) = \sum_{j_{i+1}=1}^{M} A_{j_{i+1},j_i}e^{\alpha_{j_i}(\mu_{j_{i+1}} - \frac{1}{2}\sigma_{j_{i+1}}^2)\tau + \frac{1}{2}\sigma_{j_{i+1}}^2\tau\alpha_{j_i}^2}.$$

To conclude the chapter, we consider the special case, $M = 1$, of the Switching Black–Scholes model, so that the Markov chain, $Z$, is constant, and the asset, $S$, evolves according to the ordinary Black–Scholes model,

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t).$$

The following corollary shows that for this case, the Esscher call price equals the Black-Scholes call price.

**Corollary 5.4.4.** The Esscher price of a call option expiring at time $T$ with strike price, $K$, on an asset, $S$, which evolves according to the Switching Black–Scholes model with $M = 1$ and initial price, $S_0$, is,

$$H_0(S_0) = S_0 N(d_1) - Ke^{-rT}N(d_2),$$
where,
\[
d_1 = \frac{\log(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma \sqrt{T}},
\]
and \(d_2 = d_1 - \sigma \sqrt{T}\).

Proof. When \(M = 1\), the vector, \(\alpha\), of Esscher parameters is just a scalar, \(\alpha\), satisfying,
\[
e^{\alpha(\mu - \frac{1}{2}\sigma^2)T + \frac{1}{2}\sigma^2 \alpha^2} \left\{ e^{\mu T + \sigma^2 \alpha^2} - e^{rT} \right\} = 0.
\]
This implies that \(\mu T + \sigma^2 \alpha^2 = rT\), and hence,
\[
\alpha = \frac{r - \mu}{\sigma^2}.
\]

From Theorem 5.4.3, the Esscher price of the call option described above is,
\[
H_0(S_0) = S_0 \exp \left\{ \sum_{i=1}^{N} \left[ (\mu - r)\tau_i + \alpha \sigma^2 \tau_i \right] \right\} N \left( d_j^+ + \alpha \sigma \sqrt{T} \right) - Ke^{-rT} N \left( d_j^- + \alpha \sigma \sqrt{T} \right),
\]
as, for \(M = 1\), the probability \(p(j_1, j_2, \ldots, j_{N+1})\) defined in Theorem 5.4.3 equals 1, and \(K_N(j_2, N+1) = \sigma \sqrt{T}\). Using \(\alpha = \frac{r - \mu}{\sigma^2}\), we have,
\[
\sum_{i=1}^{N} \left[ (\mu - r)\tau_i + \alpha \sigma^2 \tau_i \right] = (\mu - r)T + \frac{r - \mu}{\sigma^2} \sigma^2 T = 0,
\]
and from Equation (5.14),
\[
d_j^+ + \alpha \sigma \sqrt{T} = \frac{\log(S_0/K) + (\mu \pm \frac{1}{2}\sigma^2)T}{\sigma \sqrt{T}} + \frac{r - \mu}{\sigma \sqrt{T}} = \frac{\log(S_0/K) + (r \pm \frac{1}{2}\sigma^2)T}{\sigma \sqrt{T}}.
\]
Therefore,
\[
H_0(S_0) = S_0 N (d_1) - Ke^{-rT} N (d_2),
\]
as required.
Chapter 6

Option Pricing via Minimum Entropy

In previous chapters, we have discussed option pricing when the asset price follows the Switching Black–Scholes model, using both the mean-variance hedging and Esscher methods. In this chapter, we develop option pricing formulae using the minimum entropy method to choose a martingale measure.

In Section 6.1 we discuss the history of entropy and the motivation for using the method of minimum entropy to price options. Previous work in this area is also mentioned. Mathematical terminology is developed in Section 6.2, where we define and prove the existence of the minimum entropy martingale measure and find its general structure. Section 6.3 then explains how the minimum entropy martingale measure can be found.

Finally, in Section 6.4, option pricing via minimum entropy is considered. Due to the complex form of the martingale measure, as will become apparent, we focus on a few special cases, one of which assumes a discrete approximation to the asset price.

Throughout the chapter, we make the assumption that the asset price follows the Switching Black–Scholes model and,

\[ E[S_k | \mathcal{F}_{k-1}] \geq e^{rr} S_{k-1}, \]
for \( k = 1, 2, \ldots, N \). This condition is satisfied, for example, when all the possible asset drifts are greater than the risk-free rate. We write the Switching Black–Scholes asset price as,

\[
S_k = S_{k-1}e^{X_k},
\]

where \( X_k = (\mu(Z_k) - \frac{1}{2}\sigma(Z_k)^2)\tau + \sigma(Z_k)\Delta W_k \) (from Equation (2.3)).

### 6.1 Why Entropy?

Entropy techniques were first developed by E. Jaynes ([51], 1957) in a statistical physics setting, and have since been applied to a variety of areas, including chemistry ([1]), pattern recognition ([53], [54]), and computer system modelling ([32]). The applications make use of the principle of maximum entropy, which provides a method for estimating an unknown probability distribution of a random variable on the probability space \( \{x_i : i \in \mathbb{N}\} \), given certain expectations or bounds on these expectations. The principle states that the estimate \( q \) of the unknown distribution should be chosen to maximise the entropy,

\[
-\sum_i q(x_i) \log q(x_i),
\]

given the constraints.

Later, (see for example [37], [57]), this principle was generalised to include the case when a prior estimate \( p \) of the unknown distribution is given. The so-called principle of minimum cross-entropy, or relative entropy, states that the estimate \( q \) should be chosen to minimise the relative entropy,

\[
\sum_i q(x_i) \log (q(x_i)/p(x_i)).
\]

This can be generalised to the case of continuous probability distributions, and it is this formulation which is used to choose the minimum entropy martingale measure.

At first glance, this may seem a rather arbitrary choice for the distribution estimate. Jaynes, in [52], justifies the method by explaining that the maximum entropy distribution “agrees with what is known, but expresses ‘maximum uncertainty’ with respect to all other matters”. More recently, it was shown by
Shore and Johnson ([76]) that optimising any function other than entropy will lead to inconsistency, unless the maxima of that function and the entropy function agree. Further justification for choosing the minimum entropy martingale measure for pricing comes from its correspondence to pricing by utility maximisation, with the utility function,

\[ u(x) = \frac{1}{a} e^{-a x}, \]

for some \( a > 0 \) (see Frittelli [33]).

Pricing via minimum entropy has been investigated from numerous viewpoints. Frittelli ([33]) and Grandits and Rheinländer ([38]) give alternative characterisations of the minimum entropy martingale measure, but these formulations are often not practical. Authors including Buchen and Kelly ([11]) and Avellaneda ([3]) derive a minimum entropy martingale measure which is consistent with a set of observed option prices.

More relevant to this thesis are the option pricing formulae developed using the minimum entropy criterion. Work in this area includes that by Miyahara ([62]), in the case when the asset price follows a geometric Lévy process (that is, \( S_t = S_0 e^{X_t} \), where \( X_t \) is a Lévy process), and by Czyli ([42]), when the asset can only take discrete values. The remainder of this chapter develops option pricing formulae using the Switching Black–Scholes model and minimum entropy criterion.

### 6.2 The Minimum Entropy Martingale Measure

The relative entropy of a probability measure \( Q \) with respect to a probability measure \( P \) is defined as,

\[
H(Q|P) = \begin{cases} 
    E\left[\frac{dQ}{dP} \ln \left(\frac{dQ}{dP}\right)\right] & \text{if } Q \ll P, \\
    \infty & \text{otherwise.}
\end{cases}
\] (6.1)
The minimum entropy martingale measure is chosen to minimise the relative entropy with respect to the real world measure. A mathematical formulation is given in the following definition.

**Definition 6.2.1.** For the Switching Black–Scholes model, the minimum entropy martingale measure, $Q$, is defined by,

$$
\frac{dQ}{dP}\bigg|_{\mathcal{F}_n} = \tilde{\Lambda}_1 \tilde{\Lambda}_2 \ldots \tilde{\Lambda}_n, \quad \text{for } n = 1, \ldots, N,
$$

where $\tilde{\Lambda}_k$ minimises the relative entropy,

$$
E[\Lambda_k \ln \Lambda_k | \mathcal{F}_{k-1}],
$$

subject to the constraints,

$$
\begin{align*}
\Lambda_k &> 0, \\
E[\Lambda_k | \mathcal{F}_{k-1}] &= 1, \\
E[\Lambda_k S_k | \mathcal{F}_{k-1}] &= e^{rr} S_{k-1}.
\end{align*}
$$

We check later that this is a valid martingale measure.

From the definition, we see that the minimum entropy Radon-Nikodym derivative is defined to be a product of one-period Radon-Nikodym derivatives, each chosen to minimise the relative entropy in one time period. For the Switching Black–Scholes model, this is more natural than minimising the entropy over the entire time span, as each time period involves a change in the market. A possible avenue for further study would be to find the relationship between these two formulations.

The following theorem shows that each $\tilde{\Lambda}_k$ is of exponential form – a well known property of entropy minimisation. An alternative proof in the case of a discrete asset price is given in Gzyl ([42]).

**Theorem 6.2.2.** Let $k \in \{1, 2, \ldots, N\}$. Define,

$$
\tilde{\Lambda} = \frac{e^{-\lambda_k S_k}}{E[e^{-\lambda_k S_k} | \mathcal{F}_{k-1}]},
$$
where \( \lambda_k \geq 0 \) is the unique \( \mathcal{F}_{k-1} \)-measurable random variable such that,

\[
E\left[(S_k - e^{r_k} S_{k-1})e^{-\lambda_k(S_k - e^{r_k} S_{k-1})} \bigg| \mathcal{F}_{k-1}\right] = 0,
\]

(6.3)

and \( S \) is an asset evolving according to the Switching Black–Scholes model. Then \( \tilde{\Lambda} \) minimises,

\[
E[\Lambda \ln \Lambda | \mathcal{F}_{k-1}],
\]

subject to,

\[
\Lambda > 0,
\]

\[
E[\Lambda | \mathcal{F}_{k-1}] = 1,
\]

(6.4)

\[
E[\Lambda S_k | \mathcal{F}_{k-1}] = e^{r_k} S_{k-1}.
\]

Proof. We show that there exists a unique \( \mathcal{F}_{k-1} \)-measurable \( \lambda_k \geq 0 \) satisfying Equation (6.3) in Lemma 6.2.3.

We check that \( \tilde{\Lambda} \) actually satisfies the conditions (6.4). Clearly, \( \tilde{\Lambda} > 0 \) as long as the expectation in the denominator is finite. This holds as \( \lambda_k \geq 0 \). The condition \( E[\Lambda | \mathcal{F}_{k-1}] = 1 \) follows immediately from the definition of \( \tilde{\Lambda} \), and \( E[\Lambda S_k | \mathcal{F}_{k-1}] = e^{r_k} S_{k-1} \) is established from a rearrangement of Equation (6.3).

We proceed to show that \( \tilde{\Lambda} \) is the solution to the minimisation problem. Let \( f(x) = x \ln x \). Then \( f'(x) = 1 + \ln x \) and \( f''(x) = 1/x > 0 \) for \( x > 0 \). Therefore \( f \) is convex on \([0, \infty)\), and so,

\[
f(x) \geq f(y) + (x - y)f'(y),
\]

for all \( x, y \in [0, \infty) \). Setting \( x \) to be any \( \Lambda \) satisfying the conditions (6.4), \( y = \tilde{\Lambda} \) and taking expectations gives,

\[
E[\Lambda \ln \Lambda | \mathcal{F}_{k-1}] \geq E\left[\Lambda \ln \tilde{\Lambda} \bigg| \mathcal{F}_{k-1}\right] + E\left[(\Lambda - \tilde{\Lambda})(1 + \ln \tilde{\Lambda}) \bigg| \mathcal{F}_{k-1}\right].
\]

Because \( E[\Lambda - \tilde{\Lambda} | \mathcal{F}_{k-1}] = E[\Lambda | \mathcal{F}_{k-1}] - E[\tilde{\Lambda} | \mathcal{F}_{k-1}] = 1 - 1 = 0 \), the second term becomes \( E[(\Lambda - \tilde{\Lambda}) \ln \tilde{\Lambda} | \mathcal{F}_{k-1}] \). Using the form of \( \tilde{\Lambda} \) and the conditions (6.4) we
see that,
\[
\begin{align*}
E \left[ (\Lambda - \bar{\Lambda}) \ln \bar{\Lambda} \middle| \mathcal{F}_{k-1} \right] \\
= E \left[ (\Lambda - \bar{\Lambda}) \left( -\lambda_k S_k - \ln E \left[ e^{-\lambda_k S_k} \middle| \mathcal{F}_{k-1} \right] \right) \middle| \mathcal{F}_{k-1} \right] \\
= -\lambda_k E \left[ (\Lambda - \bar{\Lambda}) S_k \middle| \mathcal{F}_{k-1} \right] - \ln E \left[ e^{-\lambda_k S_k} \middle| \mathcal{F}_{k-1} \right] E \left[ (\Lambda - \bar{\Lambda}) \middle| \mathcal{F}_{k-1} \right] \\
= 0.
\end{align*}
\]

Therefore,
\[
E [\Lambda \ln \Lambda \middle| \mathcal{F}_{k-1}] \geq E \left[ \bar{\Lambda} \ln \bar{\Lambda} \middle| \mathcal{F}_{k-1} \right],
\]
and \( \bar{\Lambda} \) minimises \( E [\Lambda \ln \Lambda \middle| \mathcal{F}_{k-1}] \) subject to the conditions (6.4).

In the proof of this theorem, we made use of the following lemma.

**Lemma 6.2.3.** Consider a Switching Black–Scholes market as defined in Chapter 2. Then for \( k \in \{1, 2, \ldots, N\} \), the equation,
\[
E \left[ (S_k - e^{r^*} S_{k-1}) e^{-(\lambda_k - e^{r^*}) S_k} \middle| \mathcal{F}_{k-1} \right] = 0, \tag{6.5}
\]
has a unique \( \mathcal{F}_{k-1} \)-measurable solution \( \lambda_k \geq 0 \).

**Proof.** Let \( g(\lambda) = E \left[ (S_k - e^{r^*} S_{k-1}) e^{-\lambda (S_k - e^{r^*} S_{k-1})} \middle| \mathcal{F}_{k-1} \right] \). Then,
\[
g'(\lambda) = E \left[ -(S_k - e^{r^*} S_{k-1})^2 e^{-\lambda (S_k - e^{r^*} S_{k-1})} \middle| \mathcal{F}_{k-1} \right] < 0,
\]
so \( g \) is decreasing. Also,
\[
g(\lambda) = E \left[ I(S_k > e^{r^*} S_{k-1})(S_k - e^{r^*} S_{k-1}) e^{-\lambda (S_k - e^{r^*} S_{k-1})} \middle| \mathcal{F}_{k-1} \right] + E \left[ I(S_k < e^{r^*} S_{k-1})(S_k - e^{r^*} S_{k-1}) e^{-\lambda (S_k - e^{r^*} S_{k-1})} \middle| \mathcal{F}_{k-1} \right] \\
:= J_1(\lambda) + J_2(\lambda).
\]

As \( \lambda \to \infty \), \( J_1(\lambda) \to 0 \) and \( J_2(\lambda) \to -\infty \), so \( g(\lambda) \to -\infty \). Similarly, as \( \lambda \to -\infty \), \( J_1(\lambda) \to \infty \) and \( J_2(\lambda) \to 0 \), so that \( g(\lambda) \to \infty \). Therefore there is a unique \( \lambda_k \) such that \( g(\lambda_k) = 0 \). The assumption \( E[S_k \middle| \mathcal{F}_{k-1}] \geq e^{r^*} S_{k-1} \) ensures \( \lambda_k \geq 0 \), as \( g(0) = E[S_k \middle| \mathcal{F}_{k-1}] - e^{r^*} S_{k-1} \).
To complete the proof, we need to show that \( \lambda_k \) is \( \mathcal{F}_{k-1} \)-measurable. Substituting \( S_k = S_{k-1} e^{X_k} \), where \( X_k = (\mu(Z_k) - \frac{1}{2} \sigma(Z_k)^2) \tau + \sigma(Z_k) \Delta W_k \), into Equation (6.5), we obtain,

\[
E \left[ S_{k-1} (e^{X_k} - e^{\tau r}) e^{-\lambda_k S_{k-1} (e^{X_k} - e^{\tau r})} \left| \mathcal{F}_{k-1} \right. \right] = 0.
\]

Equivalently, with \( \alpha^{(k-1)} = \lambda_k S_{k-1} \), conditioning first on \( Z_k = e_j \), and setting \( S_{k-1} = s \) and \( Z_{k-1} = e_i \) in the \( \sigma \)-algebra \( \mathcal{F}_{k-1} \), we require that,

\[
\sum_{j=1}^M A_{ji} \int_{-\infty}^{\infty} \left( e^{(\mu_j - \frac{1}{2} \sigma_j^2) \tau + \sigma_j \sqrt{\tau} x} - e^{r \tau} \right) \\
\times e^{-\alpha^{(k-1)}} \left( e^{(\mu_j - \frac{1}{2} \sigma_j^2) \tau + \sigma_j \sqrt{\tau} x} - e^{\tau r} \right) e^{-\frac{1}{2} x^2} dx = 0. \tag{6.6}
\]

From this equation we can see that the solution \( \alpha^{(k-1)} \) depends only on \( e_i = Z_{k-1} \), and therefore \( \lambda_k = \alpha^{(k-1)}/S_{k-1} \) is \( \mathcal{F}_{k-1} \)-measurable. \( \square \)

We now check that Definition 6.2.1 defines a valid martingale measure. From the definition and Theorem 6.2.2 we have,

\[
\frac{dQ}{dP} \bigg|_{\mathcal{F}_n} = \tilde{\Lambda}_1 \tilde{\Lambda}_2 \ldots \tilde{\Lambda}_n,
\]

where,

\[
\tilde{\Lambda}_k = \frac{e^{-\lambda_k S_k}}{E \left[ e^{-\lambda_k S_k} \left| \mathcal{F}_{k-1} \right. \right]}, \quad k = 1, 2, ..., N, \tag{6.7}
\]

and \( \lambda_k \geq 0 \) is the unique \( \mathcal{F}_{k-1} \)-measurable random variable such that,

\[
E \left[ (S_k - e^{r \tau} S_{k-1}) e^{-\lambda_k (S_k - e^{r \tau} S_{k-1})} \left| \mathcal{F}_{k-1} \right. \right] = 0.
\]

For \( Q \) to be a martingale measure, we need the discounted asset price to be a \((Q, (\mathcal{F}_k))\)-martingale and \( E \left[ \frac{dQ}{dP} \right] = 1 \). We check these conditions.

The definition and Bayes rule gives,

\[
E^Q[S_{k+1}|\mathcal{F}_k] = \frac{E \left[ \tilde{\Lambda}_1 \ldots \tilde{\Lambda}_k \tilde{\Lambda}_{k+1} S_{k+1} \left| \mathcal{F}_k \right. \right]}{E \left[ \tilde{\Lambda}_1 \ldots \tilde{\Lambda}_k \tilde{\Lambda}_{k+1} \left| \mathcal{F}_k \right. \right]}.
\]

From Equation (6.7), \( \tilde{\Lambda}_1 \tilde{\Lambda}_2 \ldots \tilde{\Lambda}_k \) is \( \mathcal{F}_k \)-measurable and thus cancels out. The constraints of the definition then show that the denominator equals 1 and the
numerator is $e^{rt}S_k$, so $E^Q[S_{k+1}|\mathcal{F}_k] = e^{rt}S_k$. Similarly, $E\left[\frac{dQ}{dP}\right] = 1$ follows from $\frac{dQ}{dP} = \lambda_1\lambda_2\ldots\lambda_N$ and the property that $E\left[\lambda_k | \mathcal{F}_{k-1}\right] = 1$.

Thus Definition 6.2.1 indeed defines a martingale measure, $Q$. In the next section, we discuss how the change of measure can be found.

### 6.3 Finding the Change of Measure

We have shown that the minimum entropy martingale measure is defined by the equation,

$$\frac{dQ}{dP} = \prod_{k=1}^{N} \frac{e^{-\lambda_k S_k}}{E\left[e^{-\lambda_k S_k} | \mathcal{F}_{k-1}\right]},$$

where the predictable variables $\lambda_k$ are chosen such that,

$$E\left[(S_k - e^{rt}S_{k-1})e^{-\lambda_k(S_{k-1} - e^{rt}S_{k-1})} | \mathcal{F}_{k-1}\right] = 0. \quad (6.8)$$

To completely specify the change of measure, we need to compute the variables $\lambda_k$ for $k = 1, 2, \ldots, N$.

Due to the structure of Equation (6.8), it is more convenient to work with the variables $\alpha^{[k-1]} = \lambda_k S_{k-1}$. The following theorem expresses the change of measure in terms of these variables.

**Theorem 6.3.1.** For the Switching Black–Scholes asset price model, the minimum entropy martingale measure, $Q$, can be computed from,

$$\frac{dQ}{dP} = \prod_{k=1}^{N} \frac{e^{-i\alpha_k Z_{k-1}}e^{X_k}}{E\left[e^{-i\alpha_k Z_{k-1}}e^{X_k} | \mathcal{F}_{k-1}\right]}, \quad (6.9)$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_M)^T$ is the vector of entropy parameters and $\alpha_i$ is the solution to the equation,

$$\sum_{j=1}^{M} A_{ji} \int_{-\infty}^{\infty} \left(e^{i(\mu_j - \frac{1}{2}\sigma_j^2)x + \frac{1}{2}\sigma_j^2x - e^{rt}} \right) \times e^{-\alpha_i(e^{i(\mu_j - \frac{1}{2}\sigma_j^2)x + \frac{1}{2}\sigma_j^2x} - e^{rt})} e^{-\frac{1}{2}x^2} dx = 0, \quad (6.10)$$

for $i = 1, 2, \ldots, M$. 

Proof. From Equation (6.6) in the proof of Lemma 6.2.3, \( \alpha^{(k-1)} = \lambda_k S_{k-1} \) satisfies the equation,

\[
\sum_{j=1}^{M} A_{ji} \int_{-\infty}^{\infty} \left( e^{(\mu_j - \frac{1}{2}\sigma_j^2)x + \sigma_j \sqrt{\tau}x - e^{\gamma x}} \right) e^{-\alpha^{(k-1)}(e^{(\mu_j - \frac{1}{2}\sigma_j^2)x + \sigma_j \sqrt{\tau}x - e^{\gamma x}} - e^{\frac{1}{2}x^2}} dx = 0,
\]

where \( Z_{k-1} = e_i \). The result is immediate from the definition of \( \alpha_i \) and the relation \( S_k = S_{k-1} e^{X_k} \).

We have shown in Lemma 6.2.3 that a unique solution \( \alpha_i \) to Equation (6.10) exists, but it is not clear how this solution can be found. An analytical solution is not obvious, due to the form of the integrand. Instead, we must turn to numerical methods. In particular, we can use the method of interval bisection, with initial points \( \overline{\alpha} = 0 \), which gives a positive value of the integral, and \( \underline{\alpha} \), chosen to give a negative value of the integral (such an \( \alpha \) exists as the integral is decreasing to \( -\infty \)).

### 6.4 Option Pricing

Having discussed the definition of the minimum entropy martingale measure and found its structure, we now consider pricing a contingent claim, \( H_N \), using this method. At time 0, the minimum entropy price of the contingent claim is,

\[
H_0 = E^Q \left[ H_N e^{-\gamma T} \big| \mathcal{F}_0 \right].
\]

Using Bayes rule and the change of measure given in Theorem 6.3.1, we obtain,

\[
H_0 = E \left[ \prod_{k=1}^{N} \frac{e^{-(\alpha_k, Z_{k-1})e^{X_k}}}{e^{-(\alpha_k, Z_{k-1})e^{X_k}} \left| \mathcal{F}_{k-1} \right]} H_N e^{-\gamma T} \bigg| \mathcal{F}_0 \right]. \tag{6.11}
\]

As in the mean-variance and Esscher settings, it is theoretically possible to calculate this price in two ways. Method 1 introduces intermediate prices \( H_{n-1} = E^Q \left[ H_n e^{-\gamma T} \big| \mathcal{F}_{n-1} \right] \) for \( n = 1, 2, \ldots, N \), and calculates the price by conditioning back one step at a time. Method 2 calculates the joint conditional
density function of \((S_N, Z_N)\) given \(\mathcal{F}_0\) under the martingale measure, \(Q\), using a forward recursion procedure. Both have been demonstrated in the preceding chapters. However, if we apply either of these procedures using the minimum entropy martingale measure, we repeatedly encounter integrals of the form,

\[
\int_{-\infty}^{\infty} e^{-ae^{bx} - \frac{1}{2}b^2x^2} \, dx,
\]

for \(a, b > 0\). An analytical solution to this integral is not known, which makes pricing by the minimum entropy method difficult and impractical for the general Switching Black–Scholes model.

We could overcome this problem by using numerical integration, as will be discussed in Section 6.4.1 for \(N = 1\). Unfortunately, in general we need to calculate similar integrals at least \(N\) times, which is computationally infeasible due to exponential explosion of the number of computations required.

Another option is to approximate the integrand by some other suitable function for which we can compute the integral. This is a topic for further investigation.

A third alternative is to replace the asset price by an approximation giving rise to integrals which can be calculated analytically. In Section 6.4.2 we consider this approach by discretising the asset price.

### 6.4.1 Pricing Over One Time Step (ie. \(N = 1\))

When there is only one time step, the price on a contingent claim \(H_n\) at time \(n-1\) is,

\[
H_{n-1}(s, \epsilon_i) = \mathbb{E}^Q[H_n e^{-rT} | S_{n-1} = s, Z_{n-1} = \epsilon_i].
\]  

(6.12)

As shown in the following lemma, the joint conditional density function, \(\psi_n\), of \((S_n, Z_n)\) given \(S_{n-1} = s\) and \(Z_{n-1} = \epsilon_i\) under the martingale measure, \(Q\), can easily be calculated. This can then be applied to compute the price (6.12).

**Lemma 6.4.1.** Let \(n \in \{1, 2, \ldots, N\}\). The joint conditional density function of \((S_n, Z_n)\) given \(S_{n-1} = s\) and \(Z_{n-1} = \epsilon_i\) under the minimum entropy martingale
measure, \( Q \), is,

\[
\psi_n(s, e_i, x, e_j) = \begin{cases} 
\frac{1}{c(i)} \frac{A_{ji}}{\sqrt{2\pi}} e^{-\frac{(\log(x/s) - \frac{1}{2} \sigma_j^2 x)}{\sigma_j^2}} \frac{1}{\sigma_j \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{(\log(x/s) - \frac{1}{2} \sigma_j^2 x)}{\sigma_j \sqrt{2\pi}} \right)^2}, & x > 0 \\
0, & x \leq 0,
\end{cases}
\]

where \( e_j \in \mathcal{H} \), \( x \in \mathbb{R} \) and,

\[
c(i) = E \left[ e^{-(\alpha \cdot Z_{n-1}) e^{X_n}} \left| Z_{n-1} = e_i \right. \right]
\]

\[
= \sum_{j=1}^{M} A_{ji} \int_{-\infty}^{\infty} e^{-\alpha e^{(\mu_j - \frac{1}{2} \sigma_j^2) x} e^{\frac{1}{2} \sigma_j^2 x}} \, dx.
\]

Notice that \( \psi_n \) is independent of \( n \).

**Proof.** Let \( V_n \) be an arbitrary \( \mathcal{F}_n \)-measurable random variable. Using Bayes' rule and properties of the entropy Radon-Nikodym derivative given in Equation (6.9) of Theorem 6.3.1,

\[
E^Q[V_n | \mathcal{F}_{n-1}] = \frac{E \left[ e^{-(\alpha \cdot Z_{n-1}) e^{X_n}} V_n \left| \mathcal{F}_{n-1} \right. \right]}{E \left[ e^{-(\alpha \cdot Z_{n-1}) e^{X_n}} \left| \mathcal{F}_{n-1} \right. \right]}.
\]

Note that \( e^{X_n} = \frac{S_n}{S_{n-1}} \). Using the joint conditional density function of \((S_n, Z_n)\) given \( S_{n-1} = s \) and \( Z_{n-1} = e_i \) under the real world probability measure, from Theorem 2.2.1, and the definition of \( c \), we obtain,

\[
E^Q[V_n | S_{n-1} = s, Z_{n-1} = e_i] = \frac{1}{c(i)} \sum_{j=1}^{M} \int_{0}^{\infty} A_{ji} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\log(x/s) - \mu_j)^2}{2\sigma_j^2}} \times e^{-\frac{1}{2} \left( \frac{(\log(x/s) - \frac{1}{2} \sigma_j^2 x)}{\sigma_j \sqrt{2\pi}} \right)^2} V_n(x, e_j) \, dx.
\]

Choosing \( V_n(S_n, Z_n) = I(S_n \leq s', Z_n = e_{j'}) \) gives the result. \( \square \)

With this lemma, we can immediately find the minimum entropy price of a call option one period before expiry;

**Theorem 6.4.2.** Let \( S \) be an asset which evolves according to the Switching Black–Scholes model, with initial asset price, \( S_0 \), and initial market state, \( e_i \).
Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_M)^T$ be the vector of entropy parameters defined in Theorem 6.3.1. When there is no intermediate switching, the minimum entropy price of a call option with strike price, $K$, on the asset, $S$, one period before expiry is,

$$H_0(S_0, c_i) = \frac{e^{-\tau r}}{c(i)} \sum_{j=1}^{M} A_{ji} \left\{ S_0 e^{(\mu_j - \frac{1}{2} \sigma_j^2)\tau} \int_{-\infty}^{\infty} e^{\sigma_j \sqrt{\tau} y - \alpha_i e^{(\mu_j - \frac{1}{2} \sigma_j^2)\tau + \frac{1}{2} \sigma_j^2 \tau + \frac{1}{2} y^2} dy - K \int_{-\infty}^{\infty} e^{-\alpha_i e^{(\mu_j - \frac{1}{2} \sigma_j^2)\tau + \frac{1}{2} \sigma_j^2 \tau + \frac{1}{2} y^2} dy \right\}, \quad (6.13)$$

where,

$$d = \frac{\ln(S_0/K) + (\mu_j - \frac{1}{2} \sigma_j^2)\tau}{\sigma_j \sqrt{\tau}},$$

and,

$$c(i) = \sum_{j=1}^{M} A_{ji} \int_{-\infty}^{\infty} e^{-\alpha_i e^{(\mu_j - \frac{1}{2} \sigma_j^2)\tau + \frac{1}{2} \sigma_j^2 \tau} e^{\frac{1}{2}y^2} dx.$$  

Proof. From Equation (6.12) and Lemma 6.4.1, the call option price is,

$$\frac{e^{-\tau r}}{c(i)} \sum_{j=1}^{M} A_{ji} \int_{-\infty}^{\infty} (x - K) e^{-\frac{\sigma_j^2}{2} \ln(S_0/K) - (\mu_j - \frac{1}{2} \sigma_j^2)\tau + \frac{1}{2} \sigma_j^2 \tau} e^{-\frac{1}{2} (\ln(S_0/K) - (\mu_j - \frac{1}{2} \sigma_j^2)\tau)^2} dx.$$  

Changing variables to $y = \frac{\ln(x/S_0) - (\mu_j - \frac{1}{2} \sigma_j^2)\tau}{\sigma_j \sqrt{\tau}}$ gives the result. \hfill \square

The call price in Theorem 6.4.2 involves three integrals which we are unable to calculate analytically. (Two are seen directly in the call price (6.13), and the third is involved in calculating $c(i)$). However, we could use numerical integration to calculate these integrals and thus compute the minimum entropy call option price.

Unfortunately, this procedure can not easily be generalised to $N$ periods, as the number of calculations required to numerically evaluate the integrals explodes exponentially. An alternative solution is to discretise the asset price, which will now be discussed.

### 6.4.2 Pricing with a Discretised Asset Price

The Switching Black–Scholes model has the property that,

$$S_k = S_{k-1} e^{(\mu(Z_k) - \frac{1}{2} \sigma^2(Z_k)^2)\tau + \sigma(Z_k) \Delta W_k},$$

where $Z_k$ is a standard normal random variable.
where \( \Delta W_k \), the change in Brownian motion from time \( k-1 \) to time \( k \), is normally distributed with mean 0 and variance \( \tau \). Pricing by minimum entropy gives rise to integrals of the form,

\[
\int_{-\infty}^{\infty} e^{-ae^{bx} - \frac{1}{2}x^2} \, dx,
\]

which makes the entropy method intractable for our model. In other models, such as the binomial asset pricing model, this problem does not arise.

In this section, to investigate the entropy pricing method further, we approximate \( \Delta W_k \) by a discrete distribution,

\[
\Delta W_k \approx \begin{cases} 
\sqrt{\tau} & \text{with probability } 1/2, \\
-\sqrt{\tau} & \text{with probability } 1/2.
\end{cases}
\]

We are then able to calculate minimum entropy call option prices over \( N \) switching periods.

**The Change of Measure**

Recall from Theorem 6.3.1 that the minimum entropy martingale measure, \( Q \), is defined by,

\[
\frac{dQ}{dP} = \prod_{k=1}^{N} \frac{e^{-(\alpha_k Z_{k-1})e^{X_k}}}{\mathbb{E}\left[e^{-(\alpha_k Z_{k-1})e^{X_k}} \mid \mathcal{F}_{k-1}\right]},
\]

where \( X_k = (\mu(Z_k) - \frac{1}{2}\sigma(Z_k)^2)\tau + \sigma(Z_k)\Delta W_k \) and \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_M)^T \) is the vector of entropy parameters, with \( \alpha_i \) being the solution to the equation,

\[
\mathbb{E}\left[(e^{X_k} - e^{\tau \sqrt{\tau}})e^{-\alpha_i(e^{X_k} - e^{\tau \sqrt{\tau}})} \mid Z_{k-1} = e_i\right] = 0.
\]

Applying our approximation for \( \Delta W_k \), the expectation becomes,

\[
\sum_{j=1}^{M} A_{ij}^k \left[ \left( e^{(\mu_j - \frac{1}{2}\sigma_j^2)^\tau + \sigma_j \sqrt{\tau} - e^{\tau \sqrt{\tau}}} \right) e^{-\alpha_i \left(e^{(\mu_j - \frac{1}{2}\sigma_j^2)^\tau + \sigma_j \sqrt{\tau} - e^{\tau \sqrt{\tau}}} - e^{\tau \sqrt{\tau}}\right)} + \left( e^{(\mu_j - \frac{1}{2}\sigma_j^2)^\tau - \sigma_j \sqrt{\tau} - e^{\tau \sqrt{\tau}}} \right) e^{-\alpha_i \left(e^{(\mu_j - \frac{1}{2}\sigma_j^2)^\tau - \sigma_j \sqrt{\tau} - e^{\tau \sqrt{\tau}}} - e^{\tau \sqrt{\tau}}\right)} \right] = 0. \tag{6.14}
\]

When this equation is solved for \( i = 1, 2, \ldots, M \) to give the vector \( \alpha \), the change of measure can be computed. As in the continuous case, \( \alpha_i \) can be found using
the method of interval bisection.

**Pricing**

Recall from Equation (6.11) that the minimum entropy price of a contingent claim, $H_N$, at time $0$ is,

$$
H_0 = E \left[ \prod_{n=1}^{N} \frac{e^{-(\alpha, z_{n-1})e^{\gamma n}}}{E^{(\alpha, z_{n-1})e^{\gamma n}|F_{n-1}}} H_N e^{-rT} \bigg| F_0 \right].
$$

Also, setting $H_{n-1} = E^Q[H_n|F_{n-1}]$ for $n = 1, 2, \ldots, N$, the value of the claim $H_N$ at time $t_{n-1}$,

$$
H_{n-1} = \frac{e^{-r\tau}E[H_n e^{-(\alpha, z_{n-1})e^{\gamma n}}|F_{n-1}]}{E[e^{-(\alpha, z_{n-1})e^{\gamma n}}|F_{n-1}]}.
$$

(6.15)

Observing that the denominator of (6.15) depends only on $Z_{n-1}$, we make the following definition.

**Definition 6.4.3.** For $i \in \{1, 2, \ldots, M\}$, define,

$$
c(i) = E \left[ e^{-(\alpha, z_{n-1})e^{\gamma n}} \bigg| Z_{n-1} = e_i \right].
$$

This quantity is independent of $n$, and computing the expectation gives,

$$
c(i) = \sum_{j=1}^{M} \frac{A_{ji}}{2} \left( e^{-\alpha_i e^{\rho_j e^{\gamma} + \sigma_j \sqrt{\tau}}} + e^{-\alpha_i e^{\rho_j e^{\gamma} - \sigma_j \sqrt{\tau}}} \right).
$$

In the following theorem, we calculate the expectations in Equation (6.15) to give the minimum entropy price at any time of a call option expiring at time $T$.

**Theorem 6.4.4.** Let $S$ be an asset which evolves according to the discretised Switching Black–Scholes model, and let $H$ be a call option on this asset expiring at time $T$ with strike price, $K$. The minimum entropy price of $H$ at any time $(N - n)\tau$, given $S_{N-n} = s$ and $Z_{N-n} = e_{j_n+1}$, is,

$$
H_{N-n}(s, e_{j_n+1}) = e^{-rn\tau} \sum_{j_1, j_2, \ldots, j_n=1}^{M} \frac{A_{j_1, j_2, \ldots, j_n} A_{j_n, j_{n+1}} \cdots A_{j_{n+1}, j_{n+2}}}{2^n c(j_2) \cdots c(j_{n+1})} \times \sum_{p=1}^{2^n} \left( s e^{F_n(j_1, \ldots, j_n) + K^n_p(j_1, \ldots, j_n) - K} \right) e^{-F^n_p(j_1, \ldots, j_{n+1})},
$$

(6.16)
where,

\[ E_n(j_1, \ldots, j_n) = \sum_{k=1}^{n} (\mu_{j_k} - \frac{1}{2} \sigma_{j_k}^2) \tau, \]

\[ K_n^p(j_1, \ldots, j_n) = \sum_{k=1}^{n} (1 - 2p_{k-1}) \sigma_{j_k} \sqrt{\tau}, \]

\[ F_n^p(j_1, \ldots, j_{n+1}) = \sum_{k=1}^{n} \alpha_{j_{k+1}} e^{E_1(j_k) + (1-2p_{k-1}) \sigma_{j_k} \sqrt{\tau}}, \]

\( \alpha_i \) is the solution to Equation (6.14), and the numbers \( p_{k-1} \in \{0, 1\} \) come from the binary representation of \( p - 1 \),

\[ p - 1 = \sum_{i=0}^{n-1} p_i 2^i. \]

**Proof.** The proof is by induction of \( n \). For \( n = 1 \), from the recursive definition (6.15), we have,

\[ H_{N-1} = \frac{e^{-rr}}{c(Z_{N-1})} E \left[ H_N e^{-(\alpha, Z_{N-1}) e^{(\mu(z_N) - \frac{1}{2} \sigma^2(z_N)^2) \tau + \sigma(z_N) \Delta W_N}} \mid \mathcal{F}_{N-1} \right]. \]

As \( H \) is a call option, its value at expiry is \( H_N = (S_N - K)^+ \). Setting \( S_{N-1} = s \) and \( Z_{N-1} = e_{j_1} \) in the \( \sigma \)-algebra \( \mathcal{F}_{N-1} \), conditioning first on \( Z_N = e_{j_1} \) and using the representation \( S_N = se^{(\mu_{j_1} - \frac{1}{2} \sigma_{j_1}^2) \tau + \sigma_{j_1} \Delta W_N} \) gives,

\[ H_{N-1}(s, e_{j_2}) = e^{-rr} \sum_{j_1=1}^{M} \frac{A_{j_1j_2}}{c(j_2)} E \left[ \left( \left( s e^{(\mu_{j_1} - \frac{1}{2} \sigma_{j_1}^2) \tau + \sigma_{j_1} \Delta W_N - K} \right)^{+} - \alpha_{j_2} e^{E_1(j_1) + \sigma_{j_1} \sqrt{\tau}} \right) e^{-\alpha_{j_2} e^{E_1(j_1) + \sigma_{j_1} \sqrt{\tau}}} \right]. \]

Applying our discretised approximation of \( \Delta W_N \) we get,

\[ e^{-rr} \sum_{j_1=1}^{M} \frac{A_{j_1j_2}}{2c(j_2)} \left\{ \left( s e^{E_1(j_1) + \sigma_{j_1} \sqrt{\tau} - K} \right)^{+} e^{-\alpha_{j_2} e^{E_1(j_1) + \sigma_{j_1} \sqrt{\tau}}} \right. \]

\[ \left. + \left( s e^{E_1(j_1) - \sigma_{j_1} \sqrt{\tau} - K} \right)^{+} e^{-\alpha_{j_2} e^{E_1(j_1) - \sigma_{j_1} \sqrt{\tau}}} \right\}. \]

As \( K_1^1(j_1) = \sigma_{j_1} \sqrt{\tau} \) and \( K_1^2(j_1) = -\sigma_{j_1} \sqrt{\tau} \), this gives Equation (6.16) for \( n = 1 \).

Next, suppose the result is true for arbitrary \( k \). Using Equation (6.15), the definition of \( H_{n-1} \), as before,

\[ H_{N-k-1}(s, e_{j_{k+2}}) = e^{-rr} \frac{e^{-T}}{c(j_{k+2})} E \left[ e^{-\alpha_{j_{k+2}} e^{(\mu(z_{N-k}) - \frac{1}{2} \sigma^2(z_{N-k})^2) \tau + \sigma(z_{N-k}) \Delta W_{N-k}}} \right. \]

\[ \left. \times H_{N-k}(S_{N-k}, Z_{N-k}) \mid S_{N-k} = s, Z_{N-k} = e_{j_{k+2}} \right]. \]
Condition also on \( Z_{N-k} = e_{j_{k+1}} \) and note that
\[
S_{N-k} = S_{N-k-1} e^{(\mu_{j_{k+1}} + \frac{1}{2} \sigma_{j_{k+1}}^2) T + \sigma_{j_{k+1}} \Delta W_{N-k}}
\]
to obtain,
\[
e^{-rT} \sum_{j_{k+1} = 1}^{M} \frac{A_{j_{k+1}j_{k+2}}}{c(j_{k+2})} \left[ H_{N-k} \left( se^{(\mu_{j_{k+1}} + \frac{1}{2} \sigma_{j_{k+1}}^2) T + \sigma_{j_{k+1}} \Delta W_{N-k}, e_{j_{k+1}})} \times e^{-\alpha_{j_{k+2}} e^{E_{1}(j_{k+1}) + \sigma_{j_{k+1}} \Delta W_{N-k}}} \right].
\]
Calculating the expectation gives,
\[
e^{-rT} \sum_{j_{k+1} = 1}^{M} \frac{A_{j_{k+1}j_{k+2}}}{2c(j_{k+2})} \times \left\{ H_{N-k} \left( se^{(\mu_{j_{k+1}} + \frac{1}{2} \sigma_{j_{k+1}}^2) T + \sigma_{j_{k+1}} \Delta W_{N-k}, e_{j_{k+1}})} e^{-\alpha_{j_{k+2}} e^{E_{1}(j_{k+1}) + \sigma_{j_{k+1}} \Delta W_{N-k}}} \right.
\+ H_{N-k} \left( se^{(\mu_{j_{k+1}} + \frac{1}{2} \sigma_{j_{k+1}}^2) T - \sigma_{j_{k+1}} \Delta W_{N-k}, e_{j_{k+1}}) e^{-\alpha_{j_{k+2}} e^{E_{1}(j_{k+1}) - \sigma_{j_{k+1}} \Delta W_{N-k}}} \right).\right\}
\]
Now, from the induction hypothesis, \( H_{N-k} \) has the form in Equation (6.16), and from the definition of \( K_n^p \) we see that,
\[
K_{n+1}^p(j_1, \ldots, j_{n+1}) = \begin{cases} 
K_n^p(j_1, \ldots, j_n) + \sigma_{j_{n+1}} \sqrt{T}, & p = 1, 2, \ldots, 2^n, \\
K_n^{p-2^n}(j_1, \ldots, j_n) - \sigma_{j_{n+1}} \sqrt{T}, & p = 2^n + 1, \ldots, 2^{n+1}.
\end{cases}
\]
The result then follows. \( \square \)

Immediately from this theorem we can write down the minimum entropy price of a call option at time 0.

**Corollary 6.4.5.** Let \( S \) be an asset which evolves according to the discretised Switching Black–Scholes model, with initial asset price, \( s \), and initial market state, \( e_{j_{N+1}} \). The minimum entropy price of a call option expiring at time \( T \) with strike price, \( K \), on the asset, \( S \), is,
\[
H_0(s, e_{j_{N+1}}) = e^{-rT} \sum_{j_1,j_2,\ldots,j_N=1}^{M} \frac{A_{j_1j_2}A_{j_2j_3} \cdots A_{j_Nj_{N+1}}}{2^N c(j_2) \cdots c(j_{N+1})} \times \sum_{p=1}^{2^N} \left( se^{(\mu_{j_k} + \frac{1}{2} \sigma_{j_k}^2) T + \sum_{k=1}^{N-1}[1-2p_k-1] \sigma_{j_k} \sqrt{T}} - K \right)^+
\times e^{-\sum_{k=1}^{N} \alpha_{j_{k+1}} e^{(\mu_{j_k} + \frac{1}{2} \sigma_{j_k}^2) T + \sum_{k=1}^{N-1}[1-2p_k-1] \sigma_{j_k} \sqrt{T}}},
\]
where for \( i = 1, 2, \ldots, M \), \( \alpha_i \) satisfies the equation,

\[
\sum_{j=1}^{M} \frac{A_{ij}}{2} \left[ \left( e^{(\mu_j - \frac{1}{2}\sigma_j^2)t + \sigma_j \sqrt{\tau} - e^{\tau \gamma}} \right) e^{-\alpha_i \left( e^{(\mu_j - \frac{1}{2}\sigma_j^2)t + \sigma_j \sqrt{\tau} - e^{\tau \gamma}} \right)} + \left( e^{(\mu_j - \frac{1}{2}\sigma_j^2)t - \sigma_j \sqrt{\tau} - e^{\tau \gamma}} \right) e^{-\alpha_i \left( e^{(\mu_j - \frac{1}{2}\sigma_j^2)t - \sigma_j \sqrt{\tau} - e^{\tau \gamma}} \right)} \right] = 0,
\]

the function \( c \) is defined by,

\[
c(i) = \sum_{j=1}^{M} \frac{A_{ij}}{2} \left( e^{-\alpha_i e^{(\mu_j - \frac{1}{2}\sigma_j^2)t + \sigma_j \sqrt{\tau}}} + e^{-\alpha_i e^{(\mu_j - \frac{1}{2}\sigma_j^2)t - \sigma_j \sqrt{\tau}}} \right), \tag{6.17}
\]

and the numbers \( p_{k-1} \in \{0, 1\} \) come from the binary representation of \( p - 1 \),

\[
p - 1 = \sum_{i=0}^{N-1} p_i 2^i.
\]

**Proof.** This follows directly from the previous theorem. \( \square \)

As a topic for further study, the success of our approximation could be measured by checking whether the limiting value of this price as the number of time steps approaches infinity is equal to the Black-Scholes price. This would be expected.

The formula for a minimum entropy call option price given in the above theorem can be written in an alternative way, which clearly shows that it is a sum of call option expiry prices weighted according to some probability. For \( i, j \in \{1, 2, \ldots, M\} \) and \( x \in \{0, 1\} \), define,

\[
q_x(i, j) = \frac{A_{ij} e^{-\alpha_i e^{(\mu_j - \frac{1}{2}\sigma_j^2)t + (1-x)\sigma_j \sqrt{\tau}}}}{2c(j)}. \tag{6.18}
\]

Then from the form of \( c(i) \) in Equation (6.17),

\[
\sum_{i=1}^{M} \{q_0(i, j) + q_1(i, j)\} = 1. \tag{6.19}
\]

Now consider the term multiplying the \((S - K)^{+}\) part in the call price,

\[
\frac{A_{j_1 j_2} \cdots A_{j_N j_{N+1}} e^{-\sum_{k=1}^{N} \alpha_{j_k+1} e^{(\mu_{j_k} - \frac{1}{2}\sigma_{j_k}^2)t + (1-2p_{k-1})\sigma_{j_k} \sqrt{\tau}}} \cdot c(j_2) \cdots c(j_{N+1})}{2^N c(j_2) \cdots c(j_{N+1})}.
\]
From our definition of \( q_x(i, j) \) this is equal to,

\[
Q(j_1, j_2, \ldots, j_{N+1}, p) := q_{p_0}(j_1, j_2)q_{p_1}(j_2, j_3)\ldots q_{p_{N-1}}(j_N, j_{N+1}),
\]

so we can write the call price as,

\[
H_0(s, e^{r_{j_{N+1}}}) = e^{-rT} \sum_{j_1, j_2, \ldots, j_{N}} \sum_{p=1}^{2^N} Q(j_1, j_2, \ldots, j_{N+1}, p) \\
\times \left( s e^{\sum_{k=1}^{N}(\mu_{j_k} - \frac{1}{2}\sigma_{j_k}^2)\tau + \sum_{k=1}^{N}(1-2p_{k-1})\sigma_{j_k}\sqrt{\tau} - K} \right)^+.
\]  \( (6.20) \)

Using induction on \( N \) and the property (6.19) of \( q_x \), it can be shown that,

\[
\sum_{j_1, j_2, \ldots, j_{N}} \sum_{p=1}^{2^N} Q(j_1, j_2, \ldots, j_{N+1}, p) = 1.
\]

Thus Equation (6.20) expresses the minimum entropy call price as a discounted weighted sum of call option values at expiry.

**Special Case – No Switching (ie. \( M=1 \))**

It is interesting to examine the special case when no switching occurs. Recall our model of the asset price,

\[
S_k = S_{k-1}e^{(\mu(Z_k) - \frac{1}{2}\sigma(Z_k)^2)\tau + \sigma(Z_k)\Delta W_k},
\]

where \( \Delta W_k \) is approximated by a discrete distribution,

\[
\Delta W_k \approx \begin{cases} \sqrt{\tau} & \text{with probability } 1/2, \\ -\sqrt{\tau} & \text{with probability } 1/2. \end{cases}
\]

If there is no switching, the Markov chain is constant and the asset price at time \( k\tau \) is equal to \( S_{k-1}e^{(\mu - \frac{1}{2}\sigma^2)\tau + \sigma \Delta W_k} \), where \( \Delta W_k = \pm \sqrt{\tau} \). This is a binomial model which can be represented by a recombining tree. Compared to the general case, there are \( n+1 \) possible values of the asset price at time step \( n \), rather than the previous \( (2M)^n \) possibilities. The following theorem gives a simplified representation of a call price resulting from this.
Theorem 6.4.6. The minimum entropy price of a call option with strike price, $K$, on an asset, $S$, following the discretised Switching Black–Scholes model with $M = 1$ is,

\[
H_0(s) = \frac{e^{-rT}}{c^N} \sum_{k=0}^{N} \binom{N}{k} \left( s e^{(\mu - \frac{1}{2}\sigma^2)r + \sigma\sqrt{T}/(N - 2k) - K} \right) + \\
\times e^{-\alpha e^{\mu - \frac{1}{2}\sigma^2}r} (ke^{-\sigma\sqrt{T}/(N - k)}e^{r/\alpha})^N, (6.21)
\]

where,

\[
c = \left( e^{-\alpha e^{\mu - \frac{1}{2}\sigma^2}r} + e^{-\alpha e^{\mu - \frac{1}{2}\sigma^2}r} \right),
\]

and $\alpha$ satisfies the equation,

\[
\left( e^{(\mu - \frac{1}{2}\sigma^2)r + \sigma\sqrt{T} - e^{r/\alpha}} \right) e^{-\alpha \left( e^{(\mu - \frac{1}{2}\sigma^2)r + \sigma\sqrt{T} - e^{r/\alpha}} \right)} + \left( e^{(\mu - \frac{1}{2}\sigma^2)r - \sigma\sqrt{T} - e^{r/\alpha}} \right) e^{-\alpha \left( e^{(\mu - \frac{1}{2}\sigma^2)r - \sigma\sqrt{T} - e^{r/\alpha}} \right)} = 0.
\]

Proof. From Equation (6.20), the price of a call for $M = 1$ is,

\[
H_0(s) = \frac{e^{-rT}}{c^N} \sum_{p=1}^{2^N} \left( s e^{\sum_{k=1}^{N} (\mu - \frac{1}{2}\sigma^2)r + \sum_{k=1}^{N} (1 - 2p_{k-1})\sigma\sqrt{T} - K} \right) + Q(p),
\]

where, for $p - 1 = \sum_{i=0}^{N-1} p_i 2^i$,

\[
Q(p) = q_{p_0}q_{p_1}\cdots q_{p_{N-1}}, \quad (6.22)
\]

and for $x \in \{0, 1\}$,

\[
q_x = \frac{e^{-\alpha e^{\mu - \frac{1}{2}\sigma^2}r + (-1)^x \sigma\sqrt{T}}}{c}.
\]

Taking the product in (6.22), we see that,

\[
Q(p) = \frac{e^{-\alpha e^{\mu - \frac{1}{2}\sigma^2}r + \sum_{k=1}^{N} (-1)^{p_k - 1}\sigma\sqrt{T}}}{c^N}.
\]

Hence the price is,

\[
H_0(s) = \frac{e^{-rT}}{c^N} \sum_{p=1}^{2^N} \left( s e^{(\mu - \frac{1}{2}\sigma^2)r + \sigma\sqrt{T} \sum_{i=1}^{N} (1 - 2p_{i-1}) - K} \right) + \\
\times e^{-\alpha e^{\mu - \frac{1}{2}\sigma^2}r \sum_{i=1}^{N} (-1)^{p_i - 1}\sigma\sqrt{T}}.
\]
The only dependence on $p$ is in the terms $\sum_{i=1}^N p_{i-1}$ and $\sum_{i=1}^N e^{(-1)^{p_{i-1}} \sigma \sqrt{\tau}}$. As $p_{i-1} \in \{0, 1\}$ and every possibility $\{p_0, p_1, \ldots, p_{N-1}\}$ is achieved, the sum $\sum_{i=1}^N p_{i-1}$ can take the value $k$ for $k = 0, 1, \ldots, N$ in $\binom{N}{k}$ ways, and in this case,

$$\sum_{i=1}^N e^{(-1)^{p_{i-1}} \sigma \sqrt{\tau}} = k e^{-\sigma \sqrt{\tau}} + (N - k) e^{\sigma \sqrt{\tau}}.$$ 

Thus the call price is given by Equation (6.21). \qed
Chapter 7

Numerical Studies of the Option Pricing Methodologies

The preceding chapters described and examined the Switching Black–Scholes model for an asset price, and applied the methods of mean-variance hedging, Esscher transforms and minimum entropy to compute the prices of options on such assets. To evaluate these option prices, in total ten parameters values must be specified. These are,

1. The strike price, $K$, of the option;
2. The time to expiry, $T$, of the option;
3. The initial asset price, $S_0$;
4. The number of possible market states, $M$;
5. The initial market state, $Z_0$;
6. The drift vector, $\mu = (\mu_1, \mu_2, \ldots, \mu_M)^T$, of the asset;
7. The volatility vector, $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_M)^T$, of the asset;
8. The number of switches, $N$;
9. The transition matrix, $A$, for the switches;
10. The continuously compounding interest rate, $r$. 

127
In this chapter, we choose particular values for these parameters so as to give some concrete results, compare the three different option pricing methodologies and analyse the Switching Black–Scholes model in more detail.

The call option prices tabulated in the following sections are computed using the C programs given in Appendix 4. These programs calculate the mean-variance, Esscher and minimum entropy call option prices using the algorithms and formulae given in Chapters 3, 5 and 6. For the minimum entropy method, the pricing formula obtained by approximating the Brownian motion component of the asset price by a discrete random variable with only two possible values is implemented.

In Section 7.1, these programs are used to compute call option prices for a particular set of parameter values. The results enable us to compare the option pricing methodologies and examine the effect each parameter has on the option price. In each scenario, we also compute a Black-Scholes price of the call option. Analysing these results, we see that significant pricing errors stem from using the traditional Black-Scholes formula to price options on an asset which exhibits switching.

The results obtained from the traditional Black-Scholes formula also indicate that Switching Black–Scholes prices should generate implied volatility smiles. We show in Section 7.2 that this is indeed the case, and investigate the effect of the various parameters on the shape of the implied volatility curve.

Throughout the chapter, topics for future research suggested by the analysis are also discussed.

### 7.1 Comparison of Call Option Prices

In this section, we compare the three methods of option pricing that have been discussed in previous chapters by choosing particular values for the ten parameters needed to specify the model and call option, and using these as inputs in the C programs to compute each call option price. By varying these pa-
rameters, we also observe their effect on the prices. The particular parameters we consider are specified in Section 7.1.1, and the resulting call option prices tabulated. These results are then interpreted in Section 7.1.2.

### 7.1.1 Parameter Specification and Call Option Prices

To compare the various option pricing methodologies and examine the effect of individual parameters on each call price, we consider a call option which expires in 3 months and has a strike price of 100. The underlying asset is assumed to switch between three possible states, with drifts and volatilities of 6\% and 20\%, 8\% and 40\%, and 10\% and 60\%, respectively, and the continuously compounding interest rate is taken to be 5\% per annum. We compute the mean-variance, Esscher and minimum entropy price for this call option in two scenarios. In the first, the transition matrix for the state changes is,

\[
A = \begin{pmatrix}
0.2 & 0.1 & 0 \\
0.8 & 0.8 & 0.8 \\
0 & 0.1 & 0.2
\end{pmatrix},
\]

which represents an asset having a strong mean-reverting tendency to the second state, and in the second, the transition matrix is taken to be,

\[
A = \begin{pmatrix}
0.4 & 0.1 & 0 \\
0.6 & 0.8 & 0.6 \\
0 & 0.1 & 0.4
\end{pmatrix},
\]

representing an asset having a weaker tendency to revert to the middle state. For each of these matrices, the three option prices are calculated when there are 2, 5, and 10 switching times, and when the initial market state is 1, 2, and 3. The Black-Scholes price is also computed in each scenario, using the initial volatility value.

Tables 7.1, 7.2 and 7.3 contain the various prices of this call option when the initial asset price is 90, 100, and 110, respectively. In each cell, the first entry corresponds to the mean-variance price, the second to the Esscher price,
and the third to the minimum entropy price. In the next section, we interpret these results.

### 7.1.2 Interpretation of the Results

The main aim of this section is to compare the three different option pricing methodologies discussed in previous chapters. However, from the results tabulated in Tables 7.1, 7.2 and 7.3, we can also deduce the effect of various parameters on the option price. After comparing the option pricing methodologies, we examine a number of these effects.

#### Comparison of the Option Pricing Methodologies

In Tables 7.1 - 7.3, the mean-variance, Esscher and minimum entropy prices for the call option discussed in Section 7.1.1 are listed, for varying initial asset prices, number of switches, transition matrices and initial market states. We see

<table>
<thead>
<tr>
<th>Initial State</th>
<th>Weakly Mean-Reverting</th>
<th>Strongly Mean-Reverting</th>
<th>Black-Scholes Price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N = 2$</td>
<td>$N = 5$</td>
<td>$N = 10$</td>
</tr>
<tr>
<td>1</td>
<td>3.13</td>
<td>3.66</td>
<td>3.90</td>
</tr>
<tr>
<td></td>
<td>3.15</td>
<td>3.67</td>
<td>3.90</td>
</tr>
<tr>
<td></td>
<td>3.58</td>
<td>3.69</td>
<td>4.04</td>
</tr>
<tr>
<td>2</td>
<td>4.00</td>
<td>4.06</td>
<td>4.10</td>
</tr>
<tr>
<td></td>
<td>4.02</td>
<td>4.07</td>
<td>4.10</td>
</tr>
<tr>
<td></td>
<td>4.55</td>
<td>4.06</td>
<td>4.15</td>
</tr>
<tr>
<td>3</td>
<td>4.94</td>
<td>4.52</td>
<td>4.33</td>
</tr>
<tr>
<td></td>
<td>4.97</td>
<td>4.53</td>
<td>4.34</td>
</tr>
<tr>
<td></td>
<td>5.53</td>
<td>4.53</td>
<td>4.39</td>
</tr>
</tbody>
</table>

\textit{Table 7.1:} Mean-variance, Esscher, minimum entropy and Black-Scholes prices for the call option discussed in Section 7.1.1, when the initial asset price is 90.
<table>
<thead>
<tr>
<th>Initial State</th>
<th>Weakly Mean-Reverting</th>
<th></th>
<th></th>
<th>Strongly Mean-Reverting</th>
<th></th>
<th></th>
<th>Black-Scholes Price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( N = 2 )</td>
<td>( N = 5 )</td>
<td>( N = 10 )</td>
<td>( N = 2 )</td>
<td>( N = 5 )</td>
<td>( N = 10 )</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7.54</td>
<td>8.21</td>
<td>8.50</td>
<td>8.15</td>
<td>8.50</td>
<td>8.62</td>
<td>4.61</td>
</tr>
<tr>
<td></td>
<td>7.56</td>
<td>8.22</td>
<td>8.51</td>
<td>8.17</td>
<td>8.50</td>
<td>8.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.34</td>
<td>8.46</td>
<td>8.56</td>
<td>7.80</td>
<td>8.77</td>
<td>8.67</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>8.60</td>
<td>8.69</td>
<td>8.74</td>
<td>8.61</td>
<td>8.69</td>
<td>8.72</td>
<td>8.55</td>
</tr>
<tr>
<td></td>
<td>8.63</td>
<td>8.70</td>
<td>8.74</td>
<td>8.64</td>
<td>8.70</td>
<td>8.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.16</td>
<td>8.98</td>
<td>8.79</td>
<td>8.22</td>
<td>8.97</td>
<td>8.77</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>9.70</td>
<td>9.22</td>
<td>9.01</td>
<td>9.07</td>
<td>8.88</td>
<td>8.81</td>
<td>12.48</td>
</tr>
<tr>
<td></td>
<td>9.73</td>
<td>9.23</td>
<td>9.02</td>
<td>9.10</td>
<td>8.87</td>
<td>8.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.32</td>
<td>9.52</td>
<td>9.07</td>
<td>8.64</td>
<td>9.17</td>
<td>8.86</td>
<td></td>
</tr>
</tbody>
</table>

**Table 7.2:** Mean-variance, Esscher, minimum entropy and Black-Scholes prices for the call option discussed in Section 7.1.1, when the initial asset price is 100.

<table>
<thead>
<tr>
<th>Initial State</th>
<th>Weakly Mean-Reverting</th>
<th></th>
<th></th>
<th>Strongly Mean-Reverting</th>
<th></th>
<th></th>
<th>Black-Scholes Price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( N = 2 )</td>
<td>( N = 5 )</td>
<td>( N = 10 )</td>
<td>( N = 2 )</td>
<td>( N = 5 )</td>
<td>( N = 10 )</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14.29</td>
<td>14.83</td>
<td>15.08</td>
<td>14.79</td>
<td>15.07</td>
<td>15.18</td>
<td>11.99</td>
</tr>
<tr>
<td></td>
<td>14.30</td>
<td>14.84</td>
<td>15.08</td>
<td>14.80</td>
<td>15.08</td>
<td>15.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.71</td>
<td>14.86</td>
<td>15.14</td>
<td>15.27</td>
<td>15.07</td>
<td>15.24</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>15.20</td>
<td>15.26</td>
<td>15.29</td>
<td>15.20</td>
<td>15.25</td>
<td>15.27</td>
<td>15.11</td>
</tr>
<tr>
<td></td>
<td>15.22</td>
<td>15.27</td>
<td>15.30</td>
<td>15.22</td>
<td>15.25</td>
<td>15.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15.67</td>
<td>15.25</td>
<td>15.36</td>
<td>15.68</td>
<td>15.25</td>
<td>15.33</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>16.21</td>
<td>15.75</td>
<td>15.54</td>
<td>15.61</td>
<td>15.42</td>
<td>15.35</td>
<td>18.85</td>
</tr>
<tr>
<td></td>
<td>16.23</td>
<td>15.76</td>
<td>15.55</td>
<td>15.63</td>
<td>15.43</td>
<td>15.36</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16.61</td>
<td>15.77</td>
<td>15.62</td>
<td>16.09</td>
<td>15.42</td>
<td>15.42</td>
<td></td>
</tr>
</tbody>
</table>

**Table 7.3:** Mean-variance, Esscher, minimum entropy and Black-Scholes prices for the call option discussed in Section 7.1.1, when the initial asset price is 110.
from these tables that, for all the parameter values considered, the Esscher price and the mean-variance price are almost identical, suggesting that the mean-variance and Esscher methods of pricing may actually be equivalent for the Switching Black–Scholes model, or at least equivalent under certain restrictions of the parameters. Thus a topic for further research would be to investigate the validity of this hypothesis (even in an approximate sense), either by numerical analysis and statistical tests, or by a direct comparison of the formulae. If found to be correct, we could then use the more efficient numerical procedure for the Esscher price to find an approximation to the mean-variance price.

Further analysis of Tables 7.1 - 7.3 reveals that the minimum entropy price is generally higher than the mean-variance and Esscher prices, although perhaps such a comparison is not valid, as the minimum entropy prices calculated in these tables are found by approximating the Brownian motion component of the asset price by a discrete random variable with only two possible values. A topic for further investigation would be to examine the effect this approximation has on the price. One would expect that as the number of switches in a given time interval increases, the accuracy of the approximation would be improved. If so, our numerical example then indicates that the minimum entropy price, when no approximation is made, may also be insignificantly different to the mean-variance and Esscher prices. This potential equivalence should be examined.

Assuming that the three pricing methodologies are not equivalent, an obvious question is, “which method of pricing should be used?” The answer to this question depends on the point of view being taken. For example, a professional writing call options would most likely be interested in using the mean-variance price, as this gives a hedging strategy which minimises the expected loss he would make at expiry. An investor, on the other hand, may be interested in using either the Esscher or minimum entropy price, as these correspond to maximising the utility of an investor’s wealth at expiry. The particular method then depends on the risk preferences of the individual.
Effect of the Initial Volatility on the Option Price

The tables clearly indicate that as the initial volatility of the asset price increases, the call option becomes more expensive. However, this effect may become less significant as the number of switching times increases (as then the effect of the initial volatility would be reduced).

Effect of the Transition Matrix on the Option Price

Tables 7.1 - 7.3 display the call option price for two different transition matrices. The first represents a weak tendency to revert to the state with drift and volatility of 8% and 40%, while the second represents a very strong tendency for this reversion. We see in the tables that the effect of changing from one matrix to the other depends on the initial state of the asset parameters. If the initial state has low drift and volatility (that is, state 1), then the call option price increases when the transition matrix changes from a weak to a strong mean reverting nature. The opposite is true when the initial asset parameters are larger than those of the mean reverting state (that is, the initial state is 3). A negligible change in the price is observed when the mean reverting state (that is, state 2) is the initial state. Again, these results agree with intuition.

The effect of the transition matrix seems to diminish as a larger number of switches in a given time interval are permitted. This is a consequence of the greater likelihood of spending time in a variety of states.

To further examine the effect of the transition matrix on a call option price, we consider a second example. Table 7.4 displays the mean-variance prices of a call option on an asset with two possible states, \( \mu_1 = 0.08 \), \( \sigma_1 = 0.20 \) and \( \mu_2 = 0.10 \), \( \sigma_2 = 0.40 \), for 36 different transition matrices, representing all possible situations. In this example, the call option was assumed to expire in 3 months and to have a strike price of 100. The initial asset price was also taken to be 100, and the initial state was 1. Six possible switching times were assumed, and the continuously compounding interest rate was 5% per annum.
### Table 7.4: Mean-variance prices of a call option on an asset following the Switching Black–Scholes model with 2 possible states, for various transition matrices. Here, $A_{11}$ is the probability of remaining in state 1, and $A_{22}$ is the probability of remaining in state 2.

We see in the table that if the transition probability $A_{11}$ is 1, so that the asset does not switch out of the initial state, the call price is always 4.61. This equals the Black-Scholes price for such a call option, as would be expected.

For the case of $A_{22} = 1$, the asset stays in state 1 for a random amount of time before being absorbed into state 2. A smaller value of $A_{11}$ means that the asset is likely to switch into state 2 more quickly. As state 2 has a higher volatility, the call price increases as $A_{11}$ decreases. For similar reasons, the call price increases as $A_{22}$ increases. These principles continue to apply for other parameter specifications.

### Effect of the Number of Switches on the Option Price

From the prices tabulated in Tables 7.1 - 7.3, it appears as though the call option price converges to a fixed value as the number of possible switches increases. This is especially evident for the mean-variance and Escher pricing methodologies, but is not so obvious for the minimum entropy price. As before, this may be due to the approximation to the Brownian motion component
which is assumed in the minimum entropy procedure.

From the definition of the Switching Black–Scholes model, we would expect this convergence to be towards a traditional Black–Scholes price for some fixed volatility value. To confirm this hypothesis and determine the volatility value are topics for further research.

Additionally, notice that for an initial state of 1, the prices inflate as the number of possible switches increases, whereas they deflate for an initial state of 3. This reflects the lessening effect of the initial state on the price as more switches are permitted.

**Effect of “Moneyness” of the Option on the Option Price**

For all the methods of pricing and all the parameter values considered, the tables show that the call option becomes more expensive as the option moves from being out of the money (that is, \( S_0 < K \)) to being in the money (that is, \( S_0 > K \)). This is a property of the price which we would desire and expect of a pricing methodology.

**Comparison of Switching Prices to Black-Scholes Prices**

In addition to giving the mean-variance, Esscher and minimum entropy call option prices, Tables 7.1 - 7.3 also display the Black-Scholes price of the option. This is the price which would be assigned to the option if the drift and volatility of the underlying asset did not vary from their initial values. The results indicate that significant pricing errors arise from using the traditional Black-Scholes formula when the underlying asset follows a Switching Black-Scholes type model. We see that when the underlying asset has initial state 1, the Black-Scholes formula notably under-prices the call option, whereas the reverse is true when the initial state is 3. This can be understood intuitively. For example, if an asset has a high likelihood of moving from low to high levels of drift and volatility, we would expect the price of a call option on this asset to be greater than if the low drift and volatility persisted. This is the case when
the initial state is 1. Moreover, the lower this probability of switching to higher levels of drift and volatility, the smaller the pricing error. This is also observed in the tables.

From these results, we might expect Black-Scholes implied volatility smiles to be generated from the Switching Black–Scholes call option prices. This topic is discussed in the next section.

7.2 Black-Scholes Implied Volatility Smiles

Recall that one of our reasons for introducing the Switching Black–Scholes model was to try and produce Black-Scholes implied volatility smiles, as observed in the market. We now use the option pricing formulae of the previous chapters to show that the Switching Black–Scholes model is indeed consistent with this observation, and examine the type of implied volatility smiles generated. We focus here on the mean-variance method of pricing, but briefly comment on results from the other methodologies at the conclusion of this section.

To examine the type of smiles which can be generated from the Switching Black–Scholes model, we consider a call option expiring in three months, on an asset which can switch between three possible states at six equally spaced times. As in Section 7.1, we take the possible drifts and volatilities in each state to be 6% and 20%, 8% and 40%, and 10% and 60%, respectively, and assume a continuously compounding interest rate of 5% per annum. The initial asset price is taken to be 100. We consider three different transition matrices,

\[
\text{matrix 1} = \begin{pmatrix} 0.2 & 0.1 & 0 \\ 0.8 & 0.8 & 0.8 \\ 0 & 0.1 & 0.2 \end{pmatrix}, \quad \text{matrix 2} = \begin{pmatrix} 0.4 & 0.1 & 0 \\ 0.6 & 0.8 & 0.6 \\ 0 & 0.1 & 0.4 \end{pmatrix},
\]

\[
\text{matrix 3} = \begin{pmatrix} 0.5 & 0.33 & 0 \\ 0.5 & 0.34 & 0.5 \\ 0 & 0.33 & 0.5 \end{pmatrix}.
\]
The first two matrices are the same as those used in Section 7.1, and represent an asset with strong (respectively, weak) mean-reversion to the middle state. The third matrix represents an asset with equal probabilities of remaining in the same state and jumping to an adjacent state.

Figures 7.1, 7.2 and 7.3 (pages 139 - 141) show the implied volatility smiles produced in each of these scenarios, when the initial state is 1, 2 and 3, respectively. As is intuitively expected, the figures show that the implied volatility smile becomes deeper (that is, the range of implied volatilities increases) as it becomes more likely for the asset parameters to switch states frequently. This is true regardless of the initial asset state.

Figure 7.4 (page 142) shows, using transition matrix 3, the effect of a varying initial state on the implied volatility curve. We see from this figure that the range of implied volatilities decreases as the initial state increases. This can be explained when the actual volatility values are considered. Recall that we are assuming the asset has volatility states of 20%, 40% and 60%. The graphs show that almost all of the implied volatilities are greater than 40%. These two facts explain the observations in Figure 7.4. The results are similar when matrices 1 and 2 are used as the transition matrix.

The effect of the time to expiry of the call option on the implied volatility curve is shown in Figure 7.5 (page 143), where transition matrix 3, initial state 1 and times to expiry of 1 week, 1 month and 3 months are used as inputs. We see from this graph that options with shorter times to expiry produce deeper implied volatility smiles.

In all of the examples considered above, the implied volatilities generated always lie between the smallest and largest possible volatilities of the asset. (In particular, between 20% and 60% in the above examples). This is consistent with the results discussed by El Karoui, Jeanblanc-Picqué and Shreve in the paper [25]. As a consequence, if we are trying to fit a Switching Black–Scholes model to a given implied volatility curve, the possible volatility states must at least cover the range of implied volatilities needed. Also, as discussed above,
the deepness of the smile can be adjusted by choosing an appropriate transition matrix, initial market state and set of possible drifts and volatilities.

If, after varying all these parameters, a given smile still cannot be replicated, we could use a different method of pricing to alter the shape of the smile generated. Figures 7.6 and 7.7 (pages 144, 145) show the type of implied volatility curves obtained using, respectively, the Esscher and minimum entropy methods of pricing. For these examples, the initial state was taken to be 1, and the other parameters were as discussed at the beginning of this section. We see that the shape of the Esscher implied volatility smile is very similar to that produced by the mean-variance method. However, the minimum entropy smile has a totally different shape. Again, we must remember that in the minimum entropy method, an approximation to the Brownian motion component of the asset price was made, which is a likely reason for the oddly shaped curve obtained in this case. Clearly, better approximations should be sought for this method.
Figure 7.1: Implied volatilities obtained using the mean-variance pricing method, for varying strike prices and transition matrices. Here, the initial market state was taken to be 1, and the other parameters were as discussed in Section 7.2.
Figure 7.2: Implied volatilities obtained using the mean-variance pricing method, for varying strike prices and transition matrices. Here, the initial market state was taken to be 2, and the other parameters were as discussed in Section 7.2.
Figure 7.3: Implied volatilities obtained using the mean-variance pricing method, for varying strike prices and transition matrices. Here, the initial market state was taken to be 3, and the other parameters were as discussed in Section 7.2.
Figure 7.4: Implied volatilities obtained using the mean-variance pricing method, for varying strike prices and initial states. Here, transition matrix 3 was used, and the other parameters were as discussed in Section 7.2.
Figure 7.5: Implied volatilities obtained using the mean-variance pricing method, for varying strike prices and times to expiry. Here, transition matrix 3 and initial state 1 were used, and the other parameters were as discussed in Section 7.2.
Figure 7.6: Implied volatilities obtained using the Esscher pricing method, for varying strike prices and transition matrices. Here, the initial market state was taken to be 1, and the other parameters were as discussed in Section 7.2.
Figure 7.7: Implied volatilities obtained using the minimum entropy pricing method, for varying strike prices and transition matrices. Here, the initial market state was taken to be 1, and the other parameters were as discussed in Section 7.2.
Chapter 8

Estimation for the Switching Black–Scholes Model

In order to apply the pricing formulae for call options on assets which evolve according to the Switching Black–Scholes model, developed in previous chapters, a number of model parameters must be estimated. For example, we need to know the drift vector, $\mu$, and volatility vector, $\sigma$, the transition times and transition matrix of the Markov chain, the initial market state, and the size of the Markov chain state space. We consider in this chapter how some of these parameters can be estimated, given observations of the asset price. In particular, we estimate the volatility vector, the transition matrix, and the state of the Markov chain.

Section 8.1 gives the preliminaries needed for this work. In Section 8.1.1, the Switching Black–Scholes asset price model is altered slightly so that the drift is fixed, but the volatility can still switch according to the hidden Markov chain. Evaluation of the estimates is considerably simplified by transforming these dynamics to give what we call the observation model. The different time scales between the observation process (continuous) and the hidden Markov chain (discrete) are reconciled in Section 8.1.2, by constructing a new continuous-
time state process which has CADLAG\(^1\) sample paths, indistinguishable from
the Markov chain at the transition times. The required semimartingale repre-
sentation of this process is then obtained, using two Dirac measures. The
preliminaries conclude in Section 8.1.3, where a new measure is defined using
Girsanov’s theorem. Throughout the chapter, Bayes rule is used to convert to
this new measure, under which the observation process is Brownian motion,
simplifying the mathematics.

Sections 8.2 and 8.3 focus respectively on estimating the market state and
model parameters. The estimate of the market state, or state of the hidden
Markov chain, at time \(t\) is obtained by directly evaluating the appropriate con-
ditional expectation. Conditioning on the information available until time \(t\)
gives a filtered estimate. However, this estimate involves stochastic integration,
which can lead to implementation difficulties. These are avoided by using the
robust filtered estimate, where the stochastic process appears as a parameter
in a deterministic integral equation. A smoothed estimate of the market state
is also calculated by conditioning on the information available to a time \(T > t\).

Having estimated the market state, maximum likelihood estimates of the
volatility vector and transition matrix are found via the EM algorithm. This
involves calculation of three conditional expectations, as discussed in Campillo
and Le Gland ([14]) and Elliott ([27]). As for the market state, we calculate fil-
tered estimates, robust filtered estimates, and smoothed estimates, using Bayes
rule to change to the new measure. Algorithms for implementation are also
given.

Throughout this chapter, time dependent processes are written with time
as a subscript: \(H = (H_t)_{0 \leq t \leq T}\). The symbol \(H_{t-}\) is used to denote the left limit
of \(H\) at time \(t\),

\[
H_{t-} = \lim_{u \uparrow t} H_u,
\]

\(^1\)A CADLAG process is right continuous and has finite limits on the left.
and the following norms are adopted (when there is no notational confusion):
\[ \| x \| = \left( \sum_{i=1}^{M} |x_i|^2 \right)^{1/2}, \quad \text{for } x \in \mathbb{R}^M; \]
\[ \| y \| = \sup \{ \| y_t \|, \ 0 \leq t \leq T \}, \quad \text{for functions } y : [0, T] \to \mathbb{R}^M; \]
\[ \| A \| = \left( \sum_{i,j=1}^{M} |A_{ij}|^2 \right)^{1/2}, \quad \text{for } M \times M \text{ square matrices } A. \]

For a vector \((x_1, x_2, \ldots, x_M)^T \in \mathbb{R}^M\), we use the notation,
\[ \text{diag}\{x_i\} = \text{diag}\{x_1, x_2, \ldots, x_M\} = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_M \end{pmatrix}. \]

The material presented in this chapter contains work produced in collaboration with Professor R.J. Elliott (University of Calgary, Canada) and Dr W.P. Malcolm (DSTO Edinburgh and University of Adelaide).

### 8.1 Preliminaries

To begin, we define the necessary models and change of measure. In Section 8.1.1, the asset price and observation models are discussed. The asset price follows a slightly modified version of the Switching Black–Scholes model, which is transformed to obtain the observation model. The hidden Markov chain, or state process, is converted into a continuous-time process in Section 8.1.2, and a semimartingale representation is found. Finally, in Section 8.1.3, a reference probability, under which we carry out all calculations, is defined via Girsanov’s theorem.

#### 8.1.1 Asset Price and Observation Models

Rather than taking the general Switching Black–Scholes model for the asset price, for this discussion we assume that the asset has constant drift, \( \mu \), which is known. The asset price dynamics are then summarised by the equation,
\[ S_t = S_0 + \int_0^t \mu S_u \, du + \int_0^t \langle \sigma, Z_u \rangle S_u \, d\widehat{W}_u, \quad (8.1) \]
where $\tilde{W}$ is Brownian motion, $Z$ is the hidden Markov chain, which will be discussed more in 8.1.2, and $\sigma = (\sigma_1, \ldots, \sigma_M)^T$ is the vector of asset volatilities. Recall that $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^M$ and the state space of $Z$ is the set of unit vectors in $\mathbb{R}^M$, so that $\langle \sigma, Z_u \rangle$ chooses one of the $\sigma_i$ values, according to the state of $Z_u$. We assume that the asset price can be directly observed.

Next, we want to transform these dynamics to obtain a new process having Brownian motion as the martingale term. This will be the observation model. With such a process, we can apply Girsanov’s theorem to find a reference probability under which the observation dynamics are Brownian motion, and then use Bayes rule to compute the parameter estimates using this new measure. This simplifies the calculations significantly.

To find such a process, we first set $x_t = \log S_t$. Using Itô’s Lemma, we see that,

$$x_t = x_0 + \int_0^t (\mu - \frac{1}{2} \langle \sigma, Z_u \rangle^2) du + \int_0^t \langle \sigma, Z_u \rangle d\tilde{W}_u.$$ 

Furthermore, $x^2$ satisfies the equation,

$$x_t^2 = x_0^2 + 2 \int_0^t x_u dx_u + \int_0^t \langle \sigma^2, Z_u \rangle du,$n

where $\sigma^2 = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2)^T$. Recall $Z_t \in \{e_1, e_2, \ldots, e_M\}$, so that $\langle \sigma, Z_t \rangle^2 = \langle \sigma^2, Z_t \rangle$. We can see from the above equation that,

$$\int_0^t \langle \sigma^2, Z_u \rangle du = x_t^2 - x_0^2 - 2 \int_0^t x_u dx_u,$$

so the integral on the left hand side can be approximated on a partition

$\{t_k : k = 0, 1, \ldots, K\}$ of $[0, T]$ by the observable quantity,

$$\tilde{Y}_{t_k} = x_{t_k}^2 - x_0^2 - 2 \sum_{t=0}^{k-1} x_t (x_{t+1} - x_t).$$

(This is obtained by approximating the stochastic integral, $\int_0^t x_u dx_u$). Assuming that the error of this approximation has the form $\epsilon W_{t_k}$, where $W$ is a Brownian motion independent of $\tilde{W}$ and $\epsilon > 0$ is a small parameter whose value is assumed known, we have,

$$\tilde{Y}_t = \int_0^t \langle \sigma^2, Z_u \rangle du + \epsilon W_t.$$
Finally, we rescale this equation to give the observation dynamics,

\[
Y_t = \int_0^t \langle \xi, Z_u \rangle \, du + W_t, \tag{8.5}
\]

where \( \xi = \frac{1}{\epsilon} \sigma^2 \) and \( \bar{Y}_t = \frac{1}{\epsilon} \bar{Y}_t \).

### 8.1.2 State Process Dynamics

Recall from Chapter 2 that the state process, \( Z \), is a discrete-time, stationary Markov chain defined on a partition \( \{ t_k : k = 0, 1, \ldots, N \} \) of \( [0, T] \) and having values in the state space \( \mathcal{H} = \{ e_1, e_2, \ldots, e_M \} \) of unit vectors in \( \mathbb{R}^M \). The transition probabilities are \( A_{ij} = P(Z_{t_k} = e_i | Z_{t_{k-1}} = e_j) \) and we can write,

\[
Z_{t_k} = AZ_{t_{k-1}} + L_{t_k}, \tag{8.6}
\]

where \( L = (L_{t_k})_{k = 1, 2, \ldots, N} \) is a martingale increment with respect to the filtration generated by \( Z \). Let us specify the initial distribution of \( Z \) by the probability, \( \Pi_0 \), where \( \Pi_0(e_j) = P(Z_0 = e_j) \).

As the Markov chain is a discrete-time process, whereas the asset price model is continuous-time, we reconcile the two different time scales by constructing a continuous-time semimartingale representation for the state process. Using a “sample and hold” procedure,

\[
Z_t = Z_{t_k} \text{ for all } t \in [t_k, t_{k+1}), \tag{8.7}
\]

the discrete-time Markov chain is converted into a continuous-time process with CADLAG sample paths, which coincides with the Markov chain at the transition times. In the following, we take \( Z \) to be this continuous-time process.

To compute a semimartingale representation for \( Z \), we begin by writing \( Z \) as a sum of its initial value, \( Z_0 \), and a second term summing the contributions at the transition times:

\[
Z_t = Z_0 + \sum_{0 < u \leq t} (Z_u - Z_{u-}).
\]
We now introduce two Dirac measures. Let \( m(\cdot) \) be a deterministic measure, defined by,

\[
m(dt) = \sum_{k=0}^{N} \delta_k(dt),
\]

where \( \delta_k(\cdot) \) is a Dirac measure with the property that \( \int_{-\infty}^{\infty} f(t) \delta_k(dt) = f(t_k) \), for \( k = 0, 1, \ldots, N \). This means that \( \int_{-\infty}^{\infty} f(t) m(dt) = \sum_{k=0}^{N} f(t_k) \). Also, denote by \( \nu \) a random measure, defined on the space \( \mathcal{H} \), which satisfies,

\[
\int_{\mathcal{H}} f(b) \nu(u, db) = f(Z_u).
\]

The value of this integral depends on the pair \((A, \Pi_0)\).

Using these measures, we may write the dynamics of \( Z \) in the form,

\[
Z_t = Z_0 + \int_{\mathcal{H}} \int_{[0,t]} (b - Z_{u^-}) \nu(u, db) m(du).
\] (8.8)

The following theorem gives a semimartingale representation for \( Z \).

**Theorem 8.1.1.** The continuous-time state process \( Z \), defined at (8.7), admits a semimartingale representation, \( Z_t = Z_0 + D_t + M_t \). Here, \( D \) is a process of bounded variation and \( M \) is a martingale with respect to the filtration defined by \( \mathcal{F}_t := \sigma\{Z_u : 0 \leq u \leq t\} \). The dynamics for these processes are, respectively,

\[
D_t = \int_{[0,t]} (AZ_{u^-} - Z_{u^-}) m(du)
= \int_{[0,t]} (A - I)Z_{u^-} m(du),
\]

\[
M_t = \int_{[0,t]} \int_{\mathcal{H}} (b - Z_{u^-}) \nu(u, db) m(du) - \int_{[0,t]} (AZ_{u^-} - Z_{u^-}) m(du).
\]

The symbol \( I \) denotes an \( M \times M \) identity matrix.

**Proof.** The representation \( Z_t = Z_0 + D_t + M_t \) is obtained by first adding, then subtracting, the term \( \int_{[0,t]} (AZ_{u^-} - Z_{u^-}) m(du) \) from the right hand side of Equation (8.8). To prove this is a semimartingale representation, we begin by
writing $M_t$ in a simpler form,

$$M_t = \int_{[0,t]} \int (b - Z_{u-}) \nu(u,db) m(du) - \int_{[0,t]} (AZ_{u-} - Z_{u-}) m(du)$$

$$= \sum_{0 < t_k \leq t} (Z_{t_k} - Z_{t_{k-}}) - \sum_{0 < t_k \leq t} (AZ_{t_k-} - Z_{t_k-})$$

$$= \sum_{0 < t_k \leq t} (Z_{t_k} - AZ_{t_{k-}}). \quad (8.9)$$

It is immediate from the last line above that $M_t$ is both $\mathcal{F}_t$-measurable and integrable. As $M$ is a pure jump process, we need only check the martingale condition at the discrete times, $t_k$. Equation (8.9) gives,

$$M_{t_k} = M_{t_{k-1}} + Z_{t_k} - AZ_{t_{k-1}}.$$

As $\mathbb{E}[Z_{t_k}|\mathcal{F}_{t_{k-1}}] = AZ_{t_{k-1}}$, from Equation (8.6), we have,

$$\mathbb{E}[M_{t_k}|\mathcal{F}_{t_{k-1}}] = M_{t_{k-1}}.$$

This shows that the process $M$ is a martingale, as required. Also, it is clear that $D$ is a process of bounded variation, so the representation $Z_t = Z_0 + D_t + M_t$ is a semimartingale decomposition for $Z_t$. \hfill \Box

Note that when we are integrating $Z_t$ with respect to a continuous measure, we may replace $Z_t$ by $Z_{t-}$ without changing the value of the integral, as the two are equal except on a set of measure zero. We shall use this fact throughout the chapter.

### 8.1.3 Reference Probability

Define the following filtrations, or ‘histories’:

$$\mathcal{F} = \{\mathcal{F}_t : 0 \leq t \leq T\}, \text{ with } \mathcal{F}_t = \sigma\{Z_u : 0 \leq u \leq t\},$$

$$\mathcal{Y} = \{\mathcal{Y}_t : 0 \leq t \leq T\}, \text{ with } \mathcal{Y}_t = \sigma\{Y_u : 0 \leq u \leq t\},$$

$$\mathcal{G} = \{\mathcal{G}_t : 0 \leq t \leq T\}, \text{ with } \mathcal{G}_t = \sigma\{Y_u, Z_u : 0 \leq u \leq t\}. $$
We introduce a new probability measure, \( P^\dagger \), on the space \((\Omega, \mathcal{F})^2\) using the Doléans-Dade exponential,

\[
\frac{dP^\dagger}{dP} \bigg|_{\mathcal{G}_t} = \Lambda_t^{-1} := \exp \left( - \int_0^t \langle \xi, Z_u \rangle \, dW_u - \frac{1}{2} \int_0^t \langle \xi, Z_u \rangle^2 \, du \right).
\] (8.11)

Girsanov's theorem then ensures that the process \( W_t + \int_0^t \langle \xi, Z_u \rangle \, du \) is Brownian motion under the measure \( P^\dagger \). In other words, the observation process, \( Y \), is Brownian motion under \( P^\dagger \) (from Equation 8.5). All the parameter estimates we calculate use the measure change,

\[
\frac{dP}{dP^\dagger} \bigg|_{\mathcal{G}_t} = \Lambda_t,
\]

to convert to the measure \( P^\dagger \), as this simplifies calculations significantly. This is achieved using Bayes rule. For example, the estimate \( E_t [\gamma_t \mid \mathcal{Y}_t] \) of a \( \mathcal{G} \)-adapted process \( \gamma = \{ \gamma_u : 0 \leq u \leq T \} \), given the observations to time \( t \), is computed by,

\[
E_t [\gamma_t \mid \mathcal{Y}_t] = \frac{E^\dagger_t [\Lambda_t \gamma_t \mid \mathcal{Y}_t]}{E^\dagger_t [\Lambda_t \mid \mathcal{Y}_t]} := \frac{\sigma_t (\gamma_t)}{\sigma_t (1)}.
\] (8.12)

We use the symbol \( \sigma_t (\gamma_t) \) to denote the expectation \( E^\dagger_t [\Lambda_t \gamma_t \mid \mathcal{Y}_t] \). As all calculations are carried out using this new measure, \( P^\dagger \) is called the reference probability.

When evaluating such expectations, we will need to know the dynamics of \( \Lambda \). The dynamics of the process \( \Lambda^{-1} \) are immediately given by,

\[
\Lambda_t^{-1} = 1 - \int_0^t \Lambda_u^{-1} \langle \xi, Z_u \rangle \, dW_u,
\]

(see, for example, Chapter 13 of Elliott [26]). Consequently, from Itô’s Lemma, the process \( \Lambda \) has dynamics,

\[
\Lambda_t = 1 + \int_0^t \Lambda_u \langle \xi, Z_u \rangle \, dY_u.
\] (8.13)

To conclude this section, we give a theorem which summarises the various dynamics under the reference probability.

---

\(^2\)The space \((\Omega, \mathcal{F})\) is the space on which the asset price is defined.
**Theorem 8.1.2.** Under the reference probability $P^1$, the observation process, $Y$, is a standard Brownian motion, and the dynamics of the state process, $Z$, remain unchanged.

**Proof.** Girsanov’s Theorem asserts that the process $Y$ is a standard Brownian motion under the measure $P^1$.

To show that the dynamics of the state process, $Z$, do not change under the measure $P^1$, let $X$ denote an arbitrary $\mathcal{F}_T$-measurable random variable and let $\phi$ denote a real-valued, bounded Borel-measurable function on $\mathbb{R}$. Then, using the Radon-Nikodym derivative and conditioning first on the $\sigma$-algebra $\mathcal{F}_T$ gives,

$$E^1[\phi(X)] = E[\phi(X)\Lambda^{-1}_T] = E\left[E[\phi(X)\Lambda^{-1}_T | \mathcal{F}_T]\right].$$

The function $\phi(X)$ is $\mathcal{F}_T$-measurable, so the expectation becomes,

$$E^1[\phi(X)] = E[\phi(X)E[\Lambda^{-1}_T | \mathcal{F}_T]].$$

Inspection of Equation (8.11) reveals that the process $\Lambda^{-1}$ is a $\mathcal{F}_T \vee \mathcal{Y}_t$-martingale, so that $E[\Lambda^{-1}_T | \mathcal{F}_T] = E\left[E[\Lambda^{-1}_T | \mathcal{F}_T \vee \mathcal{Y}_0] | \mathcal{F}_T\right] = E[\Lambda^{-1}_T | \mathcal{F}_T] = 1$. We then obtain,

$$E^1[\phi(X)] = E[\phi(X)].$$

Setting $X = Z_t$ and $\phi(X) = I(Z_t \leq \alpha)$ for any $\alpha \in \mathbb{R}$ and $t \in [0,T]$ gives,

$$P^1[Z_t \leq \alpha] = P[Z_t \leq \alpha],$$

which shows that distribution of $Z$ remains unchanged under the measure $P^1$. \hfill \Box

### 8.2 Estimation of the Market State

In this section we consider how to estimate the market state (that is, the state of the hidden Markov chain), and hence the volatility state of the model, at a given time. Section 8.2.1 finds the *filtered* estimate of the market state, that
is, the expected value of the state conditioned on the observations available up until the estimation time. The value of \( E[Z_t | \mathcal{Y}_t] \) is computed using Bayes rule to change to the reference probability,

\[
E[Z_t | \mathcal{Y}_t] = \frac{E[^t[\Lambda_t Z_t | \mathcal{Y}_t]]}{E[^t[\Lambda_t | \mathcal{Y}_t]]}.
\]

All expectations throughout the chapter are evaluated in this way.

Section 8.2.2 calculates the robust filtered estimate using a process introduced by Clark [17], which eliminates the stochastic integration involved in the filter. This gives a version of the estimate which is locally Lipschitz-continuous with respect to the observations, and so is suitable to use in practice.

In Section 8.2.3, we estimate the market state at a given time, \( t \), using all the observations available to a time \( T > t \). This is called the smooth estimate of the market state, and involves the filtered estimate, as well as an additional process incorporating the information obtained between the times \( t \) and \( T \). We will see the advantage of the smoother when estimating the model parameters in Section 8.3.

In Section 8.2.4, time-discretised versions of the filter and smoother are given.

### 8.2.1 Filtered Estimate

Using the notation \( q_t = E[^t[\Lambda_t Z_t | \mathcal{Y}_t]] \in \mathbb{R}^M \), the filtered estimate of the market state is,

\[
E[Z_t | \mathcal{Y}_t] = \frac{E[^t[\Lambda_t Z_t | \mathcal{Y}_t]]}{E[^t[\Lambda_t | \mathcal{Y}_t]]} = \frac{q_t}{\langle q_t, 1 \rangle}.
\]  

(8.16)

The fact that \( E[^t[\Lambda_t | \mathcal{Y}_t]] = \langle q_t, 1 \rangle \) follows from the identity \( \langle Z_t, 1 \rangle = 1 \). We see from Equation (8.16) that finding the dynamics of the process \( q \) is sufficient to evaluate the filtered market state estimate. The following theorem gives these dynamics.

**Theorem 8.2.1.** Let \( q_t = E[^t[\Lambda_t Z_t | \mathcal{Y}_t]] \). Then the process \( q \) satisfies the
stochastic integral equation,

\[ q_t = q_0 + \int_{[0,t]} (A - I) q_{u-} m(du) + \int_{[0,t]} \text{diag}\{\xi_i\} q_{u-} dY_u. \]

**Proof.** We first find the dynamics of \( \Lambda_t Z_t \) using the general product rule, given in Chapter 12 of Elliott’s book [26], and then take the conditional expectation under the measure \( P^\dagger \), given the information \( \mathcal{Y}_t \), to prove the result. We have,

\[ \Lambda_t Z_t = \Lambda_0 Z_0 + \int_{[0,t]} Z_{u-} \Lambda_u d\Lambda_u + \int_{[0,t]} \Lambda_u - dZ_u. \]

Substituting in the dynamics of \( Z_u \) and \( \Lambda_u \) from Theorem 8.1.1 and Equation (8.13) gives,

\[ \Lambda_t Z_t = \Lambda_0 Z_0 + \int_{[0,t]} Z_{u-} \Lambda_u - \langle \xi, Z_{u-} \rangle dY_u \]

\[ + \int_{[0,t]} \Lambda_u - dM_u + \int_{[0,t]} \Lambda_u - (A - I) Z_{u-} m(du). \]

Now take the conditional expectation under the measure \( P^\dagger \), given the information \( \mathcal{Y}_t \). From Lemma A.3.2 in Appendix 3, the third term equals zero, and in the remaining terms, we may apply Fubini’s theorem and first take the expectation before integrating. As the process \( Y \) is Brownian motion under the measure \( P^\dagger \), it has independent increments, which allows us, for example, to replace \( \mathcal{Y}_t \) by \( \mathcal{Y}_0 \) in the expectation \( E^\dagger[\Lambda_0 Z_0 | \mathcal{Y}_t] \). (Arguments such as these are given on page 261 of the book by Wong and Hajek ([80])). This gives,

\[ E^\dagger[\Lambda_t Z_t | \mathcal{Y}_t] = E^\dagger[\Lambda_0 Z_0 | \mathcal{Y}_0] + \int_{[0,t]} E^\dagger[\Lambda_u - Z_{u-} \langle \xi, Z_{u-} \rangle | \mathcal{Y}_{u-}] dY_u \]

\[ + \int_{[0,t]} (A - I) E^\dagger[\Lambda_u - Z_{u-} | \mathcal{Y}_{u-}] m(du). \]

Using the definition of \( q \) and Lemma A.3.1 of Appendix 3, we obtain,

\[ q_t = q_0 + \int_{[0,t]} (A - I) q_{u-} m(du) + \int_{[0,t]} \text{diag}\{\xi_i\} q_{u-} dY_u, \]

as required. \( \square \)

Using the dynamics for \( q \) given in this theorem, the market state estimate can be calculated via Equation (8.16). However, the dynamics of the process
$q$ involve a stochastic integral, which can lead to implementation difficulties. These are avoided by using the robust filtered estimate, which we will now define.

### 8.2.2 Robust Filtered Estimate

The robust filtered estimate is obtained by following the method introduced by Clark ([17]). The filter $q$ (of the previous section) is transformed to a new filter, $\overline{q}$, which has dynamics with the stochastic process appearing as a parameter rather than as an integrator. For this reason, the filter $\overline{q}$ is said to be robust, and Corollary 8.2.4 shows that using this filter gives a market state estimate which has local Lipschitz-continuous dependence upon the observations.

The following definition is required.

**Definition 8.2.2.** For $t \in [0, T]$, let $\Phi_t$ be the $M \times M$ diagonal matrix,

$$\Phi_t = \text{diag} \{ \phi_{t,1}, \phi_{t,2}, \ldots, \phi_{t,M} \},$$

where $\phi_{t,i} = \exp \left( \langle \xi, e_i \rangle Y_t - \frac{1}{2} \langle \xi, e_i \rangle^2 t \right)$ for $i = 1, 2, \ldots, M$. Note that,

1. $\Phi = (\Phi_t)$ is a continuous process, so we may interchange $\Phi_t$ and $\Phi_t-$.
2. $\Phi_t$ is actually a function of the observations, $Y$, and we write $\Phi_t(Y) = \Phi_t$.

The robust filter, $\overline{q}$, is then defined by the *gauge transform* (see Clark [17]),

$$\overline{q}_t = \Phi_t^{-1} q_t \text{ for all } t \in [0, T].$$

We see in following theorem that in the dynamics of $\overline{q}$, the stochastic process appears as a parameter of a deterministic integral equation.

**Theorem 8.2.3.** The process $\overline{q} = \Phi^{-1} q$ satisfies the deterministic integral equation,

$$\overline{q}_t = \overline{q}_0 + \int_{[0,t]} \Phi_u^{-1}(A - I)\Phi_u \overline{q}_{u-} m(du), \quad (8.18)$$

for all $t \in [0, T]$. Equivalently, for $k = 0, 1, \ldots, N$,

$$\overline{q}_t = \overline{q}_{t_k} = \Phi_{t_k}^{-1} A \Phi_{t_k} \overline{q}_{t_{k-1}} \text{ for all } t \in [t_k, t_{k+1}). \quad (8.19)$$
Proof. Under the measure $P^t$, $\phi_{t,i}$ is, by definition, a Doléans-Dade exponential with constant parameter $\langle \xi, e_i \rangle$. Thus we can write down the dynamics of $\Phi_t$:

$$d\Phi_t = \text{diag}\{\xi_i\}\Phi_t dY_t.$$ 

Using Itô's Lemma, $\Phi_t^{-1}$ has dynamics,

$$d\Phi_t^{-1} = \text{diag}\{\xi_i\}^2\Phi_t^{-1}dt - \text{diag}\{\xi_i\}\Phi_t^{-1}dY_t.$$ 

We then apply the product rule to find an equation for $\overline{\gamma}$:

$$\overline{\gamma}_t = \Phi_t^{-1}\overline{q}_t$$

$$= \overline{\gamma}_0 + \int_{[0,t]} \Phi_u^{-1}dq_u + \int_{[0,t]} d\Phi_u^{-1}dq_u - \text{diag}\{\xi_i\}q_{u-} dY_u$$

$$= \overline{\gamma}_0 + \int_{[0,t]} \Phi_u^{-1}(A-I)q_{u-} m(du) + \int_{[0,t]} \Phi_u^{-1}\text{diag}\{\xi_i\}q_{u-} dY_u$$

$$+ \int_{[0,t]} \text{diag}\{\xi_i\}^2\Phi_u^{-1}q_{u-} du - \int_{[0,t]} \text{diag}\{\xi_i\}\Phi_u^{-1}q_{u-} dY_u$$

$$- \int_{[0,t]} \text{diag}\{\xi_i\}\Phi_u^{-1}\text{diag}\{\xi_i\}q_{u-} du.$$ 

As all the matrices are diagonal, their order may be interchanged, proving Equation (8.18). Writing the integral as a sum, we see that,

$$\overline{\gamma}_t = \overline{\gamma}_0 + \sum_{0 < t_k < t} \Phi^{-1}_{t_k}(A-I)\Phi_{t_k}\overline{\gamma}_{t_{k-1}},$$

which shows that $\overline{\gamma}$ is constant for $t \in [t_k, t_{k+1})$. At the discrete times $t_k$ where switches occur, we have,

$$\overline{q}_{t_k} = \overline{q}_{t_{k-1}} + \Phi^{-1}_{t_k}(A-I)\Phi_{t_k}\overline{\gamma}_{t_{k-1}}$$

$$= \Phi^{-1}_{t_k}A\Phi_{t_k}\overline{\gamma}_{t_{k-1}}.$$ 

This completes the proof. $\square$

The following corollary shows that, using $\overline{\gamma}$, we can define a version of the expectation $\mathbb{E}[Z_t|\mathcal{Y}_t]$ which is “robust” in the sense that it has local Lipschitz-continuous dependence upon the observation sample path.
Corollary 8.2.4. For $0 \leq t \leq T$, the expectation $E[Z_t | \mathcal{Y}_t]$ can be computed using the process $\mathbf{q}$ according to the equation,

$$E[Z_t | \mathcal{Y}_t] = \frac{\Phi_t \mathbf{q}_t}{\langle \Phi_t \mathbf{q}_t, 1 \rangle}.$$  

This defines a version of the expectation which has local Lipschitz-continuous dependence upon the observations.

Proof. First, the equation,

$$E[Z_t | \mathcal{Y}_t] = \frac{\Phi_t \mathbf{q}_t}{\langle \Phi_t \mathbf{q}_t, 1 \rangle},$$

is obtained by applying Bayes rule to convert to the measure $P^t$, and then using the definitions of $q$ and $\mathbf{q}$.

The remainder of the proof is similar to those given in the papers by Clark ([17]) and James, Krishnamurphy and Le Gland ([49]). However, these proofs use Gronwall’s inequality to prove Lipschitz-continuity, which we cannot apply as the functions are discontinuous. Instead, we use an induction argument.

For functions $y : [0, T] \to \mathbb{R}$, which represent a particular observation sample path of the model given in Equation (8.5), define $B_t(y) = \Phi_t^{-1}(y) A \Phi_t(y)$. We see that the function $B_t(y)$ is locally Lipschitz-continuous in $y$, as $\Phi_t(y)$ is a diagonal matrix with continuous elements of the form $e^{c y - \frac{1}{2} c^2 t}$. This and the local boundedness of $\Phi_t(y)$, $B_t(y)$, and (from Equation (8.19)) $\mathbf{q}_t(y)$ will be needed in the proof.

At the first time step we use Equation (8.19) to give,

$$\|\mathbf{q}_{t_1} (y_1) - \mathbf{q}_{t_1} (y_2)\| = \|B_{t_1}(y_1) \mathbf{q}_0(y_1) - B_{t_1}(y_2) \mathbf{q}_0(y_2)\| \leq \|B_{t_1}(y_1) - B_{t_1}(y_2)\| \|\mathbf{q}_0(y_2)\|,$$

as $\mathbf{q}_0$ is constant. The local Lipschitz-continuity of $\mathbf{q}_{t_1}$ then follows from the boundedness of $\mathbf{q}_0$ and Lipschitz-continuity of $B_{t_1}(y)$.

Now suppose $\mathbf{q}_{t_{k-1}}$ is locally Lipschitz-continuous. The above properties of
\(\overline{q}_t\) and \(B_t\), combined with the induction hypothesis, give,
\[
\|\overline{q}_t(y_1) - \overline{q}_t(y_2)\| = \|B_t(y_1)\overline{q}_{t-1}(y_1) - B_t(y_2)\overline{q}_{t-1}(y_2)\|
\leq \|B_t(y_1) - B_t(y_2)\| \|\overline{q}_{t-1}(y_2)\|
+ \|B_t(2)_{t-1}(y_1)\| \|\overline{q}_{t-1}(y_1) - \overline{q}_{t-1}(y_2)\|
\leq K(\|y_1\|, \|y_2\|) \|y_1 - y_2\|,
\]
which shows that \(\overline{q}_t\) is locally Lipschitz-continuous for all \(t \in [0, T]\).

To prove local Lipschitz-continuity of \(\text{E}[Z_t | \mathcal{Y}_t]\), we require the inner product \(\langle \Phi_t \overline{q}_t, 1 \rangle\) to be bounded away from zero. As \(\overline{q}_0 = \text{E}[Z_0]\), all elements of \(\overline{q}_0\) are non-negative and at least one is positive. The matrix \(B_t\) also has non-negative elements, and, as the columns of \(A\) sum to 1, each column has at least one non-zero element. We then deduce from Equation (8.19) that \(\langle \Phi_t \overline{q}_t, 1 \rangle \geq 0\).

### 8.2.3 Smoothed Estimate

Next, we calculate a smoothed estimate of the market state. Using Bayes rule, this is given by,
\[
\text{E}[Z_t | \mathcal{Y}_T] = \frac{\text{E}[\Lambda_{0,T} Z_t | \mathcal{Y}_T]}{\text{E}[\Lambda_{0,T} | \mathcal{Y}_T]}.
\] (8.24)

The symbol \(\Lambda_{0,T}\) denotes the Radon-Nikodym derivative written according the the notation,
\[
\Lambda_{t,t'} = \exp\left(\int^t_t \langle \xi, Z_u \rangle dY_u - \frac{1}{2} \int^t_t \langle \xi, Z_u \rangle^2 du\right).
\]

The following theorem explains how to calculate the smoothed state estimate using the filtered estimate, as well as an additional process incorporating the information obtained between the times \(t\) and \(T\). The advantage of the smoother over the filter will be discussed in Section 8.3.

**Theorem 8.2.5.** Suppose that the state process \(\{Z_t : 0 \leq t \leq T\}\) and asset price process \(\{S_t : 0 \leq t \leq T\}\) have dynamics as described in Theorem 8.1.1 and Equation (8.1) respectively. If the asset price is observed according to the
transformed dynamics given in Equation (8.5), the conditional expectation of
$Z_t$, given the observations to time $T$, is,
\[
E[Z_t|\mathcal{Y}_T] = \sum_{i=1}^{M} \left\{ \frac{\langle q_t, e_i \rangle \langle v_t, e_i \rangle}{\langle q_t, v_t \rangle} \right\} e_i, \quad (8.25)
\]
for any $0 \leq t \leq T$. Here, $q_t = \Phi_t \overline{q}_t$ with $\overline{q}$ satisfying the forward deterministic integral equation,
\[
\overline{q}_t = \overline{q}_0 + \int_{[0,t]} \Phi_u^{-1} (A - I) \Phi_u \overline{q}_u m(du), \quad \overline{q}_0 = \varphi, \quad (8.26)
\]
vectors $\overline{v}_t = \Phi_t^{-1} \overline{v}_t$ with $\overline{v}_t$ satisfying the backward deterministic integral equation,
\[
\overline{v}_t = \overline{v}_T + \int_{[t,T]} \Phi_u (A - I)^T \Phi_u^{-1} \overline{v}_u m(du), \quad \overline{v}_T = \Phi_T \mathbf{1}, \quad (8.27)
\]
and $\Phi_t = \text{diag}\{\phi_{t,1}, \phi_{t,2}, \ldots, \phi_{t,M}\}$, with $\phi_{t,i} = \exp\left(\langle \xi, e_i \rangle \gamma_t - \frac{1}{2} \langle \xi, e_i \rangle^2 t\right)$.

Proof. The proof proceeds by setting $r_t = E^\dagger[\Lambda_{0,T} Z_t | \mathcal{Y}_T]$, so that the smoothed state estimate is,
\[
E[Z_t|\mathcal{Y}_T] = \frac{r_t}{\langle r_t, \mathbf{1} \rangle}. \quad (8.28)
\]
We will find an equation for $r_t$ in terms of the filter, $q$, and a new process $v = (v_t)_{0 \leq t \leq T}$, with $v_t = (v^1_t, ..., v^M_t)^T \in \mathbb{R}^M$ defined by,
\[
v^i_t = \langle v_t, e_i \rangle = E^\dagger[\Lambda_{t,T} | \mathcal{Y}_T \text{ and } Z_t = e_i].
\]
The vector $v_t$ incorporates the information observed between the times $t$ and $T$.

Using properties of the Radon-Nikodym derivative and conditional expectation, we have,
\[
r_t = E^\dagger[\Lambda_{0,T} Z_t | \mathcal{Y}_T] \\
= E^\dagger[E^\dagger[\Lambda_{0,T} \Lambda_{t,T} Z_t | \mathcal{Y}_T \vee \mathcal{F}_t] | \mathcal{Y}_T] \\
= E^\dagger[\Lambda_{0,T} Z_t E^\dagger[\Lambda_{t,T} | \mathcal{Y}_T \vee \mathcal{F}_t] | \mathcal{Y}_T].
\]

Since the process $Z$ remains Markov under the measure $P^\dagger$, the inner expectation in the last line becomes $E^\dagger[\Lambda_{t,T} | \mathcal{Y}_T \vee \sigma\{Z_t\}]$. Using the identity
\[
\sum_{i=1}^{M} \langle Z_t, e_i \rangle = 1 \text{ we obtain,}
\]

\[
 r_t = \mathbb{E}^\dagger \left[ \Lambda_{0,t} \sum_{i=1}^{M} \langle Z_t, e_i \rangle Z_t \mathbb{E}^\dagger [\Lambda_{t,T} | \mathcal{Y}_T \vee \sigma \{ Z_t \}] | \mathcal{Y}_T] \right].
\]

As the inner product \( \langle Z_t, e_i \rangle \) is zero for \( Z_t \neq e_i \), the value \( Z_t \) may be replaced by \( e_i \), giving,

\[
 r_t = \sum_{i=1}^{M} \mathbb{E}^\dagger [\Lambda_{0,t} \langle Z_t, e_i \rangle | \mathcal{Y}_T] v_t^i e_i.
\]

Under the measure \( P^\dagger \), \( Y \) is a Brownian motion and therefore has independent increments, so we may replace the \( \sigma \)-algebra \( \mathcal{Y}_T \) by \( \mathcal{Y}_t \) in the above expectation.

With Equation (8.28) and the definition and properties of the filter, \( q \), this proves Equations (8.25) and (8.26).

It remains to find the dynamics of the process \( v \). To find these directly can be difficult (see [64], [65]), so instead we exploit a duality which is motivated by the time invariance of the inner product,

\[
\langle q_t, v_t \rangle = \langle r_t, 1 \rangle = \mathbb{E}^\dagger [\Lambda_{0,T} \langle Z_t, 1 \rangle | \mathcal{Y}_T] = \mathbb{E}^\dagger [\Lambda_{0,T} | \mathcal{Y}_T].
\]

The dual process, \( \overline{v} \), is defined so that,

\[
\langle \overline{v}_t, v_t \rangle = \langle q_t, v_t \rangle \text{ for all } t \in [0, T],
\]

from which we immediately see that \( \overline{v}_t = \Phi_t v_t \). The time invariance property implies that,

\[
\langle \overline{v}_{t_{k-1}}, \overline{v}_{t_{k-1}} \rangle = \langle \overline{v}_{t_k}, \overline{v}_{t_k} \rangle,
\]

and using the relation \( \overline{q}_{t_k} = \Phi_{t_k}^{-1} A \Phi_{t_k} \overline{q}_{t_{k-1}} \) and adjoint properties gives,

\[
\langle \overline{q}_{t_{k-1}}, \overline{q}_{t_{k-1}} \rangle = \langle \overline{q}_{t_k}, \Phi_{t_k} A^T \Phi_{t_k}^{-1} \overline{q}_{t_k} \rangle.
\]

Hence \( \overline{q}_{t_{k-1}} = \Phi_{t_k} A^T \Phi_{t_k}^{-1} \overline{q}_{t_k} \). A backward induction argument then shows that the process \( \overline{v} \) satisfies the backward deterministic integral equation,

\[
\overline{v}_t = \overline{v}_T + \int_{[t,T]} \Phi_u (A - I)^T \Phi_u^{-1} \overline{v}_u m(du), \quad \overline{v}_T = \Phi_T 1.
\]

\( \square \)
8.2.4 Time-Discretised Filter and Smoother

From previous sections, the filtered estimate of the market state at time $t$ is,

$$
E[Z_t | \mathcal{Y}_t] = \frac{q_t}{\langle q_t, 1 \rangle},
$$

and the smoothed estimate is,

$$
E[Z_t | \mathcal{Y}_T] = \sum_{i=1}^{M} \left( \frac{\langle q_t, e_i \rangle \langle v_t, e_i \rangle}{\langle q_t, v_t \rangle} \right) e_i.
$$

The processes $q$ and $v$ are calculated from the relationships $q_t = \Phi_t \overline{q}_t$ and $v_t = \Phi_t^{-1} \overline{v}_t$, where $\overline{q}$ and $\overline{v}$ satisfy the deterministic integral equations given in Theorem 8.2.5, and $\Phi_t$ is an $M \times M$ diagonal matrix with diagonal elements $\phi_{t,i} = \exp(\langle \xi_i, e_i \rangle Y_t - \frac{1}{2} \langle \xi_i, e_i \rangle^2 t)$ for $i = 1, 2, \ldots, M$.

In order to calculate the estimates given real data, we need discrete-time equations for $q$ and $v$. From Theorem 8.2.3, we already have a discrete-time relationship, $\overline{q}_{k+1} = \Phi_{k+1} A \Phi_{k} \overline{q}_k$, for the process $\overline{q}$ at the Markov chain transition times. Using the equation $q_t = \Phi_t \overline{q}_t$ then gives,

$$
q_{k+1} = A \Phi_{k+1} \Phi_k^{-1} q_k.
$$

Also, we saw in the proof of Theorem 8.2.5 that $\overline{v}_{k+1} = \Phi_{k+1} A^T \Phi_k^{-1} \overline{v}_k$, and so,

$$
v_{k+1} = \Phi_{k+1}^{-1} \Phi_k A^T v_k.
$$

Using the definition of the matrix $\Phi$, we have,

$$
q_k = A \text{diag} \left\{ \exp \left( \xi_i (Y_k - Y_{k-1}) - \frac{1}{2} \xi_i^2 (t_k - t_{k-1}) \right) \right\} q_{k-1},
$$

$$
v_{k-1} = \text{diag} \left\{ \exp \left( \xi_i (Y_{k-1} - Y_{k-1}) - \frac{1}{2} \xi_i^2 (t_k - t_{k-1}) \right) \right\} A^T v_k.
$$

These are exact recursions which can be used to estimate the market state at the transition times $t_k$, $k = 0, 1, \ldots, N$. The initial values are $q_0 = E[Z_0]$ and $v_0 = 1$.

To estimate the state at a time $t$ with $t < t_k$, note that the processes $\overline{q}$ and $\overline{v}$ are piecewise constant so that $\overline{q}_t = \overline{q}_{t_k-1}$ and $\overline{v}_t = \overline{v}_{t_k-1}$. Then the state estimate at time $t$ can be obtained from the equations $q_t = \Phi_t \overline{q}_{t_k-1}$ and $v_t = \Phi_t^{-1} \overline{v}_{t_k-1}$.
8.3 Estimation of the Transition Matrix and Volatility Vector

Having estimated the market state, we now consider how the parameters of our model (that is, the vector, $\xi$, and the transition matrix, $A$) can be estimated. In Section 8.3.1, the method of estimation is explained. We use the EM algorithm to find maximum likelihood estimates of the parameters. In our setting, this involves computation of filters for three quantities relating to the state process, which are computed in Section 8.3.2. Robust filters and discrete-time recursions for the filters are also given here.

In Section 8.3.3, smoother-based estimates of the parameters are found by incorporating the process $\bar{v}$ of the previous section. This is more efficient than the filter-based method, as we need only compute two, rather than four, recursions to evaluate the estimates.

We start with a brief review of maximum likelihood estimation and the EM algorithm, and explain its form within our context.

8.3.1 Maximum Likelihood Estimates and the EM Algorithm

Suppose $\{P_\theta : \theta \in \Theta\}$ is a family of probability measures, each defined on the measurable space $(\Omega, \mathcal{F})$ and each absolutely continuous with respect to a fixed measure, $P_0$. Further, suppose $\mathcal{Y} \subset \mathcal{F}$ is a $\sigma$-algebra containing information that has been observed. The maximum likelihood estimate (MLE) for the parameter $\theta$, given the observations $\mathcal{Y}$, is then defined by,

$$\hat{\theta} := \arg\max_{\theta \in \Theta} L(\theta),$$

where $L$ is the likelihood function,

$$L(\theta) := \mathbb{E}^{P_\theta} \left[ \frac{dP_\theta}{dP_0} \bigg| \mathcal{Y} \right].$$
For any $\theta \in \Theta$ and $A \in \mathcal{Y}$, we have,

$$P_{\theta}(A) = \mathbb{E}_{\theta}^b[I(A)] = \mathbb{E}_{\theta} \left[ \frac{dP_{\theta}}{dP_0} I(A) \right] = \mathbb{E}_{\theta} \left[ I(A) \mathbb{E}_{\theta} \left[ \frac{dP_{\theta}}{dP_0} \mid \mathcal{Y} \right] \right],$$

which shows that the MLE maximises the probability of the observations actually occurring.

In its classical form (see [20], [59]), the EM algorithm provides an iterative scheme to compute MLEs. Its main feature is that the sequence of likelihoods it generates is nondecreasing. This means that one is assured of convergence towards a stationary point on the log likelihood surface. In the seminal paper [20], the EM algorithm was described as a missing, or incomplete, data algorithm. It is worth noting precisely what the term missing data means in our context. Essentially we have data at all times, so data is not missing in any real sense, rather, the data said to be missing is the indirectly observed state process, $Z$.

The algorithmic form of the EM scheme creates a sequence, where at each stage, two steps, the Expectation step and Maximisation step, are computed. It is summarised by the following algorithm.

1. Set $l = 0$ and choose $\hat{\theta}_0$.

2. E-Step : Compute the function $Q(\theta, \hat{\theta}_l) = \mathbb{E}^{\hat{\theta}_l} \left[ \ln \left( \frac{dP_{\theta}}{dP_{\hat{\theta}_l}} \right) \mid \mathcal{Y} \right]$.

3. M-Step : Find $\hat{\theta}_{l+1} = \arg\max_{\theta \in \Theta} Q(\theta, \hat{\theta}_l)$.

4. Decide to stop, or increment $l$ by 1 and continue from step 2.

When using the EM algorithm to obtain MLEs for the vector, $\xi$, and transition matrix, $A$, of our model, we need to compute conditional expectations of three quantities related to the state process. These are,
1. $O^i_t$, the amount of time spent by the process $Z$ in state $e_i$ up to time $t$:

$$O^i_t = \int_{[0,t]} \langle Z_u, e_i \rangle \, du \in \mathbb{R}.$$

2. $N^{i,j}_t$, the number of transitions $e_i \rightarrow e_j$ of $Z$ (where $i \neq j$) up to time $t$:

$$N^{i,j}_t = \int_{[0,t]} \langle Z_u-, e_i \rangle \langle dZ_u, e_j \rangle \in \mathbb{R}.$$

3. $\Gamma^i_t$, the level integrals for the state $e_i$:

$$\Gamma^i_t = \int_{[0,t]} \langle Z_u, e_i \rangle \, dY_u \in \mathbb{R}.$$

Applying the methods presented in the papers of Campillo and Le Gland ([14]) and Elliott ([27]), updated parameter estimates from the EM algorithm are given by,

$$\langle \tilde{A} e_i, e_j \rangle := \mathbb{E} [ \langle A e_i, e_j \rangle | \mathcal{Y}_T ] = \frac{\mathbb{E} [ N^{i,j}_T | \mathcal{Y}_T ]}{\mathbb{E} [ O^{i}_T | \mathcal{Y}_T ]} = \frac{\sigma_T(N^{i,j}_T)}{\sigma_T(O^{i}_T)}, \quad (8.32)$$

$$\langle \tilde{\xi}, e_i \rangle := \mathbb{E} [ \langle \xi, e_i \rangle | \mathcal{Y}_T ] = \frac{\mathbb{E} [ \Gamma^i_T | \mathcal{Y}_T ]}{\mathbb{E} [ O^{i}_T | \mathcal{Y}_T ]} = \frac{\sigma_T(\Gamma^i_T)}{\sigma_T(O^{i}_T)}, \quad (8.33)$$

for $i, j \in \{1, 2, \ldots, M\}, i \neq j$. Recall the notation, $\sigma_T(H_t) := \mathbb{E} [ L_t H_t | \mathcal{Y}_t ]$, for a $\mathbb{G}$-adapted process $H = \{H_t : 0 \leq t \leq T\}$. Also, note that the estimator for $\langle A e_i, e_j \rangle$ is defined only for the off diagonal elements of the matrix, $A$. However, one can compute the estimated values for the diagonal elements of $A$ from the identity $\sum_{i=1}^{M} \langle A e_i, e_j \rangle = 1$, which holds for all $j$.

From Equations (8.32) and (8.33), we see that to calculate updated parameter estimates, we need to compute $\sigma_T(N^{i,j}_T)$, $\sigma_T(\Gamma^i_T)$ and $\sigma_T(O^{i}_T)$. However, direct computation of these values can be difficult, so instead we apply an alternative approach presented in the paper by Elliott ([27]), in which the values $\sigma_T(N^{i,j}_T Z_T)$, $\sigma_T(\Gamma^i_T Z_T)$ and $\sigma_T(O^{i}_T Z_T)$ are computed. Then, since our state process takes values in the set of unit vectors in $\mathbb{R}^M$, we note, for example, that,

$$\langle \sigma_T(N^{i,j}_T Z_T), 1 \rangle = \langle E^\dagger [ L_T N^{i,j}_T Z_T | \mathcal{Y}_T ], 1 \rangle$$

$$= E^\dagger [ L_T N^{i,j}_T \langle Z_T, 1 \rangle | \mathcal{Y}_T ]$$

$$= \sigma_T(N^{i,j}_T).$$
Analogous equations hold for $O^i_T$ and $\Gamma^i_T$. The parameter estimates are then given by,

\[
\langle \tilde{A} e_i, e_j \rangle = \frac{\langle \sigma_T(N^i_{T} Z_T), 1 \rangle}{\langle \sigma_T(O^i_{T} Z_T), 1 \rangle}, \tag{8.34}
\]

\[
\langle \tilde{\xi}, e_i \rangle = \frac{\langle \sigma_T(\Gamma^i_{T} Z_T), 1 \rangle}{\langle \sigma_T(O^i_{T} Z_T), 1 \rangle}. \tag{8.35}
\]

In the next section, we use these equations to find filter-based parameter estimates.

### 8.3.2 Filter-based Parameter Estimates

As previously discussed, to find estimates of the vector, $\xi$, and transition matrix, $A$, we need to find the values of $\sigma_T(N^i_{T} Z_T)$, $\sigma_T(\Gamma^i_{T} Z_T)$ and $\sigma_T(O^i_{T} Z_T)$. The dynamics of these processes can be computed as for the filter $q_t = E_t | \Lambda_t Z_t |$ of Section 8.2.1. The following theorem gives these dynamics, from which the estimates of $\xi$ and $A$ can be obtained.

**Theorem 8.3.1.** For $i, j \in \{1, 2, \ldots, M\}$ and $t \in [0, T]$, the vectors $\sigma_t(N^i_{T} Z_t)$, $\sigma_t(O^i_{T} Z_t)$ and $\sigma_t(\Gamma^i_{T} Z_t)$ satisfy, respectively, the stochastic integral equations:

\[
\sigma_t(O^i_{T} Z_t) = \int_{[0,t]} (A - I) \sigma_u (O^i_{u} Z_u) m(du) + \int_{[0,t]} q_{u, -} e_i du e_i + \int_{[0,t]} \text{diag} \{ \xi_i \} \sigma_u (O^i_{u} Z_u) dY_u,
\]

\[
\sigma_t(N^i_{T} Z_t) = \int_{[0,t]} q_{u, -} e_i \langle (A - I)e_i, e_j \rangle m(du) e_j + \int_{[0,t]} (A - I) \sigma_u (N^i_{u} Z_u) m(du) + \int_{[0,t]} \text{diag} \{ \xi_i \} \sigma_u (N^i_{u} Z_u) dY_u,
\]

and,

\[
\sigma_t(\Gamma^i_{T} Z_t) = \int_{[0,t]} \text{diag} \{ \xi_i \} q_{u, -} e_i du e_i + \int_{[0,t]} (A - I) \sigma_u (\Gamma^i_{u} Z_u) m(du) + \int_{[0,t]} \text{diag} \{ \xi_i \} \sigma_u (\Gamma^i_{u} Z_u) dY_u + \int_{[0,t]} q_{u, -} e_i dY_u e_i,
\]

where $q_t = E_t | \Lambda_t Z_t |$ is the filter of Section 8.2.1.
Proof. We need to find the dynamics of $\sigma_t(H_tZ_t)$ for $H_t = N_t^{i,j}$, $O_t^i$, and $\Gamma_t^i$. The proof for each follows the same method, and is analogous to the proof of Theorem 8.2.1. We first find the dynamics of $\Lambda_t H_t Z_t$ using the general product rule, given in Chapter 12 of Elliott’s book ([26]), and then take the conditional expectation under the measure $P^i$, given the information $\mathcal{Y}_t$, to obtain the equation for $\sigma_t(H_t Z_t)$.

Case (i): $H_t = O_t^i$

We first need the dynamics of $O_t^i Z_t$. The product rule gives,

$$O_t^i Z_t = \int_{[0,t]} O_u^i \, dZ_u + \int_{[0,t]} Z_u \, dO_u^i.$$

The bracket term is zero because $Z$ is a discrete-time process and $O_t^i$ is a continuous-time process. Substituting in the dynamics of $O_t^i$ and $Z$ gives,

$$O_t^i Z_t = \int_{[0,t]} O_u^i \, dM_u + \int_{[0,t]} O_u^i \, (A - I) Z_u \, m(du) + \int_{[0,t]} Z_u \langle Z_u, e_i \rangle \, du.$$

We then apply the product rule again to find the dynamics for $\Lambda_t O_t^i Z_t,$ take the conditional expectation under the measure $P^i$ given the information $\mathcal{Y}_t$, and invoke Lemmas A.3.1 and A.3.2 of Appendix 3 to obtain the result.

Case (ii): $H_t = N_t^{i,j}$

In this case, the bracket term in the expression for $N_t^{i,j} Z_t$ is non-zero, as both $Z$ and $N_t^{i,j}$ are discrete. The bracket term is,

$$[N_t^{i,j}, Z]_t = \sum_{0 < s \leq t} \Delta N_s^{i,j} \Delta Z_s.$$

The term $\Delta N_s^{i,j}$ is non-zero only when there is a transition $e_i \rightarrow e_j$, and at such times, $\Delta Z_s = e_j - e_i$. This gives,

$$[N_t^{i,j}, Z]_t = (e_j - e_i) \sum_{0 < s \leq t} \Delta N_s^{i,j}$$

$$= (e_j - e_i) N_t^{i,j}$$

$$= (e_j - e_i) \int_{[0,t]} \langle Z_u, e_i \rangle \langle dZ_u, e_j \rangle.$$
Using the product rule and simplifying, we have,

\[
N_t^{i,j} Z_t = \int_{[0,t]} N_u^{i,j} dZ_u + \int_{[0,t]} Z_u^{i,j} dN_u + [N^{i,j}, Z]_t
\]

\[
= \int_{[0,t]} N_u^{i,j} (A - I) Z_u m(du) + \int_{[0,t]} \langle Z_u, e_i \rangle \langle dM_u, e_j \rangle e_j
\]

\[
+ \int_{[0,t]} N_u^{i,j} dM_u + \int_{[0,t]} \langle Z_u, e_i \rangle \langle (A - I) Z_u, e_j \rangle m(du) e_j.
\]

The simplification involved replacing \( Z_u (Z_u, e_i) \) by \( (Z_u, e_i) e_i \), as the value is only non-zero when \( Z_u = e_i \). The remainder of the proof follows as for the previous case.

**Case (iii):** \( H_t = \Gamma_t^i \)

For the final case, the product rule applied to \( \Gamma_t^i Z_t \) gives,

\[
\Gamma_t^i Z_t = \int_{[0,t]} Z_u (Z_u, e_i) dY_u + \int_{[0,t]} \Gamma_u^i dM_u + \int_{[0,t]} (A - I) \Gamma_u^i Z_u m(du).
\]

A second application yields,

\[
\Lambda_t \Gamma_t^i Z_t = \int_{[0,t]} \Lambda_u d(\Gamma_u^i Z_u) + \int_{[0,t]} \Gamma_u^i Z_u d\Lambda_u + [\Gamma^i Z, \Lambda]_t.
\]

The bracket term here is \( \int_{[0,t]} Z_u (Z_u, e_i) \Lambda_u (\xi, e_i) du \), as \( Y \) is Brownian motion under the measure \( P^i \). We proceed as standard to obtain the result. \( \square \)

Using the equations given in the above theorem, we can compute \( \sigma_T(N_t^{i,j} Z_T), \)

\( \sigma_T(\Gamma_t^i Z_T) \) and \( \sigma_T(O_t^i Z_T), \) and thus estimate the transition matrix and volatility vector. Since these equations involve stochastic integrals, approximation schemes are required for their implementation. However, it is better to avoid the stochastic integration by applying the gauge transformation of Clark ([17]), as we did for the filter \( q_t \), to eliminate the stochastic integrals and obtain deterministic integral equations.

For \( H_t \in \{ N_t^{i,j}, O_t^i, \Gamma_t^i \} \), the gauge transformation is given by,

\[
\overline{\sigma}(H_t Z_t) = \Phi_t^{-1} \sigma_t(H_t Z_t).
\]
Here, as before, $\Phi_t$ is the diagonal matrix $\Phi_t = \text{diag}\{\phi_{t,1}, \phi_{t,2}, \ldots, \phi_{t,M}\}$ with diagonal elements, $\phi_{t,i}$, defined by,

$$\phi_{t,i} = \exp\left(\langle \xi, e_i \rangle Y_t - \frac{1}{2} \langle \xi, e_i \rangle^2 t \right).$$

The dynamics of the transformed processes $\bar{\sigma}_t(N^{i,j}_t Z_t)$, $\bar{\sigma}_t(\Gamma^i_t Z_t)$ and $\bar{\sigma}_t(O^i_t Z_t)$ are given in the Theorem 8.3.2.

**Theorem 8.3.2.** For $i, j \in \{1, 2, \ldots, M\}$ and $t \in [0, T]$, the dynamics of $\bar{\sigma}_t(H_t Z_t) = \Phi^{-1}_t \sigma_t(H_t Z_t)$, for $H_t = N^{i,j}_t$, $O^i_t$ and $\Gamma^i_t$, are given, respectively, by,

$$\bar{\sigma}_t(N^{i,j}_t Z_t) = \int_{[0,t]} \Phi^{-1}_u \langle q_u, e_i \rangle \langle (A - I) e_i, e_j \rangle m(du) e_j + \int_{[0,t]} \Phi^{-1}_u (A - I) \Phi_u \bar{\sigma}_{u-} (N^{i,j}_{u-} Z_{u-}) m(du),$$

$$\bar{\sigma}_t(O^i_t Z_t) = \int_{[0,t]} \langle \bar{\eta}_u, e_i \rangle du e_i + \int_{[0,t]} \Phi^{-1}_u (A - I) \Phi_u \bar{\sigma}_{u-} (O^i_{u-} Z_{u-}) m(du),$$

and,

$$\bar{\sigma}_t(\Gamma^i_t Z_t) = \int_{[0,t]} \Phi^{-1}_u (A - I) \Phi_u \bar{\sigma}_{u-} (\Gamma^i_{u-} Z_{u-}) m(du) + \int_{[0,t]} \langle \bar{\eta}_u, e_i \rangle dY_u e_i.$$

**Proof.** The proof is analogous to the proof of Theorem 8.2.3, which gives the dynamics of $\Phi^{-1}_t$ and then uses the product rule. We also use the identity,

$$\Phi^{-1}_u \langle q_u, e_i \rangle e_i = \langle \bar{\eta}_u, e_i \rangle e_i.$$

\[\square\]

The following corollary shows that using the processes $\bar{\sigma}_t(H_t Z_t)$ for $H_t = N^{i,j}_t$, $O^i_t$ and $\Gamma^i_t$, we can define versions of the expectations $E[H_t Z_t | \mathcal{Y}_t]$ which are “robust” in the sense that they have local Lipschitz-continuous dependence upon the observation sample path.
Corollary 8.3.3. For $0 \leq t \leq T$ and $H_t \in \{N_t^{i,j}, O_t^i, \Gamma_t^i\}$, the expectation $E[H_tZ_t|\mathcal{Y}_t]$ can be computed using the process $\overline{\sigma}_t(H_tZ_t)$ according to the equation,

$$E[H_tZ_t|\mathcal{Y}_t] = \Phi_t \overline{\sigma}_t(H_tZ_t) \Phi_t^{-1} \mathbf{1}.$$

This defines a version of the expectation which has local Lipschitz-continuous dependence upon the observations.

Proof. Arguments similar to those in the proof of Corollary 8.2.4 prove this result. The inequality,

$$\|\langle q_t(y), e_i \rangle\| \leq \|q_t(y)\|,$$

which follows from the norm definitions, is also needed. For the cases $H_t = O_t^i$ and $H_t = \Gamma_t^i$, the value of $\overline{\sigma}_t(H_tZ_t)$ is not piecewise constant, so both the switching times and times between must be considered. \(\square\)

To obtain the estimates of $A$ and $\xi$ in practice, we need to discretise the equations for $\overline{\sigma}_t(H_tZ_t)$, $H_t = N_t^{i,j}$, $O_t^i$, $\Gamma_t^i$, and calculate recursions for the filters $\sigma_t(H_tZ_t)$. These recursions are given in Theorem 8.3.4.

Theorem 8.3.4. At the switching times \(\{t_k : k = 0, 1, \ldots, N\}\), the filters $\sigma_t(H_tZ_t)$, for $H_t \in \{N_t^{i,j}, O_t^i, \Gamma_t^i\}$, satisfy, respectively, the recursive equations,

$$\sigma_{t_k}(N_t^{i,j}Z_t) = A\Phi_{t_k} \Phi_{t_k}^{-1} \sigma_{t_{k-1}}(N_{t_{k-1}}^{i,j}Z_{t_{k-1}}) + \langle q_{t_{k-1}}, e_i \rangle \langle (A - I)e_i, e_j \rangle e_j,$$

$$\sigma_{t_k}(O_t^iZ_t) = A\Phi_{t_k} \Phi_{t_k}^{-1} \sigma_{t_{k-1}}(O_{t_{k-1}}^iZ_{t_{k-1}}) + \Phi_{t_k} \Phi_{t_k}^{-1} \langle q_{t_{k-1}}, e_i \rangle (t_k - t_{k-1})e_i,$$

$$\sigma_{t_k}(\Gamma_t^iZ_t) = A\Phi_{t_k} \Phi_{t_k}^{-1} \sigma_{t_{k-1}}(\Gamma_{t_{k-1}}^iZ_{t_{k-1}}) + \Phi_{t_k} \Phi_{t_k}^{-1} \langle q_{t_{k-1}}, e_i \rangle (Y_{t_k} - Y_{t_{k-1}})e_i.$$

Here, the initial points, $\sigma_{t_0}(H_0Z_0)$, are all zero.

Proof. From the equation for $\overline{\sigma}_t(N_t^{i,j}Z_t)$ in Theorem 8.3.2, we see that,

$$\overline{\sigma}_{t_k}(N_t^{i,j}Z_t) = \overline{\sigma}_{t_{k-1}}(N_{t_{k-1}}^{i,j}Z_{t_{k-1}}) + \Phi_{t_k}^{-1} \langle q_{t_{k-1}}, e_i \rangle \langle (A - I)e_i, e_j \rangle e_j$$

$$+ \Phi_{t_k}^{-1} (A - I) \Phi_{t_k} \overline{\sigma}_{t_{k-1}}(N_{t_{k-1}}^{i,j}Z_{t_{k-1}})$$

$$= \Phi_{t_k}^{-1} A \Phi_{t_k} \overline{\sigma}_{t_{k-1}}(N_{t_{k-1}}^{i,j}Z_{t_{k-1}}) + \Phi_{t_k}^{-1} \langle q_{t_{k-1}}, e_i \rangle \langle (A - I)e_i, e_j \rangle e_j.$$
Using the identity $\bar{\sigma}_t(N_t^{i,j} Z_t) = \Phi_t^{-1} \sigma_t(N_t^{i,j} Z_t)$ proves the equation for $\sigma_{t_k}(N_{t_k}^{i,j} Z_{t_k})$.

Similarly, using the other results of Theorem 8.3.2 and the fact that $\mathbf{q}$ is piecewise constant gives,

$$\bar{\sigma}_{t_k}(O_{t_k}^i Z_{t_k}) = \Phi_{t_k}^{-1} A \Phi_{t_k} \bar{\sigma}_{t_{k-1}}(O_{t_{k-1}}^i Z_{t_{k-1}}) + \langle \mathbf{q}_{t_{k-1}}, e_i \rangle (t_k - t_{k-1}) e_i,$$

$$\bar{\sigma}_{t_k}(\Gamma_{t_k}^i Z_{t_k}) = \Phi_{t_k}^{-1} A \Phi_{t_k} \bar{\sigma}_{t_{k-1}}(\Gamma_{t_{k-1}}^i Z_{t_{k-1}}) + \langle \mathbf{q}_{t_{k-1}}, e_i \rangle (Y_k - Y_{t_{k-1}}) e_i.$$

Multiplying on the left by the matrix $\Phi_{t_k}$ completes the proof.

Using the recursive equations given in this theorem, we can calculate $\sigma_T(N_T^{i,j} Z_T)$, $\sigma_T(O_T^i Z_T)$ and $\sigma_T(\Gamma_T^i Z_T)$ and then evaluate the parameter estimates using Equations (8.34) and (8.35):

$$\langle \tilde{A} e_i, e_j \rangle = \frac{\langle \sigma_T(N_T^{i,j} Z_T), 1 \rangle}{\langle \sigma_T(O_T^i Z_T), 1 \rangle},$$

$$\langle \tilde{\xi}, e_i \rangle = \frac{\langle \sigma_T(\Gamma_T^i Z_T), 1 \rangle}{\langle \sigma_T(O_T^i Z_T), 1 \rangle}.$$

The filter-based EM algorithm is summarised by the following steps:

**Step 1** Set $l = 0$.

**Step 2** Choose an initial matrix $\tilde{A}_0$ and an initial vector $\tilde{\xi}_0$.

**Step 3** Using the estimates $\tilde{A}_l$, $\tilde{\xi}_l$, apply the recursions for the quantities $q_{t_k}, \sigma_{t_k}(N_{t_k}^{i,j} X_{t_k}), \sigma_{t_k}(O_{t_k}^i X_{t_k})$ and $\sigma_{t_k}(\Gamma_{t_k}^i X_{t_k})$, and compute the updates $\tilde{A}_{l+1}$ and $\tilde{\xi}_{l+1}$ according to the above two equations.

**Step 4** Decide to stop, or, increment $l$ by 1 and continue from step 3.

This algorithm is stopped when it is decided that the parameter estimates have the desired accuracy, using some suitable criterion. For example, we could decide to stop the algorithm when the difference between the two final estimates is less that 0.001.

### 8.3.3 Smoother-based Parameter Estimates

In this section we use smoother-based methods to calculate estimates of the parameters $\xi$ and $A$. In the last section, we saw that the parameter estimates
are calculated from the values of \( \langle \sigma_T(H_T Z_T), 1 \rangle \), for \( H_T = N_T^{i j}, O_T^i, \) and \( \Gamma_T^i \).

The smoother-based method makes use of the following identity,

\[
\langle \sigma_T(H_T Z_T), 1 \rangle = \langle \sigma_T(H_T Z_T), \Phi_T^{-1} \Phi_T 1 \rangle \\
= \langle \Phi_T^{-1} \sigma_T(H_T Z_T), \Phi_T 1 \rangle \\
= \langle \sigma_T(H_T Z_T), \overline{v}_T \rangle.
\]

Using this identity, with Equations (8.34) and (8.35), the smoothed estimates of the parameters \( A \) and \( \xi \) are given by,

\[
\langle \tilde{A}e_i, e_j \rangle = \frac{\langle \sigma_T(N_T^{i j} Z_T), \overline{v}_T \rangle}{\langle \sigma_T(O_T^i Z_T), \overline{v}_T \rangle},
\]

(8.41)

\[
\langle \tilde{\xi}, e_i \rangle = \frac{\langle \sigma_T(\Gamma_T^{i j} Z_T), \overline{v}_T \rangle}{\langle \sigma_T(O_T^i Z_T), \overline{v}_T \rangle}.
\]

(8.42)

The following theorem calculates these estimates by determining the dynamics of \( \langle \sigma_t(H_t Z_t), \overline{v}_t \rangle \), for \( H_t = N_t^{i j}, O_t^i, \) and \( \Gamma_t^i \).

**Theorem 8.3.5.** Suppose we have a sampled data set \( \{Y_{t_k} : k = 0, 1, \ldots, N\} \) from the observation period [0, T], where the sampling instants are coincident with the transition times of the hidden Markov chain. Given estimates \( \tilde{A}_n \) and \( \tilde{\xi}_n \) of the parameters \( A \) and \( \xi \), MLE updates, \( \tilde{A}_{n+1} \) and \( \tilde{\xi}_{n+1} \), are given, respectively, by,

\[
\langle \tilde{A}_{n+1} e_i, e_j \rangle = \frac{\sum_{k=1}^{N} \langle (\tilde{A}_n - I)e_i, e_j \rangle \langle q_{t_{k-1}}, e_i \rangle \langle v_{t_k}, e_j \rangle}{\sum_{k=1}^{N} \langle q_{t_{k-1}}, e_i \rangle \langle v_{t_{k-1}}, e_i \rangle (t_k - t_{k-1})},
\]

and,

\[
\langle \tilde{\xi}_{n+1}, e_i \rangle = \frac{\sum_{k=1}^{N} Y_{t_k} \langle q_{t_{k-1}}, e_i \rangle \langle (\tilde{A}_n - I)^T v_{t_k}, e_i \rangle}{\sum_{k=1}^{N} \langle q_{t_{k-1}}, e_i \rangle \langle v_{t_{k-1}}, e_i \rangle (t_k - t_{k-1})}
\]

\[
- \frac{\sum_{k=1}^{N} Y_{t_k} \langle (\tilde{A}_n - I)q_{t_{k-1}}, e_i \rangle \langle v_{t_k}, e_i \rangle}{\sum_{k=1}^{N} \langle q_{t_{k-1}}, e_i \rangle \langle v_{t_{k-1}}, e_i \rangle (t_k - t_{k-1})},
\]

for \( i, j \in \{1, 2, \ldots, M\}, i \neq j \). The values of \( q_{t_k} \) and \( v_{t_k} \) are calculated from the recursions,

\[
q_{t_k} = A \Phi_{t_k} \Phi_{t_{k-1}}^{-1} q_{t_{k-1}},
\]

\[
v_{t_{k-1}} = \Phi_{t_k} \Phi_{t_{k-1}}^{-1} A^T v_{t_k},
\]
using the estimates \( \hat{A}_n \) and \( \hat{\xi}_n \), and the initial values \( q_0 = E[Z_0] \) and \( v_T = 1 \).

**Proof.** To calculate estimates of the matrix \( A \) and vector \( \xi \), we see from Equations (8.41) and (8.42) that we need to compute the values of \( \langle \mathbf{\sigma}_T(O^i_{T} Z_T), \mathbf{\nu}_T \rangle \), \( \langle \mathbf{\sigma}_T(N^{i,j}_{T} Z_T), \mathbf{\nu}_T \rangle \) and \( \langle \mathbf{\sigma}_T(\Gamma^{i}_{T} Z_T), \mathbf{\nu}_T \rangle \). We will find equations for these values involving only the processes \( q \) and \( v \).

First, to find the value of \( \langle \mathbf{\sigma}_T(O^i_{T} Z_T), \mathbf{\nu}_T \rangle \), the product rule gives,

\[
d\langle \mathbf{\sigma}_t(O^i_{t} Z_t), \mathbf{\nu}_t \rangle = \langle d\mathbf{\sigma}_t(O^i_{t} Z_t), \mathbf{\nu}_t \rangle + \langle \mathbf{\sigma}_t(O^i_{t} Z_t), d\mathbf{\nu}_t \rangle.
\]

From Equation (8.27), the process \( \mathbf{\nu} \) satisfies,

\[
d\mathbf{\nu}_t = - \Phi_t (A - I)^T \Phi_t^{-1} \mathbf{\nu}_t m(dt).
\]

Making this substitution for \( d\mathbf{\nu}_t \), and the appropriate substitution for \( d\mathbf{\sigma}_t(O^i_{t} Z_t) \) from Theorem 8.3.2, gives,

\[
d\langle \mathbf{\sigma}_t(O^i_{t} Z_t), \mathbf{\nu}_t \rangle = \langle \mathbf{\theta}_t, e_i \rangle \, dt \, e_i + \Phi_t^{-1} (A - I) \Phi_t \mathbf{\sigma}_{t-}(O^i_{t-} Z_{t-}) \, m(dt), \mathbf{\nu}_t \rangle
- \langle \mathbf{\sigma}_{t-}(O^i_{t-} Z_{t-}), \Phi_t (A - I)^T \Phi_t^{-1} \mathbf{\nu}_t \rangle \, m(dt)\rangle
= \langle \mathbf{\theta}_t, e_i \rangle \langle \mathbf{\nu}_t, e_i \rangle \, dt
+ \langle \Phi_t^{-1} (A - I) \Phi_t \mathbf{\sigma}_{t-}(O^i_{t-} Z_{t-}) \, m(dt), \mathbf{\nu}_t \rangle
- \langle \Phi_t^{-1} (A - I) \Phi_t \mathbf{\sigma}_{t-}(O^i_{t-} Z_{t-}) \, m(dt), \mathbf{\nu}_t \rangle
= \langle \mathbf{\theta}_t, e_i \rangle \langle \mathbf{\nu}_t, e_i \rangle \, dt.
\]

Then,

\[
\langle \mathbf{\sigma}_T(O^i_{T} Z_T), \mathbf{\nu}_T \rangle = \int_{[0,T]} \langle \mathbf{\theta}_u, e_i \rangle \langle \mathbf{\nu}_u, e_i \rangle \, du.
\]

This integral is easily computed, as the processes \( \theta \) and \( \nu \) are piecewise constant.

Also, as \( \mathbf{\theta}_t = \Phi_t^{-1} \mathbf{\theta}_t \) and \( \mathbf{\nu}_t = \Phi_t v_t \), we have \( \langle \mathbf{\theta}_t, e_i \rangle \langle \mathbf{\nu}_t, e_i \rangle = \langle \mathbf{\theta}_t, e_i \rangle \langle \mathbf{\nu}_t, e_i \rangle \). This gives,

\[
\langle \mathbf{\sigma}_T(O^i_{T} Z_T), \mathbf{\nu}_T \rangle = \sum_{k=1}^{N} \langle q_{t_{k-1}}, e_i \rangle \langle v_{t_{k-1}}, e_i \rangle \, (t_k - t_{k-1}). \tag{8.43}
\]

Similarly,

\[
\langle \mathbf{\sigma}_T(N^{i,j}_{T} Z_T), \mathbf{\nu}_T \rangle = \int_{[0,T]} \langle (A - I) e_i, e_j \rangle \langle \mathbf{\theta}_u, e_i \rangle \langle v_u, e_j \rangle \, m(du)
= \sum_{k=1}^{N} \langle (A - I) e_i, e_j \rangle \langle q_{t_{k-1}}, e_i \rangle \langle v_{t_{k}}, e_j \rangle. \tag{8.44}
\]
To calculate \( \langle \sigma_T (\Gamma^i_T Z_T), \overline{v}_T \rangle \) we introduce a new process,

\[
\widetilde{\sigma}_t (\Gamma^i_T Z_t) := \sigma_t (\Gamma^i_T Z_t) - \langle \eta_t, e_i \rangle Y_t e_i.
\]

Using integration by parts in the equation for \( \widetilde{\sigma}_t (\Gamma^i_T Z_t) \) in Theorem 8.3.2, the dynamics of \( \widetilde{\sigma}_t (\Gamma^i_T Z_t) \) are,

\[
d\widetilde{\sigma}_t (\Gamma^i_T Z_t) = \Phi_t^{-1} (A - I) \Phi_t \sigma_{t-} (\Gamma^i_T Z_{t-}) \, m(dt) \\
- Y_t \langle \Phi_t^{-1} (A - I) \Phi_t \sigma_{t-}, e_i \rangle \langle e_i, v_t \rangle m(dt) \\
- \langle \sigma_{t-} (\Gamma^i_T Z_{t-}), \Phi_t (A - I)^T \Phi_t^{-1} \overline{v}_t \rangle m(dt).
\] (8.45)

The product rule and the equation for \( d\overline{v}_t \) then give,

\[
d \langle \widetilde{\sigma}_t (\Gamma^i_T Z_t), \overline{v}_t \rangle = \langle \Phi_t^{-1} (A - I) \Phi_t \sigma_{t-} (\Gamma^i_T Z_{t-}), \overline{v}_t \rangle m(dt) \\
- Y_t \langle \Phi_t^{-1} (A - I) \Phi_t \sigma_{t-}, e_i \rangle \langle e_i, \overline{v}_t \rangle m(dt) \\
- \langle \sigma_{t-} (\Gamma^i_T Z_{t-}), \Phi_t (A - I)^T \Phi_t^{-1} \overline{v}_t \rangle m(dt)
\] (8.46)

Taking adjoints and substituting \( \widetilde{\sigma}_t (\Gamma^i_T Z_t) = \sigma_t (\Gamma^i_T Z_t) - \langle \eta_t, e_i \rangle Y_t e_i \) gives,

\[
d \langle \widetilde{\sigma}_t (\Gamma^i_T Z_t), \overline{v}_t \rangle = \langle \sigma_{t-} (\Gamma^i_T Z_{t-}), \Phi_t (A - I)^T \Phi_t^{-1} \overline{v}_t \rangle m(dt) \\
- Y_t \langle \Phi_t^{-1} (A - I) \Phi_t \sigma_{t-}, e_i \rangle \langle e_i, \overline{v}_t \rangle m(dt) \\
- \langle \sigma_{t-} (\Gamma^i_T Z_{t-}), \Phi_t (A - I)^T \Phi_t^{-1} \overline{v}_t \rangle m(dt) \\
+ Y_t \langle \sigma_{t-}, e_i \rangle \langle e_i, \Phi_t (A - I)^T \Phi_t^{-1} \overline{v}_t \rangle m(dt)
\] (8.47)

so that,

\[
\langle \sigma_T (\Gamma^i_T Z_T), \overline{v}_T \rangle = \int_{[0,T]} Y_t \langle \sigma_{t-}, e_i \rangle \langle e_i, \Phi_t (A - I)^T \Phi_t^{-1} \overline{v}_t \rangle m(dt) \\
- \int_{[0,T]} Y_t \langle \Phi_t^{-1} (A - I) \Phi_t \sigma_{t-}, e_i \rangle \langle e_i, v_t \rangle m(dt).
\]

Using \( \sigma_t (\Gamma^i_T Z_t) = \widetilde{\sigma}_t (\Gamma^i_T Z_t) + \langle \eta_t, e_i \rangle Y_t e_i \) in the inner product \( \langle \sigma_T (\Gamma^i_T Z_T), \overline{v}_T \rangle \),
and the diagonal nature of the matrix $\Phi_t$, gives,

\[
\langle \mathbf{\sigma}_T(\Gamma_T^T Z_T), \mathbf{\sigma}_T \rangle = \int_{[0,T]} Y_t \langle q_{t-}, e_i \rangle \langle (A - I)^T v_t, e_i \rangle m(dt) + Y_T \langle q_T, e_i \langle v_T, e_i \rangle \\
- \int_{[0,T]} Y_t \langle (A - I) q_{t-}, e_i \rangle \langle v_t, e_i \rangle m(dt)
= \sum_{k=1}^{N} Y_{t_k} \langle q_{t_{k-1}}, e_i \rangle \langle (A - I)^T v_{t_k}, e_i \rangle + Y_T \langle q_T, e_i \langle v_T, e_i \rangle \\
- \sum_{k=1}^{N} Y_{t_k} \langle (A - I) q_{t_{k-1}}, e_i \rangle \langle v_{t_k}, e_i \rangle.
\] (8.48)

Substituting Equations (8.43), (8.44) and (8.48) into (8.41) and (8.42) proves the result.

The discrete-time equations for $q$ and $v$ are given in Section 8.2.4.

In general, the smoothed estimates offer an improvement over the filtered estimates, as all of the information on the interval $[0, T]$ is being used to compute estimates within this interval. Further, we see from the previous theorem that to obtain smoother-based estimates of the parameters $A$ and $\xi$, we need only compute recursions for the processes $q$ and $v$. This is a significant advantage over the filter-based scheme, which requires four recursions, for the processes $q$, $\sigma(N^{ij} X)$, $\sigma(O^i X)$ and $\sigma(\Gamma^i X)$.

In summary, the smoother-based EM algorithm reads:

**Step 1** Set $l = 0$.

**Step 2** Choose an initial matrix $\tilde{A}_0$ and an initial vector $\tilde{\xi}_0$.

**Step 3** Using the estimates $\tilde{A}_l$, $\tilde{\xi}_l$, apply the recursions for the quantities $q_{t_k}$ and $v_{t_k}$, and compute the updates $\tilde{A}_{l+1}$ and $\tilde{\xi}_{l+1}$ according to Theorem 8.3.5.

**Step 4** Decide to stop, or, increment $l$ by 1 and continue from step 3.
Chapter 9

Conclusion and Future Directions

In this thesis, a “Switching Black-Scholes” model of a price process was proposed and studied. This model extends the ordinary Black-Scholes model by allowing the drift and volatility parameters to vary between a finite number of possible values at known times, according to the state of a hidden Markov chain.

After describing the model, the majority of the thesis investigated various methods of pricing options when the underlying asset evolves according to this model. As the Markov chain incorporates a second source of uncertainty into the Black-Scholes model, the Switching Black-Scholes market is incomplete, and no unique option pricing methodology exists. In this thesis, three alternatives, the methods of mean-variance hedging, Esscher transforms and minimum entropy, were applied to price call options in a Switching Black-Scholes market. Using particular parameter values, and C programs, to compute the prices, the pricing methods were also compared, and the Switching Black-Scholes model was shown to generate implied volatility smiles.

In order to apply the option pricing formulae, the parameters of the Switching Black-Scholes model must be estimated. A smaller portion of the thesis examined this problem. Maximum likelihood estimates of the possible volatility states and transition matrix of the Markov chain were found using filtering
and reference probability techniques. An estimate of the state of the hidden Markov chain was also computed.

Arising from the work in this thesis are many possible topics for further research. Firstly, the fit of the model to real world data should be considered. This could be examined in various ways. For example, the model could be directly calibrated using the estimation algorithms in this thesis, and then the prices obtained from the option pricing formulae compared to the market prices, or the model could be calibrated so that the implied volatility curves obtained from the option pricing formulae replicate the market implied volatility curves. The second alternative here is equivalent to calibrating the model so that the computed and market option values agree.

In applying these methods for calibration, we would be assuming the size of the Markov chain state space and the times of the switches are known. But is this a reasonable assumption? Another possible research topic would involve estimation of these parameters\(^1\). In reality, it would also be wise to be aware of the consequences of incorrectly estimating the model parameters, and to know the accuracy of the estimates obtained.

Secondly, more research needs to be undertaken on implementation of the option pricing formulae. Currently, the computation time required for the C programs which compute the call option values explodes exponentially as the size of the Markov chain state space (which represents the number of possible drifts and volatilities) and the number of switches increase. To avoid this problem, it would be appropriate to search for more efficient algorithms to compute the price, or to find suitable approximations, and error estimates, for the call option pricing formulae. Two possible methods for doing this would be to use the method of Hull and White (see [48]), or to determine terms involved in the sums which are effectively zero. This is particularly important for the method of minimum entropy, as computation of the price in this case involves numerical

\(^1\)The size of the Markov chain state space could be estimated using the compensated log-likelihood methods discussed in the PhD thesis of Berlian Setiawaty ([75]).
approximation to integrals of the form,

\[ \int_{-\infty}^{c} e^{-\frac{1}{2}x^2 - at^2} dx. \]

Regarding the general model, it would be relevant to find a theorem characterising the martingale measures of the Switching Black–Scholes model, and hence find the spread of arbitrage-free prices for a contingent claim in this market. If this were not possible, an idea of the range of arbitrage-free prices would be suggested by applying other option pricing methodologies, such as the method of Elliott and Madan ([28]), and evaluating the results.

Finally, there are the obvious extensions to the model, such as allowing multiple assets, an unlimited number of market states, and stochastic interest rates. When including stochastic interest rates, it would be natural to also use a Markov switching model, with either the same or different switching times and transition probabilities as for the asset parameters.

Although by no means complete, the above discussion indicates the variety of research topics arising from this thesis on the Switching Black–Scholes model. Hopefully, in the future, some of these topics will be addressed.