Modelling Power Market and Pricing Electricity Derivatives in a Regime Switching Framework

Ahmed Sayfeddine HAMADA

Thesis submitted for the degree of
Doctor of Philosophy
in
Applied Mathematics
at
The University of Adelaide
(Faculty of Mathematical and Computer Sciences)

Department of Applied Mathematics

March 20, 2014
Contents

Signed Statement v

Acknowledgements vi

Dedication vii

Abstract viii

1 Introduction 1

2 The Stochastic discount factor 3
   2.1 Introduction ......................................................... 3
   2.2 Presenting the model and preliminary results .................... 6
      2.2.1 The bond price and short rate .............................. 7
      2.2.2 Risk-neutral measure .......................................... 11
   2.3 Deriving the stock price .......................................... 17
      2.3.1 Dividends payment model ..................................... 17
      2.3.2 Deriving the stock price ..................................... 18
      2.3.3 Explicit stock price ........................................... 20
      2.3.4 The martingale condition .................................... 24
      2.3.5 The stock price dynamics .................................... 25
   2.4 Pricing European stock options ................................. 26
      2.4.1 The pricing of a call option ................................. 26
2.4.2 Occupation times of a Markov process . . . . . . . . . . . . . 29
2.4.3 The characteristic function of the occupation time . . . . . . . 31
2.4.4 Parseval’s theorem and Fourier transforms . . . . . . . . . . . 34
2.4.5 The hedging of a European call . . . . . . . . . . . . . . . . 35
2.4.6 Exchange rates . . . . . . . . . . . . . . . . . . . . . . . . . . 38
2.4.7 Bond dynamics . . . . . . . . . . . . . . . . . . . . . . . . . . 41
2.4.8 The forward measure . . . . . . . . . . . . . . . . . . . . . . . 46
2.5 Results summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . 47
2.5.1 Risk neutral measure . . . . . . . . . . . . . . . . . . . . . . . 47
2.5.2 Stock price . . . . . . . . . . . . . . . . . . . . . . . . . . . . 47
2.5.3 European stock option . . . . . . . . . . . . . . . . . . . . . . 47
2.5.4 Bond prices and dynamics . . . . . . . . . . . . . . . . . . . . . 48
2.5.5 Exchange rates . . . . . . . . . . . . . . . . . . . . . . . . . . 48
2.5.6 Forward rates . . . . . . . . . . . . . . . . . . . . . . . . . . . 48
2.5.7 Forward measure . . . . . . . . . . . . . . . . . . . . . . . . . 49
2.5.8 Forward rates . . . . . . . . . . . . . . . . . . . . . . . . . . . 49
2.6 Conclusion and further research . . . . . . . . . . . . . . . . . . . . 49

3 Stochastic discount factor in power market 51
3.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51
3.2 Presenting the model . . . . . . . . . . . . . . . . . . . . . . . . . . 53
3.3 Pricing of the forward . . . . . . . . . . . . . . . . . . . . . . . . . . 55
3.4 Pricing of options on forward . . . . . . . . . . . . . . . . . . . . . . 59
3.5 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 63

4 Power pricing using compensated jump processes 65
4.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 65
4.2 Presenting the model . . . . . . . . . . . . . . . . . . . . . . . . . . 66
4.3 Power derivatives pricing . . . . . . . . . . . . . . . . . . . . . . . . . 67
4.3.1 Risk-neutral measure . . . . . . . . . . . . . . . . . . . . . . . . 67
5 Power pricing using filtering theory

5.1 Introduction ........................................... 76
5.2 Presenting the model ................................. 77
5.3 Forward pricing ....................................... 79
  5.3.1 Risk premium and change of measure .......... 80
  5.3.2 Recursive filter for hidden Markov process ..... 84
5.4 Pricing of options on forward ..................... 92
  5.4.1 Expectation based approximation ............... 92
  5.4.2 Deriving the call price .......................... 97
5.5 Parameter estimation ............................... 98
  5.5.1 Discretisation of the observed process $y_t$ .... 99
  5.5.2 Estimation ....................................... 103
  5.5.3 Estimation of the drift parameter $\mu$ .......... 104
  5.5.4 Estimation of the volatility term $\sigma$ ......... 110
  5.5.5 Estimation of the speed of return $\kappa$ ......... 113
  5.5.6 Estimation of the transition rate matrix of the process .... 114
5.6 Empirical analysis and simulation ................. 116
  5.6.1 Seasonality ..................................... 119
  5.6.2 Identifying the spikes .......................... 120
  5.6.3 Negative prices .................................. 120
  5.6.4 Sample path simulation .......................... 121
5.7 Conclusion ........................................... 121
Signed Statement

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

SIGNED: ..........................  DATE:  ..........................
Acknowledgements

I wish to thank my supervisor, Professor Robert Elliott, for his continuous support as well as patience in times when progress was rather slow.

Big thanks go to Professor Zudi Lu and Professor John van der Hoek for their valuable remarks, and to my brother Dr. Mahmoud Hamada for all his support and expertise.

Many thanks also go to the academic and research staff and colleagues at the university of Adelaide who shared many enlightening discussions.

I am grateful for the financial support I have received from the University of Adelaide (AFSI) and the Australian Research Council (ARC).

I wish to thank all my other friends in Adelaide for making this part of my life such a great time.
I dedicate this work to my parents,
sisters and brother
for their continuous support.
Abstract

The deregulation of power market has led to an increase in risk for both consumers and producers when trading the underlying. Random price variations require a proper risk hedging strategy; related securities like forwards, options and swaps are the main derivatives that investors resort to in order to reduce the risk. The electricity spot price however has a particular behaviour, a consequence of the physical nature of the underlying. The non elastic offer rate causes the market equilibrium price to jump to extreme high or low levels in addition to the mean reversion and seasonality effects.

After the Introduction to the thesis contents and the background given in Chapter I, Chapter II and III develop pricing using a stochastic discount factor with applications to power derivatives. Because of the multiple sources of randomness, the power market is incomplete and any risk neutral probability measure is not unique. Pricing derivatives under the historical measure using a stochastic discount factor is one way to overcome this issue. Chapter IV investigate a different type of power pricing model. We suggest a general form for the spot price model where the randomness is given by a compensated pure jump process. Chapter V considers a new model for electricity spot price driven by a unobserved Markov jump process and the jumps are modelled using an independent Markov chain driving the jump size. In the presence of an unobserved process, the calculation of the forward and option on forward prices is performed using filtering theory. A conclusion of the thesis is given in Chapter VI.
Chapter 1

Introduction

The power market has seen a remarkable expansion in terms of trading volume and varieties of related financial instruments since the later decades of the last century. The migration from a government controlled sector into a liberated market has brought the power producers and consumers from a situation where prices were predictable, being just the marginal cost of production, into a highly volatile market. The deregulation of the electricity market has left investors faced with randomness in the underlying commodity price as a result of the changing free market equilibrium between supply and demand.

All power resources have the same common characteristic: a non elastic output; It usually requires a few hours to change the generated electricity level. However demand varies much more quickly. This causes the market to lose equilibrium and consequently the power price can quickly jump to extremely high or low levels. This behaviour of power prices motivates the models we present in this thesis. Hopefully models presented in this thesis will contribute to hedging against market risk.

The thesis is organised as follows: Chapter II presents a new way of pricing derivatives when the underlying model has several sources of randomness so we cannot consider a complete market. We introduce a Markov regime switching stochastic discount factor as an alternative approach to price derivatives. We also derive the price of different financial instruments such as bonds, forwards, stocks with divi-
dends and options. A hedging strategy for the European call options is derived.

In Chapter III we introduce a Markov regime switching electricity spot price model.
We combine this with a stochastic discount factor model to calculate the price of a
forward contract. Finally, we price a call option on a forward contract.

Chapter IV presents a new general model for power prices. Modelling the spot price
as an exponential of a mean reverting, compensated jump process, we define a risk
neutral measure for such a model. Within such a framework, we develop the price of
a forward contract, derive the dynamics of the latter process and use it to determine
the price of a swap contract.

Chapter V adopts a different approach for power pricing. We introduce a mean
reverting, Markov regime switching model for the spot price. The long term mean
switches from one regime to another depending on an unobserved Markov process.
The observable jumps are modelled using an independent Markov jump chain acting
on a jump size vector. To derive the forward price, we introduce a recursive filter
which calculates the expected value of the future spot price. This depends on the
non-observable Markov process. A call option price is also derived in this model.
Finally we develop estimators for the model parameters using filtering theory.

Chapter VI concludes the thesis and suggests further possible research.
Chapter 2

The Stochastic discount factor

2.1 Introduction

In this chapter, we investigate the use of a stochastic discount factor as introduced in Cochrane and Hansen [13]. This can price assets in a frictionless-market paradigm, and model future price distributions under a non-arbitrage condition. Gourieroux and Monfort [36] considered pricing derivatives in discrete time when the discount factors are exponential affine functions of an underlying state variable. Bansal [1] introduced stochastic discount factor pricing as the inter-temporal marginal rate of substitution. In this work, we develop a kernel pricing approach, where the stochastic discount factor has continuous Markov chain-driven switching dynamics. Since the early work of Qandt [58], who estimated the switching point of time series data using the Likelihood function, and the work of Hamilton [14] who introduced Markov processes in modelling econometric series, several extensions have developed regime switching models as an econometric tool. These models proved useful in describing randomness resulting from possible future structural economic changes. Additionally, regime switching models are appropriate for describing financial and economic data having distributions with a fat tail as an outcome of introducing Markov chains to drive the switching over different states.
While asset pricing should be performed in a non arbitrage framework, in Harrison and Kreps [41] and Harrison and Pliska [38, 39] was shown the relationship between the no-arbitrage principle and the existence of an equivalent martingale measure, as well as the equivalence between model completeness and uniqueness of an equivalent martingale measure. Unlike the Black and Scholes [6] and Merton [51] models, where perfect hedging in a complete model was established for European style options, a model with regime switching loses completeness in the presence of several sources of uncertainty, or risk. In other words, for such a model there is more than one equivalent martingale measure. This is the case in our proposed model where the state switching is driven by a continuous time Markov chain.

Rather than using an equivalent martingale measure another approach, pioneered by Gerber and Shiu [35], consists of using the concept of the Esscher transform to find an equivalent probability measure under which the discounted price of each primitive security is a martingale. Previous literature uses Markov modulated regime switching for derivative pricing in cases where the underlying dynamics follows a Lévy process; for option pricing with regime switching see Elliott et al. [21]. For option pricing with pure jump processes under regime switching see Elliott and Osakwe [22]. For optimal portfolio allocation under regime switching see Elliott and Siu [23]. For a risk measure under a PDE approach with regime switching we refer to Elliott et al. [24].

A major part of this chapter addresses the pricing of European call options, assuming that the stock price at a fixed time is the net present value (NPV) of all future dividend payments. In other words, we consider the stock price itself as a derived quantity and the underlying dividends are also modelled as a regime switching geometric Brownian motion. This approach is more realistic than considering a straightforward stock model since, by construction, the model includes the stock price changes that result from discrete dividend payments. In fact, we show in this work that a Black and Scholes-like formula is obtained when calculating the option price, with parameters describing the regime switching.
Modelling the dividend payments has become a subject of interest in recent years. Previous work on asset pricing theory following the same approach includes Graziano and Rogers [37], who considered option pricing where the stock is the present value of future dividends, under the maximisation of the agent consumption CRRA utility using a state price density. Veronesi [63] considered a two-state continuous hidden Markov chain when modelling the dividend payments and used the rational expectation equilibrium to derive the equilibrium stock price function. In our work we start from modelling the stock price as the net present value of future regime switching dividend payments and we price a European call option using a stochastic discount function. The results we obtain are semi-analytical, given by an inverse Fourier transform to derive the conditional characteristic function.
2.2 Presenting the model and preliminary results

Suppose \((\Omega, \mathcal{F}, \{\mathcal{F}\}, \mathbb{P})\) is a complete filtered probability space, where \(\mathbb{P}\) is the real world probability measure. We assume that the economy jumps in time from one state to another, and that its evolution is described by a Markov chain \(X \in \mathbb{R}^N\). Without loss of generality, the state space of \(X\) can be represented by the set of unit vectors \(S = \{e_1, e_2, \ldots, e_N\}\) where \(e_i = (0, \ldots, 1, \ldots, 0)'\) and 1 is located at the \(i^{th}\) position of the vector. This means that when the economy is in the \(i^{th}\) state \(X_t = e_i\).

From Elliott et al.[20], we have the following semi-martingale representation for \(\{X_t\}_{t \geq 0}\):
\[
X_t = X_0 + \int_0^t AX_u du + M_t.
\] (2.2.1)

Here \(A = [a_{ji}]_{i,j=1,2,\ldots,N}\) is the transition rate matrix for the homogeneous Markov chain \(X_t\). That is if, for \(s,t > 0\) : \(P_{ji}(s) = P(X_{t+s} = j|X_t = i)\), and \(i \neq j\), then
\[
\frac{dP_{ji}(s)}{ds}|_{s=0^+} = a_{ji} \quad \forall i,j = 1, \ldots, N\] independently of \(s\). (2.2.2)

The process \(M\) is an \(\mathbb{R}^N\)-valued martingale with respect to the filtration generated by \(\{X_t\}\). In the regime switching model discussed below we shall assume that the parameters switch according to the state of the economy, that is, they are determined by \(X\).

A stochastic discount factor is an adapted process \(\pi = \{\pi_t|t \geq 0\}\) with the property that for any asset price process \(\{A_t|t \geq 0\}\) and for \(s \leq t\):
\[
\pi_s A_s = \mathbb{E}[\pi_t A_t|\mathcal{F}_s].
\] (2.2.3)

Here the expectation is with respect to the historical ‘real world’ probability \(\mathbb{P}\), and the \(\sigma\)-field \(\mathcal{F}_t\) represents the information about the price process up to time \(t\). Since, the process \(\pi\) is adapted, we can write:
\[
\pi_{s,t} = \frac{\pi_t}{\pi_s}
\]
and present an alternative definition of the stochastic discount factor \( \pi_{s,t} \) by:

\[
A_s = \mathbb{E}[\pi_{s,t}A_t|F_s]. \tag{2.2.4}
\]

That is, for a price process \( \{A_t\}_{t \geq 0} \), the price at time \( s \) of the asset worth \( A_t \) at time \( t \), is the expectation of the discounted asset price back to time \( s \) conditioning on the information available up to time \( s \).

Consider parameter vectors \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_N)' \in \mathbb{R}^N \) and \( \theta = (\theta_1, \theta_2, \ldots, \theta_N)' \in \mathbb{R}^N \). Suppose \( W = \{W_t, t \geq 0\} \) is a one dimensional Brownian motion and consider a stochastic discount function (SDF) of the following form:

\[
\pi_{t,s} = \exp \left( \int_t^s \langle \theta, X_u \rangle du + \int_t^s \langle \gamma, X_u \rangle dW_u \right). \tag{2.2.5}
\]

Here \( \langle \theta, X_u \rangle \) and \( \langle \gamma, X_u \rangle \) denote scalar products in \( \mathbb{R}^N \). \( \gamma^2 \) will denote the vector with components \( \gamma_i^2 \).

\( \pi_{t,s} \) is a generalisation of the stochastic discount factor \( \Lambda \) which provides the Girsanov transform to the risk neutral probability in the Black and Scholes model, where

\[
\Lambda_t = \exp \left( -\int_0^t \theta dW_s - \frac{1}{2} \int_0^t \theta^2 ds \right) \quad \text{with} \quad \theta = \frac{\mu - r}{\sigma}.
\]

Then, by Ito’s lemma

\[
d\pi_{t,u} = \pi_{t,u} \left[ \left( \langle \theta, X_u \rangle + \frac{1}{2} \langle \gamma^2, X_u \rangle \right) du + \langle \gamma, X_u \rangle dW_u \right]. \tag{2.2.6}
\]

In our model there are two sources of randomness: the Brownian motion and the Markov chain. Therefore, we define the following filtrations which are taken to be right continuous and complete. Write \( \{F_t^X\} \) for the filtration generated by the chain \( \{X_t\}_{t \geq 0} \) and \( \{F_t^W\} \) the filtration generated by the Brownian motion \( \{W_t\}_{t \geq 0} \). Denote by \( \{F_t\} \) the filtration \( \{F_t = \{F_t^X \lor F_t^W\} \).

We shall now use our model for the SDF to derive some results and derivative prices.

### 2.2.1 The bond price and short rate

Using the above model for the stochastic discount factor, we shall calculate the price of the zero-coupon bond, and consequently calculate the forward interest rate.
Please refer to [65] for the definitions and relation between the bond price and the short rate.

The bond price

The arbitrage-free price of a bond with maturity \( T \) at time \( t \leq T \) is the value of the financial instrument with payoff equal to 1 at time \( T \). See [65]. Therefore, the price can be expressed using the SDF as:

\[
P(t, T) = \mathbb{E}[\pi_{t,T}|\mathcal{F}_t]. \tag{2.2.7}
\]

Proposition 1. Define the matrix \( K \):

\[
K = A + \text{diag}(\theta + \frac{1}{2}\gamma^2).
\]

Then the price of a zero-coupon bond is:

\[
P(t, T) = \mathbb{E}[\pi_{t,T}|\mathcal{F}_t] = \left\langle \exp(K(T-t))X_t, 1 \right\rangle. \tag{2.2.8}
\]

Proof. With \( \pi \) scalar and \( X \) a vector. Consider the variable vector

\[
H_{t,T} = \pi_{t,T}X_T. \tag{2.2.9}
\]

so that \( \pi_{t,T} = \langle H_{t,T}, 1 \rangle = \pi_{t,T}\langle X_T, 1 \rangle \).

Recall from (2.2.1) that the dynamics of the Markov chain \( X \) are:

\[
dX_u = AX_u du + dM_u.
\]

As \( \pi \) has no jumps:

\[
dH_{t,u} = \pi_{t,u}dX_u + X_u d\pi_{t,u}
\]

\[
= \pi_{t,u}AX_u du + \pi_{t,u}dM_u + X_u\pi_{t,u} \left[ \langle \gamma, X_u \rangle dW_u + \langle \theta + \frac{1}{2}\gamma^2, X_u \rangle du \right]
\]

\[
= \left( A + \langle \theta + \frac{1}{2}\gamma^2, X_u \rangle \right) H_{t,u} du + \pi_{t,u}dM_u + H_{t,u}\langle \gamma, X_u \rangle dW_u
\]

\[
= \left( A + \text{diag}(\theta + \frac{1}{2}\gamma^2) \right) H_{t,u} du + \text{diag}(\gamma)H_{t,u}dW_u + \pi_{t,u}dM_u.
\]
The integral form of $H_{t,T}$ is then:

$$H_{t,T} = X_t + \int_t^T \left( A + \text{diag}(\theta + \frac{1}{2} \gamma^2) \right) H_{t,u} du + \int_t^T \text{diag}(\gamma) H_{t,u} dW_u + \int_t^T \pi_{t,u} dM_u.$$ 

The last two integrals are martingales with respect to $\{\mathcal{F}_t\}$; see [48, 56] for details; and [18, 20] for the case of Markov and jump processes. Hence:

$$\mathbb{E}[\int_t^T \text{diag}(\gamma) H_{t,u} dW_u | \mathcal{F}_t] = \mathbb{E}[\int_t^T \pi_{t,u} dM_u | \mathcal{F}_t] = 0,$$

Therefore

$$\mathbb{E}[H_{t,T} | \mathcal{F}_t] = X_t + \int_t^T \left( A + \text{diag}(\theta + \frac{1}{2} \gamma^2) \right) \mathbb{E}[H_{t,u} | \mathcal{F}_t] du$$

(2.2.10)

With $K = \left( A + \text{diag}(\theta + \frac{1}{2} \gamma^2) \right)$.

Consequently, we obtain the following vector ordinary differential equation for $\mathbb{E}[H_{t,T} | \mathcal{F}_t]$:

$$\mathbb{E}[H_{t,T} | \mathcal{F}_t] = X_t + \int_t^T K \mathbb{E}[H_{t,u} | \mathcal{F}_t] du.$$

Suppose $\Psi(t, T)$ is the solution of the linear matrix ordinary differential equation:

$$\frac{d\Psi(t, T)}{dT} = K\Psi(t, T) \quad (2.2.11)$$

$$\Psi(t, t) = I_N.$$

Again, as $K$ is constant the solution of this system, see [20, 8] for details, is the exponential matrix:

$$\Psi(t, T) = \exp(K(T - t)).$$

Then

$$\mathbb{E}[H_{t,T} | \mathcal{F}_t] = \Psi(t, T)X_t \in \mathbb{R}^N \quad (2.2.12)$$

For more details on the definition, properties and methods of calculation of the exponential of a matrix, refer to [55] or see (A.1). Therefore,

$$\mathbb{E}[\pi_{t,T} | \mathcal{F}_t] = \mathbb{E}[\langle H_{t,T}, 1 \rangle | \mathcal{F}_t]$$

$$= \langle \mathbb{E}[H_{t,T} | \mathcal{F}_t], 1 \rangle.$$
So
\[ P(t, T) = \mathbb{E}[^\pi_t \pi_T | \mathcal{F}_t] = \left\langle \exp(K(T-t)) X_t, 1 \right\rangle. \]

The forward price and short rate

The implied forward rate is the function of a time-dependent spot interest rate that is consistent with the market price of instruments. So in terms of forward rates, we have:

**Corollary 1.** The forward rate and short rate are recursively given by:

\[ f(t, T) = -\left\langle K \exp(K(T-t)) X_t, 1 \right\rangle \left\langle \exp(K(T-t)) X_t, 1 \right\rangle \]

and
\[ r_t = -\left\langle (\theta + \frac{1}{2} \gamma^2), X_t \right\rangle. \]

**Proof.** The Bond price is given by:
\[ P(t, T) = \exp(-\int_t^T f(t, u)du), \]
so
\[ f(t, T) = -\frac{\partial \log(P(t, T))}{\partial T} = -\frac{\left\langle K \exp(K(T-t)) X_t, 1 \right\rangle}{\left\langle \exp(K(T-t)) X_t, 1 \right\rangle}. \]

Thus
\[ r_t = f(t, t) = -\langle K X_t, 1 \rangle \]
\[ = -\left\langle \left( A + \text{diag}(\theta + \frac{1}{2} \gamma^2) \right) X_t, 1 \right\rangle \]
\[ = -\left\langle \left( \text{diag}(\theta + \frac{1}{2} \gamma^2) \right) X_t, 1 \right\rangle \]
\[ = -\left\langle (\theta + \frac{1}{2} \gamma^2), X_t \right\rangle. \]

This follows because \( A \) is a rate matrix, see [20], so for all \( j, 1 \leq j \leq N \):
\[ \sum_{i=1}^{N} A_{ij} = 0. \]
2.2.2 Risk-neutral measure

Suppose \( r = (r_t, t \geq 0) \) is the interest rate given in (2.2.14). Consider the process:

\[
\beta_{0,t} = \exp \left( \int_0^t r_u du \right).
\]

Then \( d\beta_t = r_t \beta_t dt \).

**Notation 1.** In the following and for ease of notation, we write \( \pi_t := \pi_{0,t} \) and \( \beta_{u,t} := \frac{\beta_u}{\beta_t} \).

**Lemma 1.** \( \beta\pi \) is a \( \mathbb{P} \)-martingale.

**Proof.** We know that

\[
d\beta_{t,u} = r_u \beta_{t,u} du
\]

and

\[
d\pi_{t,u} = \pi_{t,u} \left[ \langle \gamma, X_u \rangle dW_u + \left( \langle \theta, X_u \rangle + \frac{1}{2} \langle \gamma^2, X_u \rangle \right) du \right]
\]

So

\[
d(\beta_u \pi_u) = \beta_u d\pi_u + \pi_u d\beta_u
\]

\[
= \beta_u \pi_u \left[ \langle \gamma, X_u \rangle dW_u + \left( \langle \theta, X_u \rangle + \frac{1}{2} \langle \gamma^2, X_u \rangle \right) du \right] + \beta_u \pi_u r_u du
\]

\[
= \beta_u \pi_u \left[ r_u + \langle \theta, X_u \rangle + \frac{1}{2} \langle \gamma^2, X_u \rangle \right] du + \beta_u \pi_u \langle \gamma, X_u \rangle dW_u.
\]

However,

\[
r_t = - \left\langle \left( \text{diag}(\theta + \frac{1}{2} \gamma^2) \right) X_t, 1 \right\rangle = - \left\langle (\theta + \frac{1}{2} \gamma^2), X_t \right\rangle.
\]

Therefore

\[
d(\beta_u \pi_u) = \beta_u \pi_u \langle \gamma, X_u \rangle dW_u,
\]

and \( \beta_t \pi_t \) is a \( \mathbb{P} \)-martingale.
Definition 1. Define $\mathbb{P}^*$, a new probability measure, by:

$$\frac{d\mathbb{P}^*}{d\mathbb{P}}|_{\mathcal{F}_T} = \frac{\beta_T \pi_T}{\mathbb{E}[\beta_T \pi_T]}.$$  \hspace{1cm} (2.2.17)

$\mathbb{P}^*$ is the risk neutral probability in our model. This is a consequence of the following result.

Lemma 2. For any asset $\mathcal{A}$, the discounted value $\mathcal{A}_t \beta_t^{-1}$ is a martingale under the risk-neutral measure $\mathbb{P}^*$ that is:

$$\forall s \leq t, \mathbb{E}^*[\frac{\mathcal{A}_t}{\beta_t} | \mathcal{F}_s] = \frac{\mathcal{A}_s}{\beta_s}.$$  

Proof. $\pi$ is a stochastic discount factor for any asset, including $\beta$, and for $0 \leq s \leq t$: $\pi_s \mathcal{A}_s = \mathbb{E}[\pi_t \mathcal{A}_t | \mathcal{F}_s]$. Therefore, by Bayes’ rule for conditional expectation, see Appendix A.3 or [20, 61] for further discussions, we have:

$$\mathbb{E}^*[\frac{\mathcal{A}_t}{\beta_t} | \mathcal{F}_s] = \frac{\mathbb{E}[\pi_t \beta_t \frac{\mathcal{A}_t}{\beta_t} | \mathcal{F}_s]}{\mathbb{E}[\pi_t \beta_t | \mathcal{F}_s]} = \frac{\pi_s \mathcal{A}_s}{\pi_s \beta_s} = \frac{\mathcal{A}_s}{\beta_s}.$$  

We used above the result from Lemma 1, that $\beta \pi$ is a $\mathbb{P}$-martingale. \hfill $\square$

Remark 1. The result from Lemma 2 justifies calling $\mathbb{P}^*$ a risk neutral measure.

Theorem 1. Under $\mathbb{P}^*$, $X$ has the same dynamics as under $\mathbb{P}$:

$$dX_t = AX_t dt + M^*_t.$$  

Here $M^*$ is a martingale under $\mathbb{P}^*$.

Proof. To prove the theorem we show that for $0 \leq t \leq T$, the dynamics under $\mathbb{P}^*$ of $X_T$ are the same as under $\mathbb{P}$.

Consider a new vector variable $Y_{t,T}$:

$$Y_{t,T} = \beta_{t,T} \pi_{t,T} X_T = \beta_{t,T} H_T.$$
We have previously calculated the dynamics of $H_{t,u} = \pi_{t,u}X_u$. As there are no quadratic variation terms between $H_{t,u}$ and $\beta_{T,U}$, and:

$$
\begin{cases}
    dH_{t,u} = (A + \text{diag}(\theta + \frac{1}{2}\gamma^2)) H_{t,u}du + \text{diag}(\gamma) H_{t,u}dW_u + \pi_{t,u}dM_u \\
    d\beta_{t,u} = r_u\beta_{t,u}du,
\end{cases}
$$

we have:

$$
dY_{t,u} = d(\beta_{t,u}H_{t,u})
= d\beta_{t,u}H_{t,u} + \beta_{t,u}dH_{t,u}
= r_u\beta_{t,u}H_{t,u}du + \left(A + \text{diag}(\theta + \frac{1}{2}\gamma^2)\right) \beta_{t,u}H_{t,u}du
+ \text{diag}(\gamma) \beta_{t,u}H_{t,u}dW_u + \pi_{t,u}\beta_{t,u}dM_u.
$$

The integral form of $Y_{t,T}$ is:

$$
Y_{t,T} = X_t + \int_t^T \left(r_u + A + \text{diag}(\theta + \frac{1}{2}\gamma^2)\right) Y_{t,u}du
+ \int_t^T \text{diag}(\gamma) Y_{t,u}dW_u + \int_t^T \pi_{t,u}\beta_{t,u}dM_u.
$$

On the other hand, from (2.2.15):

$$
\left(r_u + A + \text{diag}(\theta + \frac{1}{2}\gamma^2)\right) Y_{t,u} = \left(A + \text{diag}(\theta + \frac{1}{2}\gamma^2)\right) Y_{t,u} - \langle KX_u, 1 \rangle Y_{t,u} = AY_{t,u}.
$$

Therefore,

$$
Y_{t,T} = X_t + \int_t^T AY_{t,u}du + \int_t^T \text{diag}(\gamma) Y_{t,u}dW_u + \int_t^T \pi_{t,u}\beta_{t,u}dM_u.
$$

Conditioning each side on $\mathcal{F}_t$:

$$
\mathbb{E}[Y_{t,T}|\mathcal{F}_t] = X_t + \int_t^T A\mathbb{E}[Y_{t,u}|\mathcal{F}_t]du.
$$

Since from Bayes’ Theorem (see Appendix A.3):

$$
\mathbb{E}^*[X_T|\mathcal{F}_t] = \frac{\mathbb{E}[\beta_{T|T}X_T|\mathcal{F}_t]}{\mathbb{E}[\beta_{T|T}|\mathcal{F}_t]} = \mathbb{E}[\beta_{t,T|T}X_T|\mathcal{F}_t] = \mathbb{E}[Y_{t,T}|\mathcal{F}_t],
$$
Therefore,

\[ E^*[X_T|\mathcal{F}_t] = X_t + \int_t^T A E^*[X_u|\mathcal{F}_t] du. \]

Suppose \( \Phi(t, T) \) is the solution of the Linear system of ordinary differential equations:

\[
\frac{d\Phi(t, T)}{dT} = A\Phi(t, T) \quad \Phi(t, t) = I_N.
\]

Then

\[ E^*[X_T|\mathcal{F}_t] = \Phi(t, T)X_t. \]

The result follows from the semi-martingale decomposition of the Markov Chain \( X \), see [20].

Lemma 3. Let

\[ W^*_t = W_t - \int_0^t (\gamma, X_u) du. \]

Then \( W^* \) is a martingale under \( \mathbb{P}^* \).

Proof. From the change of probability given in (2.2.17), and by applying Bayes’ rule for conditional expectation, we have for \( 0 \leq t \leq u \):

\[
E^*[W_u|\mathcal{F}_t] = \frac{\mathbb{E}[\beta_{T,T}^T W_u|\mathcal{F}_t]}{\mathbb{E}[\beta_{T,T}^T|\mathcal{F}_t]} = \frac{\mathbb{E}[\beta_{T,T}^T W_u|\mathcal{F}_t]}{\beta_{t,t}} = \mathbb{E}[\beta_{t,T,T}^T W_u|\mathcal{F}_t] = \mathbb{E}[\mathbb{E}[\beta_{t,T,T}^T W_u|\mathcal{F}_u]|\mathcal{F}_t] = \mathbb{E}[\beta_{t,u}^T W_u|\mathcal{F}_t].
\]

Now

\[
d(\beta_{t,u}^T W_u) = W_u d(\beta_{t,u}^T W_u) + \beta_{t,u}^T W_u dW_u + d\langle W, \beta_{t,u}^T \rangle u.
\]

So from (2.2.16)

\[
d(\beta_u^T) = \beta_u^T (\gamma, X_u) dW_u,
\]
Therefore
\[ d\langle W, \beta_t, \pi_t, \rangle_u = \beta_{t,u} \pi_{t,u} \langle \gamma, X_u \rangle du , \]
and we have
\[ d(\beta_{t,u} \pi_{t,u} W_u) = W_u \beta_{t,u} \pi_{t,u} \langle \gamma, X_u \rangle dW_u + \beta_{t,u} \pi_{t,u} dW_u + \beta_{t,u} \pi_{t,u} \langle \gamma, X_u \rangle du . \]

Therefore
\[ \beta_{t,T} \pi_{t,T} W_T = W_t + \int_t^T W_u \beta_{t,u} \pi_{t,u} \langle \gamma, X_u \rangle dW_u + \int_t^T \beta_{t,u} \pi_{t,u} dW_u + \int_t^T \beta_{t,u} \pi_{t,u} \langle \gamma, X_u \rangle du . \]
Conditioning both sides on \( F_t \):
\[ \mathbb{E}[\beta_{t,T} \pi_{t,T} W_T | F_t] = W_t + \int_t^T \mathbb{E}[\beta_{t,u} \pi_{t,u} \langle \gamma, X_u \rangle | F_t] du . \]
That is
\[ \mathbb{E}^*[W_T | F_t] = W_t + \int_t^T \mathbb{E}^*[\langle \gamma, X_u \rangle | F_t] du . \]

Subtracting \( \mathbb{E}^*[\int_0^T \langle \gamma, X_u \rangle du | F_t] \) from each side of the equality we obtain:
\[ \mathbb{E}^*[W_T - \int_0^T \langle \gamma, X_u \rangle du | F_t] = W_t - \int_0^t \langle \gamma, X_u \rangle du . \]

**Corollary 2.** \( W_t^* = W_t - \int_0^t \langle \gamma, X_u \rangle du \), is a Brownian motion under \( \mathbb{P}^* \).

**Proof.** We have proved that \( W_t^* \) is a martingale. We shall now prove that it is a Brownian motion.

Applying Itô’s lemma, see [56]:
\[ d((W_t^*)^2) = 2(W_t - \int_0^t \langle \gamma, X_u \rangle du) dW_t + dt \]
So
\[ (W_t^*)^2 - t = 2 \int_0^t W_u^* dW_u \] and is a \( \mathbb{P} \)-martingale.

By Lévy’s characterisation theorem (see Appendix A.4) we conclude that \( W_t^* \) is a \( \mathbb{P}^* \) Brownian motion. \( \square \)
**Theorem 2.** Let

\[ \omega_t = (\beta_t \pi_t)^{-1}. \]

Then \( \omega \) is a martingale under \( \mathbb{P}^* \) and

\[ \omega_t = 1 + \int_0^t \omega_u \langle -\gamma, X_u \rangle dW_u^* \]

where \( W^* \) is the \( \mathbb{P}^* \)-Brownian motion of Corollary 2.

**Proof.** We have:

\[ \beta_t^{-1} = \exp\left(-\int_0^t r_u du\right). \]

So

\[ d(\beta_t^{-1}) = -r_t \beta_t^{-1} dt. \]

Also

\[ \pi_t^{-1} = \exp\left(\int_t^T \langle -\gamma, X_u \rangle dW_u + \int_t^T \langle -\theta, X_u \rangle du\right). \]

The dynamics of \( \pi^{-1} \) are similar to the dynamics of \( \pi \) after replacing \( \gamma \) by \(-\gamma\) and \( \theta \) by \(-\theta\).

Therefore we can apply Itô’s lemma or write:

\[ d\pi_u^{-1} = \pi_t^{-1}\left[ \langle -\gamma, X_u \rangle dW_u + \langle -\theta, X_u \rangle + \frac{1}{2} \langle \gamma^2, X_u \rangle du \right]. \]

Because there are no quadratic variation terms between \( \pi \) and \( \beta \):

\[ d(\beta_u \pi_u)^{-1} = \beta_u^{-1} d\pi_u^{-1} + \pi_u^{-1} d\beta_u^{-1} \]

\[ = (\beta_t \pi_t)^{-1} \left[ \langle -\gamma, X_t \rangle dW_t + \langle -\theta, X_t \rangle + \frac{1}{2} \langle \gamma^2, X_t \rangle du \right] - r_t (\beta_t \pi_t)^{-1} dt \]

\[ = (\beta_t \pi_t)^{-1} \langle -\gamma, X_t \rangle dW_t + (\beta_t \pi_t)^{-1} \left[ \langle -\theta, X_t \rangle + \frac{1}{2} \langle \gamma^2, X_t \rangle - r_t \right] dt. \]

On the other hand, from (2.2.15):

\[ r_t = -\left( \langle \theta + \frac{1}{2} \gamma^2, X_t \rangle \right). \]
Therefore, the dynamics of \((\pi \beta)^{-1}\) are:

\[
d(\beta_u \pi_u)^{-1} = (\beta_t \pi_t)^{-1}(\gamma, X_t) dW_t + (\beta_t \pi_t)^{-1}(\gamma^2, X_t) dt
\]

\[
= (\beta_t \pi_t)^{-1}(\gamma, X_t)[dW_t + (\gamma, X_t) dt].
\]

The result follows from Corollary 2.

\[\square\]

### 2.3 Deriving the stock price

In this Chapter we shall also assume the state of the economy influences the dividends paid on the asset. This assumption agrees with economic facts. When the economy is in a prosperous state, companies cut dividend payments for shareholders in order to increase investment spending, and vice-versa.

#### 2.3.1 Dividends payment model

Let \(\phi = (\phi_1, \phi_2, \ldots, \phi_N)' \in \mathbb{R}^N\) and \(\tau = (\tau_1, \tau_2, \ldots, \tau_N)' \in \mathbb{R}^N\). We suppose the dividend payments \(\delta\) on asset in our model follows the regime switching geometric Brownian motion with dynamics of the form:

\[
\frac{d\delta_u}{\delta_u} = (\tau, X_u) du + (\phi, X_u) dW_u.
\]

This implies:

\[
\delta_{t,T} = \delta_t \exp\left(\int_t^T (\phi, X_u) dW_u + \int_t^T \left(\tau - \frac{1}{2}\phi^2, X_u\right) du\right).
\]

Note that the stochastic discount function and the dividend payments are modelled using the same Brownian motion \(W\).

The stock price is the present value of all future dividend payments. That is, under the real world probability,

\[
S_t = \mathbb{E}\left[\int_t^\infty \pi_{t,u} \delta_{t,u} du | \mathcal{F}_t\right].
\]

**Remark 2.** The above expectation is taken under real word probability measure since we are using the stochastic discount factor.
2.3.2 Deriving the stock price

Lemma 4. Define the matrix $B$:

$$B = A + \text{diag}(L).$$

where $L = (\tau + \theta + \frac{1}{2}\gamma^2 + \phi\gamma) \in \mathbb{R}^N$. Then the stock price can be written:

$$S_t = \langle \delta_t \int_0^\infty \exp\{uB\}du X_t, 1 \rangle. \quad (2.3.3)$$

Proof. Let $D_{t,u} = \pi_{t,u}\delta_{t,u}$ be the dividend payment made at time $u$ discounted to $t$, so:

$$S_t = \int_t^\infty \mathbb{E}[D_{t,u} | \mathcal{F}_t] du. \quad (2.3.4)$$

Then:

$$dD_{t,u} = \pi_{t,u}d\delta_{t,u} + \delta_{t,u}d\pi_{t,u} + d\langle \pi_{t,\cdot}, \delta_{t,\cdot} \rangle_u$$

$$= D_{t,u} (\langle \phi, X_u \rangle dW_u + \langle \tau, X_u \rangle du) + D_{t,u} \langle \phi, X_u \rangle \langle \gamma, X_u \rangle du$$

$$+ D_{t,u} \left( \langle \gamma, X_u \rangle dW_u + \left[ \langle \theta, X_u \rangle + \frac{1}{2} \gamma^2, X_u \right] du \right).$$

So

$$dD_{t,u} = D_{t,u} \left[ \langle \phi + \gamma, X_u \rangle dW_u + \langle \tau + \theta + \frac{1}{2} \gamma^2 + \phi\gamma, X_u \rangle du \right].$$

That is:

$$\frac{dD_{t,u}}{D_{t,u}} = \left( \phi + \gamma \right)dW_u + \left( \tau + \theta + \frac{1}{2} \gamma^2 + \phi\gamma \right)du, X_u \right). \quad (2.3.5)$$

Consider the vector variable $Z$ defined by:

$$Z_{t,T} = D_{t,T}X_T.$$

Then, by Itô’s lemma for jump processes, see [18]:

$$d(Z_{t,u}) = dD_{t,u}X_u + D_{t,u} - dX_u + (\Delta X_u)(\Delta D_u).$$

Here $\Delta X_u$ and $\Delta D_u$ are the jumps at time $u$ of the processes $X$ and $D$, respectively. However, $D$ is a continuous process, so $\Delta D_{t,u} = 0$, and
\[
Z_{t,T} = \delta_t X_t + \int_t^T X_u - D_{t,u} - D_{t,u} - dX_u \\
= \delta_t X_t + \int_t^T X_u - D_{t,u} \left[ \langle \phi + \gamma, X_u \rangle dW_u + \langle \tau + \theta + \frac{1}{2} \gamma^2 + \phi \gamma, X_u \rangle du \right] \\
+ \int_t^T D_{t,u} - dX_u.
\]

Using the dynamics of \( X \) from (2.2.1):

\[
Z_{t,T} = \delta_t X_t + \int_t^T Z_{t,u} - \langle \phi + \gamma, X_u \rangle dW_u + \int_t^T Z_{t,u} - \langle \phi + \gamma, X_u \rangle dW_u \\
+ \int_t^T D_{t,u} - A_u X_u du + \int_t^T D_{t,u} - dM_u.
\]

Then:

\[
Z_{t,T} = \delta_t X_t + \int_t^T (\langle L, X_u \rangle + A) Z_{t,u} du + \int_t^T D_{t,u} - dM_u + \int_t^T Z_{t,u} - \langle \phi + \gamma, X_u \rangle dW_u.
\]

Now, \( \langle L, X_u \rangle Z_{t,u} = \text{diag}(L) Z_{t,u} \), and the last two integrals are martingales, so:

\[
\mathbb{E}[Z_{t,T} | \mathcal{F}_t] = \delta_t X_t + \int_t^T (\text{diag} + A) \mathbb{E}[Z_{t,u} | \mathcal{F}_t] du.
\]

Therefore,

\[
\mathbb{E}[Z_{t,T} | \mathcal{F}_t] = \delta_t X_t + \int_t^T B \mathbb{E}[Z_{t,u} | \mathcal{F}_t] du. \tag{2.3.6}
\]

Let \( \mathcal{D}(t, T) \) be the matrix solution of the linear system of ordinary differential equations:

\[
\frac{d\mathcal{D}(t, T)}{dt} = B \mathcal{D}(t, T) \]

\[
\mathbb{E}[Z_{t,T} | \mathcal{F}_t] = \delta_t.
\]

Here, since the matrix \( B \) is constant, \( \mathcal{D}(t, T) = \exp(B(T - t)) \in \mathbb{R}^{N \times N} \). Therefore, the solution of the differential equation (2.3.6) is:

\[
\mathbb{E}[Z_{t,T} | \mathcal{F}_t] = \delta_t \mathcal{D}_{t,T} X_t.
\]
So,
\[ \mathbb{E}[Z_{t,T}|\mathcal{F}_t] = \delta_{t,t} \exp(B(T - t)) X_t. \]

Using the above expression of \( \mathbb{E}[Z_{t,T}|\mathcal{F}_t] \) we can obtain \( \mathbb{E}[D_{t,T}|\mathcal{F}_t] \) as:
\[
\mathbb{E}[D_{t,T}|\mathcal{F}_t] = \mathbb{E}[\langle D_{t,T}X_T, 1 \rangle|\mathcal{F}_t] \\
= \mathbb{E}[\langle Z_{t,T}, 1 \rangle|\mathcal{F}_t] \\
= \langle \mathbb{E}[Z_{t,T}|\mathcal{F}_t], 1 \rangle \\
= \langle \delta_{t,t} \exp\{B(T - t)\}X_t, 1 \rangle.
\]

Consequently,
\[
S_t = \int_t^\infty \mathbb{E}[\pi_{t,u}\delta_{t,u}|\mathcal{F}_t]du \\
= \int_t^\infty \mathbb{E}[D_{t,u}|\mathcal{F}_t]du \\
= \int_t^\infty \langle \mathbb{E}[Z_{t,u}|\mathcal{F}_t], 1 \rangle du \\
= \int_t^\infty \langle \delta_{t,u} \exp\{B(u - t)\}X_t, 1 \rangle du \\
= \langle \delta_t \int_t^\infty \exp\{B(u - t)\}X_t, 1 \rangle \\
= \langle \delta_t \int_0^\infty \exp\{uB\}X_t, 1 \rangle.
\]

2.3.3 Explicit stock price

We now introduce conditions which ensure the stock price process (2.3.3) is well defined, that is, so that the integral converges. Recall \( B = A + \text{diag}(L) \) and \( L = \tau + \theta + \frac{1}{2} \gamma^2 + \phi \gamma. \)

**Assumption 1.** With \( \theta \) the vector appearing in the first integral of the stochastic discount function, we assume that the components of \( \theta \) are large and negative and such that:
\[
\sum_{k=1}^{N} \theta_k \leq - \left( \sum_{k=1}^{N} (\tau_k + \frac{1}{2}\gamma_k^2 + \phi_k\gamma_k) + \max_{i,j} |a_{ji}| \right).
\] (2.3.7)

The parameter \( \theta \) represents the drift term in the stochastic discount factor, therefore it is non positive.

**Lemma 5.** Suppose \( \theta \) is such that Assumption 1 holds. Then, for each \( j \in \{1, 2, \ldots, r\} \) : \( \Re(\lambda_j) < 0 \).

**Proof.** Let \( \lambda_j \) be the \( j \)th eigenvalue of the matrix \( B \), and \( \alpha_j = (\alpha_j^1, \ldots, \alpha_j^N) \) the associated eigenvector to \( \lambda_j \). Then:

\[ B\alpha_j = \lambda_j \alpha_j. \]

Without loss of generality, we can assume:

\[ |\alpha_j|^2 = \alpha_j'\alpha_j = 1. \]

From the definition of \( B \):

\[ B\alpha_j = A\alpha_j + \text{diag}(L)\alpha_j. \]

If \( \bar{\alpha}_j \) is the conjugate vector of \( \alpha_j \), then:

\[ \bar{\alpha}_j'B\alpha_j = \bar{\alpha}_j'A\alpha_j + \sum_{k=1}^{N} L_k|\alpha_j^k|^2. \]

So with \( \Re(.) \) denoting the real part, then for \( j \in \{1, 2, \ldots, r\} \)

\[ \Re(\lambda_j) = \Re(\bar{\alpha}_j'A\alpha_j) + \sum_{k=1}^{N} L_k|\alpha_j^k|^2 \]

and as \( L \) is the diagonal matrix with \( k \)th entry \( \theta_k + \tau_k + \gamma_k\phi_k + \frac{1}{2}\gamma_k^2 \),

\[ \left( \Re(\bar{\alpha}_j'A\alpha_j) + \sum_{k=1}^{N} \theta_k + \sum_{k=1}^{N} (\tau + \gamma\phi + \frac{1}{2}\phi^2) \right). \] (2.3.8)
On the other hand,

\[ \Re(\bar{\alpha}_j' \alpha_j) \leq |\bar{\alpha}_j' \alpha_j| \leq |\alpha_j|^2 \cdot \max_{i,j} |a_{i,j}| = \max_{i,j} |a_{ji}|. \]

The result follows from Assumption 1. \(\square\)

**Remark 3.** We can write the eigenvalues of the matrix \(B\) as:

\[ \forall 0 \leq j \leq r : \lambda_i = -\rho_i + i\sigma_i. \]

where \(\rho_i > 0\).

**Theorem 3.** Recall \(B = A + \text{diag}(L)\) and \(L = \theta + \tau + \gamma \phi + \frac{1}{2} \gamma^2\). Then with a parameter vector \(\theta\) satisfying the conditions of Lemma 5, the stock price at time \(t\) can be expressed in the form:

\[ S_t = \delta_t \langle VX_t, 1 \rangle \]  \hspace{1cm} (2.3.9)

where \(V\) is a constant matrix.

**Proof.** Using the Jordan normal decomposition, the matrix \(B\) can be written (see [59]) as:

\[ B = PJP^{-1}. \]  \hspace{1cm} (2.3.10)

Here \(P\) is an invertible matrix and \(J\) is a block diagonal matrix:

\[ J = \begin{pmatrix} J_1 & & & \\ & \ddots & & \\ & & 0 & \\ 0 & & \ddots & J_r \end{pmatrix}. \]
Recall $r$ is the number of eigenvalues of $B$, and the block terms are of the form:

$$J_i = \begin{pmatrix} \lambda_i & 1 \\ \lambda_i & \ddots & \ddots \\ \vdots & \ddots & 1 \\ \lambda_i \\ \end{pmatrix}_{m_i \times m_i}.$$

$J_i$ corresponds to the generalized eigenspace generated by $\lambda_i$.

From this decomposition:

$$\exp(uB) = \sum_{j \geq 0} \frac{(uB)^j}{j!}$$

$$= P \begin{pmatrix} \exp(uJ_1) & 0 & \cdots & 0 \\ 0 & \exp(uJ_1) & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \exp(uJ_r) \\ \end{pmatrix} P^{-1}. $$

For the $i^{th}$ block $J_i$ in the previous matrix, note:

$$J_i = \lambda_i I_{m_i} + N_i,$$

where $I$ is an $m_i \times m_i$ identity matrix and $N_i$ is a nilpotent matrix of order $m_i$.

Then:

$$\exp(uJ_i) = \exp(\lambda_i u) \exp^{uN_i}.$$ 

Using this decomposition, the integral in the stock price becomes:

$$\int_0^\infty \exp(uB)du = \int_0^\infty P \begin{pmatrix} e^{\lambda_1 u} \exp(uN_1) & 0 & \cdots & 0 \\ 0 & e^{\lambda_1 u} \exp(uN_1) & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & e^{\lambda_1 u} \exp(uN_r) \\ \end{pmatrix} P^{-1} du.$$ 

Now the integral of the $i^{th}$ diagonal term is:

$$\int_0^\infty e^{\lambda_i u} \exp(uN_i)du = \int_0^\infty \exp(\lambda_i u) \sum_{k=0}^{m_i} \frac{u^k}{k!} N_i^k du$$

$$= \sum_{k=0}^{m_i} \int_0^\infty \exp(\lambda_i u) \frac{u^k}{k!} N_i^k du.$$
On the other hand, using Remark 3, note that:

\[ \left| \int_0^\infty e^{-(\rho_i + \sigma_i)u} u^k du \right| \leq \int_0^\infty |e^{-\rho_i u} u^k| |e^{iu}| du < \infty. \]

Therefore,

\[ \int_0^\infty \exp(\lambda_i u) \frac{u^k}{k!} du = \left( \frac{-1}{\lambda_i} \right)^k \int_0^\infty \exp(\lambda_i u) du = \left( \frac{-1}{\lambda_i} \right)^{k+1}. \]

and

\[ \int_0^\infty e^{uB} du = P \begin{pmatrix} \Gamma_1 & 0 \\ \vdots & \vdots \\ 0 & \Gamma_r \end{pmatrix} P^{-1}. \quad (2.3.11) \]

Here

\[ \Gamma_i = \sum_{k=0}^{m_i} \left( \frac{-1}{\lambda_i} \right)^{k+1} N^k_i. \quad (2.3.12) \]

Writing

\[ V = P \begin{pmatrix} \Gamma_1 & 0 \\ \vdots & \vdots \\ 0 & \Gamma_r \end{pmatrix} P^{-1}, \]

the stock price at time \( t \) is:

\[ S_t = \delta_t (V X_t, 1) = \delta_t 1' V X_t. \quad (2.3.13) \]

2.3.4 The martingale condition

From (2.2.5), the dynamics of \( \pi \) are:

\[ d\pi_{t,u} = \pi_{t,u} \left[ (\gamma, X_u) dW_u + (\theta + \frac{1}{2} \gamma^2, X_u) du \right]. \]

Therefore, the SDF is a martingale if and only if:

\[ \theta = -\frac{1}{2} \gamma^2. \]
That is for each $n, 1 \leq n \leq N : \theta_n = \frac{-1}{2} \lambda_n^2$.

Under the martingale condition, the vector $L$ becomes:

$$L = \tau + \phi \gamma.$$ 

This means that the stock price depends on the stochastic discount function only through its randomness generated by the Markov chain and the Brownian motion, and not through its deterministic component.

### 2.3.5 The stock price dynamics

We showed in Theorem 3 that the matrix $B$ can be diagonalised as $B = P \text{diag}(\lambda) P^{-1}$ and the stock price can be written as: $S_t = \mathbf{1}'MX_t$, where the matrix $M$ is defined as:

$$M = \delta_t P \begin{pmatrix} \Gamma_1 & 0 \\ \vdots & \ddots \\ 0 & \Gamma_r \end{pmatrix} P^{-1}.$$ 

where $\Gamma_i, i \in 1, \ldots, r$ are square matrices defined in (2.3.12). Therefore, the stock price can also be expressed as:

$$S_t = \delta_t U'X_t = \delta_t \zeta_t.$$ 

where $\zeta_t = U'X_t$.

Another way to express the stock price is to write

$$S_t = \delta_t U'X_t = \langle \delta_t U, X_t \rangle.$$ 

So

$$S_t = \langle Q_t, X_t \rangle,$$ 

where

$$Q_t = M' \mathbf{1}.$$
We, therefore, have an $\mathcal{F}_t$-measurable representation for the stock price, $S_t$. Also, recall from (2.2.1) and (2.3.1) using the fact that $U$ is the constant matrix defined in (2.3.14):

$$
\begin{cases}
    d\delta_t = \delta_t [\langle \phi, X_t \rangle dW_t + \langle \tau, X_t \rangle dt] \\
    d\zeta_t = U' [AX_t dt + dM_t].
\end{cases}
$$

Therefore, the dynamics of $S$ are given by:

$$dS_t = d(\delta_t \zeta_t) = \zeta_t d(\delta_t) + \delta_t d(\zeta_t) = S_t [\langle \phi, X_t \rangle dW_t + \langle \tau, X_t \rangle dt] + \delta_t U' dM_t + \delta_t U' AX_t dt = S_t \langle \tau, X_t \rangle dt + \delta_t U' AX_t dt + S_t \langle \phi, X_t \rangle dW_t + \delta_t U' dM_t.$$

### 2.4 Pricing European stock options

In the previous section, some expressions and formulae for the stock price were derived. We shall now use our model to derive the price of a European call option, and establish a hedging strategy.

#### 2.4.1 The pricing of a call option

Suppose, as above, that $S_t = \langle Q_t, X_t \rangle$ is the stock price at time $t$, and $G(S_T)$ is a European type claim with expiry $T$. Then the value of the claim at time $T$ is $C_T = G(S_T)$. We wish to find the price of the claim at time $t$, $C_t$.

For convenience we prefer to write the stock price in (2.3.15) as:

$$S_t = \delta_t \langle V, X_t \rangle.$$

Here the vector $V = (v_1, v_2, \ldots, v_N)'$ refers to the constant vector

$$V = (P^{-1})' \text{diag} \left( \frac{-1}{\lambda_1}, \frac{-1}{\lambda_2}, \ldots, \frac{-1}{\lambda_N} \right) P' 1, \quad (2.4.1)$$

depending on the eigenvalues of the matrix $B$ defined in Lemma 4.
Remark 4. The invertible matrix $P$ obtained in the matrix $B$ Jordan normal form (2.3.10) is not necessarily orthonormal, unless the matrix $B$ satisfies $\dim(\ker(B)) = 0$, so $P^{-1} \neq P'$.

Now the stock price, and consequently the European claim, at time $t$ is a function of the Markov chain $X_t$ and the level of dividend payment. The claim value at time $T$ is then:

$$C_T = G(S_T)$$
$$= G(\delta_T(V, X_T))$$
$$= \langle G(\delta_T), X_T \rangle.$$  

Here

$$G(\delta_T) = (G(\delta_T v_1), G(\delta_T v_2), \ldots, G(\delta_T v_N))' = (G_1(\delta_T), G_2(\delta_T), \ldots, G_N(\delta_T))',$$

(2.4.2)

where

$$G(\delta_T v_i) = G_i(\delta_T) = G(\delta_T \langle V, e_i \rangle).$$  

(2.4.3)

Using the stochastic discount factor, we have:

$$C_t = \mathbb{E}[\pi_{t,T}C_T | F_t]$$
$$= \mathbb{E}[\pi_{t,T} \langle G(\delta_T), X_T \rangle | F_t]$$
$$= \mathbb{E}[\mathbb{E}[\pi_{t,T} \langle G(\delta_T), X_T \rangle | F_t \lor F^X_{t,T}] | F_t].$$\hspace{1cm} (2.4.4)

Here $F_t \lor F^X_{t,T} = \sigma\{W_u, 0 \leq u \leq t; X_u, t \leq u \leq T\}$, and the filtration $F^X_{t,T}$ represents the information generated by the Markov chain $X$ between $t$ and $T$.

Writing $\theta_u = \langle \theta, X_u \rangle$ and $\gamma_u = \langle \gamma, X_u \rangle$ we have:

$$\pi_{t,T} = \exp \left( \int_t^T \theta_u du + \int_t^T \gamma_u dW_u \right)$$
$$= \exp \left( \int_t^T \theta_u du + \frac{1}{2} \int_t^T \gamma_u^2 du \right) \exp \left( \int_t^T \gamma_u dW_u - \frac{1}{2} \int_t^T \gamma_u^2 du \right).$$
CHAPTER 2. THE STOCHASTIC DISCOUNT FACTOR

Now $\theta_u$ and $\gamma_u$ are $F_t \vee F_{t,T}^X$-measurable, so:

$$
E[\pi_{t,T}(G(\delta_T), X_T)|F_t \vee F_{t,T}^X] = \exp \left( \int_t^T \theta_u du + \frac{1}{2} \int_t^T \gamma_u^2 du \right) \times E \left[ \exp \left( \int_t^T \gamma_u dW_u - \frac{1}{2} \int_t^T \gamma_u^2 du \right) \langle G(\delta_T), X_T \rangle |F_t \vee F_{t,T}^X F_t \right].
$$

Define a new probability measure $P^\gamma$ by:

$$
\frac{dP^\gamma}{dP}|_{F_t \vee F_{t,T}^X} := \exp \left( \int_t^T \gamma_u dW_u - \frac{1}{2} \int_t^T \gamma_u^2 du \right).
$$

From Girsanov’s Theorem, under $P^\gamma$, $W_t^\gamma := W_t - \int_0^t \gamma_u du$ is a Brownian Motion and

$$
E \left[ \exp \left( \int_t^T \gamma_u dW_u - \frac{1}{2} \int_t^T \gamma_u^2 du \right) \langle G(\delta_T), X_T \rangle |F_t \vee F_{t,T}^X \right] = \langle E^\gamma \left[ G(\delta_T), X_T \right] |F_t \vee F_{t,T}^X, X_T \rangle.
$$

On the other hand, for the $i^{th}$ element of $G(\delta_T)$, we have:

$$
E^\gamma \left[ G_i(\delta_T) |F_t \vee F_{t,T}^X \right] = E^\gamma \left[ G \left( v_i \delta_t \exp \left( \int_t^T \phi_u dW_u^\gamma \exp \left( \int_t^T (\tau_u - \frac{1}{2} \phi_u^2 + \phi_u \gamma_u) du \right) \right) |F_t \vee F_{t,T}^X \right].
$$

Notation 2. For convenience of notation, write:

$$
\begin{align*}
  d & = \delta_t \\
  \Delta_i & = \tau_i - \frac{1}{2} \phi_i^2 + \phi_i \gamma_i \\
  \Delta & = (\Delta_1, \ldots, \Delta_N) \\
  \epsilon & = \exp(\langle \Delta, J_{t,T} \rangle),
\end{align*}
$$

Using the above notation, (2.4.6) can be expressed as:

$$
E^\gamma \left[ G_i(\delta_T) |F_t \vee F_{t,T}^X \right] = E^\gamma \left[ G \left( v_i d \epsilon \exp(\int_t^T \phi_u dW_u^\gamma) \right) |F_t \vee F_{t,T}^X \right].
$$
2.4.2 Occupation times of a Markov process

We consider the vector of occupation times, that is the vector with entries representing the occupation of the Markov process $X$ in different states.

**Definition 2.** Define for $1 \leq i \leq n$:

$$J^i_{t,T} = \int_t^T \langle e_i, X_u \rangle du.$$  

So

$$J_{t,T} = \langle J^1_{t,T}, J^2_{t,T}, \ldots, J^N_{t,T} \rangle$$  

is the vector of the occupation times of the Markov chain of different states.

Under $\mathbb{P}^\gamma$ we know that $\int_t^T \phi_u dW_u^\gamma$ has normal distribution $\mathcal{N}(0, \sigma^2)$, where:

$$\sigma^2 = \int_t^T \phi_u^2 du = \langle \phi^2, J_{t,T} \rangle.$$

Therefore, from (2.4.4), the claim price at time $t$ is:

$$C_t = \mathbb{E} \left[ \exp \left( \int_t^T \theta_u du + \frac{1}{2} \int_t^T \gamma_u^2 du \right) \mathbb{E}^\gamma \left[ G \left( v_i d\epsilon \exp(\int_t^T \phi_u dW_u^\gamma) \right) \bigg| \mathcal{F}_t \vee \mathcal{F}^X_{t,T} \right] \bigg| \mathcal{F}_t \right],$$

where $v_i$ is the $i^{th}$ element of the constant vector defined in (2.4.1).

For a European call claim with a strike $K$, $G(x) = (x - K)^+$, so $G_i(x) = (xv_i - K)^+$. Therefore,

**Proposition 2.** Using Notation 2, and writing:

$$\eta_1^i = \frac{1}{\sigma} \log \left( \frac{K}{v_i d\epsilon} \right), \quad \eta_2^i = \eta_1^i - \sigma,$$

we have the following equality:

$$\mathbb{E}^\gamma \left[ G_i(\delta_T) \big| \mathcal{F}_t \vee \mathcal{F}^X_{t,T} \right] = e^{\frac{\sigma^2}{2} v_i d\epsilon} \mathcal{N}(-d_1^i) - K \mathcal{N}(-d_2^i). \quad (2.4.7)$$

**Remark 5.** Note in the above expression the similarity to Black and Scholes call price expression.
Proof.

\[
\mathbb{E}^\gamma \left[ G_i(\delta_T) | \mathcal{F}_t \vee \mathcal{F}_{t,T}^X \right] = \mathbb{E}^\gamma \left[ \left( v_i \exp(\int_t^T \phi_u dW_u^\gamma) - K \right)^+ | \mathcal{F}_t \vee \mathcal{F}_{t,T}^X \right]
\]

\[
= v_i \mathbb{E}^\gamma \left[ \left( \exp\left( \int_t^T \phi_u dW_u^\gamma \right) - \frac{K}{v_i \epsilon} \right)^+ | \mathcal{F}_t \vee \mathcal{F}_{t,T}^X \right]
\]

where \( Z \sim \mathcal{N}(0,1) \)

\[
= v_i \mathbb{E}^\gamma \left[ \left( \exp(\sigma Z) - \frac{K}{v_i \epsilon} \right)^+ | \mathcal{F}_t \vee \mathcal{F}_{t,T}^X \right]
\]

\[
\frac{v_i \epsilon}{\sqrt{2\pi}} \int_{\eta_2}^{+\infty} \left( e^{\sigma z} - \frac{K}{v_i \epsilon} \right)^+ e^{-\frac{z^2}{2}} dz
\]

\[
\frac{v_i \epsilon}{\sqrt{2\pi}} \int_{\eta_2}^{+\infty} \left( e^{\sigma z} - \frac{K}{v_i \epsilon} \right) e^{-\frac{z^2}{2}} dz
\]

\[
= v_i \epsilon \frac{e^{\sigma^2}}{\sqrt{2\pi}} \int_{\eta_2}^{+\infty} e^{-\frac{1}{2}(z-\sigma)^2 + \frac{\sigma^2}{2}} dz - \mathcal{N}(-\eta_2^i)
\]

\[
= e^{\sigma^2} v_i \epsilon \mathcal{N}(-\eta_1^i + \sigma) - \mathcal{N}(-\eta_2^i)
\]

\[
= e^{\sigma^2} v_i \epsilon \mathcal{N}(-\eta_1^i) - \mathcal{N}(-\eta_2^i).
\]

\[
\square
\]

In this expression, the only remaining randomness comes from the Markov chain, which appears in \( \sigma \) and \( \epsilon \). Recall that:

\[
\sigma^2 = \int_t^T \phi_u^2 du = \langle \phi^2, J_{t,T} \rangle \quad \text{and} \quad \epsilon = \exp(\langle \Delta, J_{t,T} \rangle).
\]

The expectation can be written as:

\[
\mathbb{E}^\gamma \left[ G_i(\delta_T) | \mathcal{F}_t \vee \mathcal{F}_{t,T}^X \right] = H_i(J_{t,T}^1, J_{t,T}^2, \ldots, J_{t,T}^N).
\]

Here

\[
H_i(x_1, x_2, \ldots, x_N) = e^{\langle \phi^2 + \Delta, x \rangle} v_i \epsilon \mathcal{N} \left( \sqrt{\langle \phi^2, x \rangle} - \frac{1}{\sqrt{\langle \phi^2, x \rangle}} \log \left( \frac{K}{v_i \epsilon} \right) \right) - \mathcal{N} \left( \frac{1}{\sqrt{\langle \phi^2, x \rangle}} \log \left( \frac{K}{v_i \epsilon} \right) \right).
\]
Remark 6. The index $i$ appears only in the constant terms $v_i$.

Write $H(x_1, x_2, \ldots, x_N)$ for the vector $(H_1, \ldots, H_N)$ and:

$$\exp(\int_t^T \theta_a du + \frac{1}{2} \int_t^T \gamma_a^2 du) = \exp(\langle \theta + \frac{1}{2} \gamma^2, J_{t,T} \rangle).$$

Then, the call price at time $t$ is:

$$C_t = \mathbb{E}[\exp(\langle \theta + \frac{1}{2} \gamma^2, J_{t,T} \rangle) (H(J_{t,T}^1, J_{t,T}^2, \ldots, J_{t,T}^N), X_T) | \mathcal{F}_t]$$

with

$$H^*(x_1, x_2, \ldots, x_N) = \exp(\langle \theta + \frac{1}{2} \gamma^2, x \rangle) H(x_1, x_2, \ldots, x_N).$$

Therefore, for a vector of occupational times $J_{t,T} = (J_{t,T}^1, J_{t,T}^2, \ldots, J_{t,T}^N)'$ the call price is:

$$C_t = \sum_{i=1}^N \mathbb{E}[\langle H^*(J_{t,T}), e_i \rangle | \mathcal{F} \lor X_T = e_i] \mathbb{P}(X_T = e_i).$$

(2.4.10)

2.4.3 The characteristic function of the occupation time

To calculate the expectation terms in (2.4.10), we shall use the conditional characteristic function $\hat{\Phi}_i$ of the vector $(J_{t,T}^1, J_{t,T}^2, \ldots, J_{t,T}^N)$ given $X_T = e_i$ defined in the following lemma.

Lemma 6. Consider the vector of occupation times $J_{t,T} = (J_{t,T}^1, J_{t,T}^2, \ldots, J_{t,T}^N)'$ and the Fourier transformation vector $u = (u_1, u_2, \ldots, u_N)' \in \mathbb{R}^N$. Then the conditional characteristic function of $J_{t,T}$ given $X_T = e_i$ is given by:

$$\hat{\Phi}_i = \mathbb{E}[e^{-i\langle u, J_{t,T} \rangle} | \mathcal{F}_t] = \left\langle \exp \left( (A + \text{diag}(-iu))(T - t) \right) X_T, e_i \right\rangle.$$

(2.4.11)

Here $\mathcal{F}_t = \mathcal{F}_t^W \lor \mathcal{F}_t^X$ is the information generated by the Brownian motion and Markov chain up to time $t$.

Proof. Suppose $u = (u_1, u_2, \ldots, u_N)' \in \mathbb{R}^N$, and

$$e^{-i\langle u, J_{t,T} \rangle} = \exp \left( -i \left( u_1 \int_t^T \langle X_s, e_1 \rangle ds + \cdots + u_N \int_t^T \langle X_s, e_N \rangle ds \right) \right).$$
CHAPTER 2. THE STOCHASTIC DISCOUNT FACTOR

Then
\[
d(e^{-i(u,J_t,s)}X_s) = e^{-i(u,J_t,s)}X_s \left(-i \left( u_1(X_s,e_1)ds + \cdots + u_N(X_s,e_N)ds \right) \right) + e^{-i(u,J_t,s)}(AX_s ds + dM_s).
\]

In integral form, this is:
\[
e^{-i(u,J_t,T)}X_T = X_t + \int_t^T e^{-i(u,J_t,s)}X_s \left(-i \left( \sum_{j=1}^N u_j(X_s,e_j) \right) \right) ds + \int_t^T e^{-i(u,J_t,s)}AX_s ds + \int_t^T e^{-i(u,J_t,s)}dM_s.
\]

Conditioning on the filtration \(F_t\), we have:
\[
E[e^{-i(u,J_t,T)}X_T|F_t] = X_t + \int_t^T E[e^{-i(u,J_t,s)}X_s \left(-i \left( \sum_{j=1}^N u_j(X_s,e_j) \right) \right)]|F_t]ds + \int_t^T E[e^{-i(u,J_t,s)}AX_s|F_t]ds + E[\int_t^T e^{-i(u,J_t,s)}dM_s|F_t].
\]

The last term is equal to zero because \(M\) is a martingale. Now,
\[
X_s \left(-i \left( \sum_{j=1}^N u_j(X_s,e_j) \right) \right) = diag(-iu)X_s.
\]

Write
\[
Z_s = E[e^{-i(u,J_t,s)}X_T|F_t] \in \mathbb{R}^N.
\]

Then
\[
Z_T = X_t + \int_t^T diag(-iu)E[e^{-i(u,J_t,s)}X_s|F_t]ds + \int_t^T A\mathbb{E}e^{-i(u,J_t,s)}X_s|F_t]ds
\]
\[
= X_t + \int_t^T diag(-iu)Z_s ds + \int_t^T A Z_s ds
\]
\[
= X_t + \int_t^T (A + diag(-iu)) Z_s ds \in \mathbb{R}^N.
\]

The solution of this ordinary differential equation is:
\[
Z_{t,T} = E[e^{-i(u,J_t,T)}X_T|F_t] = \exp \left( (A + diag(-iu))(T - t) \right) X_t \in \mathbb{R}^N. \quad (2.4.12)
\]
Having calculated $\mathbb{E}[e^{-i(u,J_t,T)}X_T|\mathcal{F}_t]$, we can use this result to find the conditional characteristic function of $(J_{l,T}^1, J_{l,T}^2, \ldots, J_{l,T}^N)$ given $X_T = e_i$. This follows as:

$$\langle Z_{l,T}, e_i \rangle = \langle \exp \left( (A + \text{diag}(-iu))(T - t) \right) X_t, e_i \rangle = \mathbb{E}[e^{-i(u,J_t,T)}\langle X_T, e_i \rangle|\mathcal{F}_t].$$ (2.4.13)

Equation (2.4.14)

The conditional call price determined in (2.4.8) is:

$$C_t = \mathbb{E}[(H^*(J_{l,T}^1, J_{l,T}^2, \ldots, J_{l,T}^N), X_T)|\mathcal{F}_t].$$

When $X_T = e_i$, we have

$$\langle H^*(J_{l,T}^1, J_{l,T}^2, \ldots, J_{l,T}^N), X_T \rangle = \langle H^*(J_{l,T}^1, J_{l,T}^2, \ldots, J_{l,T}^N), e_i \rangle = H^*_i(J_{l,T}^1, J_{l,T}^2, \ldots, J_{l,T}^N).$$

The call option price can, therefore, be written:

$$C_t = \sum_{i=1}^{N} \mathbb{E}[H^*_i(J_{l,T}^1, J_{l,T}^2, \ldots, J_{l,T}^N)|\mathcal{F}_t \vee X_T = e_i] \mathbb{P}(X_T = e_i).$$ (2.4.15)

On the other hand, each of the terms $\mathbb{E}[H^*_i(J_{l,T}^1, J_{l,T}^2, \ldots, J_{l,T}^N)|\mathcal{F}_t \vee X_T = e_i]$ can be calculated using the conditional characteristic function found in Lemma 6. For the $i^{th}$ component of the vector $H^*$ we obtain:

$$\mathbb{E}[H^*_i(J_{l,T})|\mathcal{F}_t \vee X_T = e_i] = \int_{[0,T-t]^N} H^*_i(y_1, \ldots, y_N)\Phi(y_1, \ldots, y_N)dy_1 \ldots dy_N$$ (2.4.16)

Where $\Phi$ is the conditional distribution function of the occupation times $J_{l,T} = (J_{l,T}^1, J_{l,T}^2, \ldots, J_{l,T}^N)'$. The last integral cannot be calculated directly as we do not have the conditional density function of the occupation times. However, the Fourier transform of this random vector was calculated in Lemma 6. Therefore, we can proceed using Parseval’s theorem.
2.4.4 Parseval’s theorem and Fourier transforms

The Fourier transform is a much used technique in option pricing due to its flexibility for calculating the expectation of discounted claims without direct knowledge of the underlying probability density.

In Carr and Madan [10] the authors show how the fast Fourier transform can be used to value options when the characteristic function is known. Lewis [49] has used the Fourier transform to price European stock options, where the underlying is a general Lévy process. The paper by Dufresne et al. [17] presents different applications of Parseval’s theorem depending on various payoff functions.

Consider a random variable $X$ with a density function $\mu_X$ with a total mass $|\mu_X| < \infty$ and $g$ a function of bounded variation such that $g \in L^1$. Then with $PV$ referring to the principal value of the integral (see Appendix B):

$$
\mathbb{E}[g(X)] = \frac{1}{2\pi} PV \int \hat{g}(-u) \hat{\mu}_X(u) du
$$

(2.4.17)

where $\hat{g}$ and $\hat{\mu}$ are, respectively, the Fourier transform of the differentiable function $g$ and the density function of $X$.

In our case the expectation of the function $H^*_i$ is reduced to an integration over the time interval $[0, T-t]$, moreover, it is straightforward to verify that $H^*_i$ satisfies the assumptions of Parseval’s Theorem. See [17], for example. So from (2.4.16):

$$
\mathbb{E}[H^*_i(Y_1, \ldots, Y_N)|\mathcal{F}_t \lor X_T = e_i] = \frac{1}{2\pi} \int_{\mathbb{R}^N} \hat{H}^*_i(-u_1, \ldots, -u_N) \hat{\Phi}_i(u_1, \ldots, u_N) du_1 \ldots du_N.
$$

(2.4.18)

Here $\hat{\Phi}_i$ is the conditional characteristic function of the vector $J_{t,T}$ given in (2.4.13):

$$
\hat{\Phi}_i = \langle Z_{t,T}, e_i \rangle = \langle \exp \left( (A + \text{diag}(-iu))(T - t) \right) X_t, e_i \rangle.
$$

The probability $\mathbb{P}(X_T = e_i)$ can be calculated from the Markov chain dynamics. From the martingale decomposition of $X_t$:

$$
dX_t = AX_t dt + dM_t.
$$
where $M_t$ is a $\mathbb{P}$-martingale.

In integral form, this is:

$$X_T = X_t + \int_t^T AX_s ds + \int_t^T dM_s.$$ 

Taking the conditional expectation on $X_t$, the solution of the system of differential equations is:

$$\mathbb{E}[X_T|X_t] = e^{A(T-t)}X_t.$$ 

Therefore,

$$\mathbb{P}(X_T = e_i|X_t) = \mathbb{E}[1_{X_T=e_i}|X_t] = \mathbb{E}[\langle X_T, e_i \rangle|X_t] = \langle \mathbb{E}[X_T|X_t], e_i \rangle = \langle e^{A(T-t)}X_t, e_i \rangle.$$ 

Therefore, we can calculate the call price at time $t$. The put option price can be derived similarly, or from the put-call parity.

### 2.4.5 The hedging of a European call

In our model, we consider two sources of randomness, the Brownian motion and the Markov jump process. Using the stochastic discount factor we can find a value for a European claim, priced by the real world probability measure. However, since we are in the presence of two sources of randomness, the model is incomplete and there is more than one risk neutral probability. Therefore, as has been shown by Harrison and Pliska [39, 38], it is not possible to replicate the claim with a portfolio composed only of a non-risky bond and the underlying stock; in other words we need to complete our model by adding assets. Here, as in [37], we choose other traded options with different maturities. Futures on the the underlying stock are also good candidate for completing the market.

Assume that the stock price process is given by (2.3.14), where $X_t$ is an $N$ state Markov chain. In order to complete the market, beside the non-risky asset and the underlying stock, we need to include $N$ other financial instruments ($\Lambda^1, \ldots, \Lambda^N$). As for the previously calculated call price these instruments will depend on the dividend
CHAPTER 2. THE STOCHASTIC DISCOUNT FACTOR

payment level, and the current state of the economy, so we consider a self financing portfolio $V_t$ of the form:

$$V_t = \alpha(t)B_t + \beta^{(0)}(t)S_t + \sum_{k=1}^{N} \beta^{(j)}_t \Lambda^j(t).$$  \hspace{1cm} (2.4.19)

Here $\alpha(t)$ is the quantity held in the portfolio from the non risky asset at time $t$, $\beta^{(0)}(t)$ is the quantity held of the underlying asset and $\beta^{(j)}(t)$ is the quantity of the instrument $\Lambda^k$ held in the portfolio at time $t$. By the arbitrage pricing, we require:

$$C(t, \delta_t, X_t) = V_t.$$

Differentiation of (2.4.19) with respect to the dividend payment gives the following equation:

$$\frac{\partial V_t}{\partial \delta_t} = \frac{\partial C_t}{\partial \delta_t} = \beta^{(0)}(t)\zeta_t + \sum_{k=1}^{N} \beta^{(j)}(t) \frac{\partial \Lambda^j(t)}{\partial \delta_t}. \hspace{1cm} (2.4.20)$$

As we have a semi-closed formula for the call price, we can calculate in (2.4.20) the derivative of the hedged claim and the other instruments so that the only unknown quantity will be the instruments’ weights $\beta^j$. On the other hand, a variation in the instruments values when the state of the economy $X$ changes will give the following system of equations

$$\Delta C_{ij}(t) = \beta^{(0)}(t)\Delta S_{ij}(t) + \sum_{k=1}^{N} \beta^{(j)}(t) \Delta \Lambda_{ij}(t).$$  \hspace{1cm} (2.4.21)

Here $i, j \in \{1, \ldots, N\}$, and

$$\begin{cases} 
\Delta C_{ij}(t) = C_j(t) - C_i(t) \\
\Delta S_{ij}(t) = S_j(t) - S_i(t). \hspace{1cm} (2.4.22)
\end{cases}$$

The variations here are with respect to the state of the economy. From (2.4.21) and (2.4.22) we can obtain a complete system of equations, that is $N+1$ linear equations in $N+1$ unknowns, allowing us to determine the weight quantities of the stock and instruments needed to hedge the call option.

In order to find the instrument weights, and consequently a hedging portfolio, we
need to calculate the derivative of the instrument price with respect to the current
dividend payment. Recall that the price of the call and the hedging instruments are
given by:

\[ C_t = \sum_{k=1}^{N} \mathbb{E}[H_i^*(J_{i,T}^1, J_{i,T}^2, \ldots, J_{i,T}^N)|\mathcal{F} \vee X_T = e_i]\mathbb{P}(X_T = e_i). \]

The final state of the economy does not depend on \( \delta_t \), and the same applies for
the conditional characteristic function of the occupation times. Therefore using
Parseval’s theorem for the above expectation terms,

\[ \frac{\partial C_t}{\partial \delta_t} = \mathbb{P}(X_T = e_i) \sum_{k=1}^{N} \frac{1}{2\pi} \int_{\mathbb{R}^N} \frac{\partial \hat{H}_i^*(-u)}{\partial \delta_t} \hat{\phi}_i(u) du. \]

Here \( u \) is the Fourier transformation vector. Recall

\[ H_i^*(x) = \exp((\theta + \frac{1}{2}(\gamma^2 + \phi^2), x))v_i \delta_t e^{N\left(\sqrt{\langle \phi^2, x \rangle} - \frac{1}{\sqrt{\langle \phi^2, x \rangle}} \log\left(\frac{K}{v_i \delta_t e^{\langle T, x \rangle}}\right)\right)} - KN\left(\frac{1}{\sqrt{\langle \phi^2, x \rangle}} \log\left(\frac{K}{v_i \delta_t e^{\langle T, x \rangle}}\right)\right) \]

and

\[ \hat{H}_i^*(x) = \int_{\mathbb{R}^N} e^{i(x \cdot u)} H_i^*(x) dx. \]

The function \( H_i^* \) satisfies Parseval’s regularity conditions [17], so:

\[ \frac{\partial \hat{H}_i^*}{\partial \delta_t}(x) = \int_{\mathbb{R}^N} e^{i(x \cdot u)} \frac{\partial H_i^*}{\partial \delta_t}(x) du. \]

Therefore

\[ \begin{cases} 
-d_1^i &= \sqrt{\langle \phi^2, x \rangle} - \frac{1}{\sqrt{\langle \phi^2, x \rangle}} \log\left(\frac{K}{v_i \delta_t e^{\langle T, x \rangle}}\right) \\
-d_2^i &= \frac{1}{\sqrt{\langle \phi^2, x \rangle}} \log\left(\frac{K}{v_i \delta_t e^{\langle T, x \rangle}}\right),
\end{cases} \]

then
\[
\frac{\partial H^*_t(x)}{\partial \delta_t} = \frac{\partial}{\partial \delta_t} \left( \exp(\langle \theta + \frac{1}{2} (\gamma^2 + \phi^2), x \rangle) v_i \epsilon N(-d_1^i) - K N(-d_2^i) \right) \\
= \exp(\langle \theta + \frac{1}{2} (\gamma^2 + \phi^2), x \rangle) \epsilon v_i \left( N(-d_1^i) + \frac{1}{\sqrt{\langle \phi^2, x \rangle}} f(-d_1^i) \right) - \frac{K}{\delta_t \sqrt{\langle \phi^2, x \rangle}} f(-d_2^i) \\
= R_t(\delta_t, x).
\]

Here \( f(x) \) is the normal distribution density function. Therefore:
\[
\frac{\partial H^*_t(x)}{\partial \delta_t} = R(x),
\]

with \( R(x) = (R_N(\delta_t, x), \ldots, R_N(\delta_t, x))' \).

Consequently, in the above calculations, we have obtained semi-analytical expression for the required quantities to find the hedging ratio.

In the following, we shall present further results derived from the stochastic discount factor model.

### 2.4.6 Exchange rates

We define the exchange rate as the rate at which one currency can be exchanged for another. See [65] for further definitions. Write \( a \) for the domestic currency and \( b \) for the foreign currency. Write \( X_t^{ab} \) for the value of 1 unit of \( a \) in terms of \( b \) units at time \( t \).

Suppose the stochastic discount factors in countries are expressed as:
\[
\pi^{a}_{t,T} = \exp \left( \int_t^T \langle \theta^a, X_u \rangle du + \int_t^T \langle \gamma^a, X_u \rangle dW_u \right),
\]
and
\[
\pi^{b}_{t,T} = \exp \left( \int_t^T \langle \theta^b, X_u \rangle du + \int_t^T \langle \gamma^b, X_u \rangle dW_u \right),
\]
CHAPTER 2. THE STOCHASTIC DISCOUNT FACTOR

where

\[ \theta^a = (\theta^a_1, \ldots, \theta^a_N)' \quad \gamma^a = (\gamma^a_1, \ldots, \gamma^a_N)' , \]

\[ \theta^b = (\theta^b_1, \ldots, \theta^b_N)' \quad \gamma^b = (\gamma^b_1, \ldots, \gamma^b_N)' . \]

Then we have the following result for \( X^{ab}_t \):

**Lemma 7.**

\[ X^{ab}_t = X^{ab}_0 \pi^a_t \pi^b_t \]

\[ = X^{ab}_0 \exp \left( \int_0^t \langle \theta^{ab}, X_u \rangle du + \int_0^t \langle \gamma^{ab}, X_u \rangle dW_u \right). \quad (2.4.23) \]

Here \( \theta^{ab} = \theta^a - \theta^b \) and \( \gamma^{ab} = \gamma^a - \gamma^b \).

**Proof.** From the definition of the stochastic discount factor, for \( s \leq t \) and for an asset of price \( A_u \) in country \( a \) at time \( u \):

\[ \mathbb{E}[\pi^a_t A_t | \mathcal{F}_s] = \pi^a_s A_s . \]

On the other hand

\[ \mathbb{E}[\pi^b_t X^{ab}_t A_t | \mathcal{F}_s] = \pi^b_s X^{ab}_s A_s . \]

So

\[ A_s = \mathbb{E}[\frac{\pi^a_t A_t}{\pi^a_s}] = \mathbb{E}[\frac{\pi^b_t X^{ab}_t A_t}{\pi^b_s X^{ab}_s}] | \mathcal{F}_s] . \]

For any asset with price process \( A_t \):

\[ \mathbb{E} \left[ \left( \frac{\pi^a_t}{\pi^a_s} - \frac{\pi^b_t X^{ab}_t}{\pi^b_s X^{ab}_s} \right) A_t | \mathcal{F}_s \right] = 0 . \]

Taking \( s = 0 \),

\[ \mathbb{E} \left[ \left( \frac{\pi^a_t}{\pi^a_0} - \frac{\pi^b_t X^{ab}_t}{X^{ab}_0} \right) A_t \right] = 0 . \]

Assuming enough liquidity and positivity of the asset price, we conclude that

\[ X^{ab}_t = X^{ab}_0 \frac{\pi^a_t}{\pi^b_t} . \]

\( \Box \)
We now find the dynamics of $X_{t}^{ab}$.

We have

$$X_{t}^{ab} = X_{0}^{ab} \exp \left( \int_{0}^{t} \langle \theta_{ab}, X_{u} \rangle du + \int_{0}^{t} \langle \gamma_{ab}, X_{u} \rangle dW_{u} \right).$$

So from Itô’s lemma

$$dX_{t}^{ab} = X_{t}^{ab} \left( \langle \theta_{ab}, X_{t} \rangle dt + \langle \gamma_{ab}, X_{t} \rangle dW_{t} \right) + \frac{1}{2} X_{t}^{ab} \langle \gamma_{ab}, X_{t} \rangle^2 dt.$$

On the other hand, we know that

$$r_{t} = -\langle \theta + \frac{1}{2} \gamma^{2}, X_{t} \rangle.$$

and

$$\theta^{ab} + \frac{1}{2} \langle \gamma^{ab} \rangle^2 = \langle \theta + \frac{1}{2} \gamma^{2}, X_{t} \rangle - \langle \gamma^{b}, X_{t} \rangle.$$

Therefore,

$$\langle \theta^{ab} + \frac{1}{2} \langle \gamma^{ab} \rangle^2, X_{t} \rangle = r^{b} - r^{a} + \langle \gamma^{b}, X_{t} \rangle.$$

Therefore,

$$dX_{t}^{ab} = X_{t}^{ab} (r^{b} - r^{a}) dt + X_{t}^{ab} \langle \gamma^{b}, X_{t} \rangle dt + X_{t}^{ab} \langle \gamma^{ab}, X_{t} \rangle dW_{t}.$$

As shown in Corollary 2, under $P_{b}^{b}$, the risk neutral measure in country $b$:

$$W_{t}^{b*} = W_{t} - \int_{0}^{t} \gamma_{s}^{b} ds.$$

is a Brownian motion. Therefore

$$dX_{t}^{ab} = X_{t}^{ab} (r^{b} - r^{a}) dt + X_{t}^{ab} \langle \gamma^{ab}, X_{t} \rangle dW_{t}^{b*}. \quad (2.4.24)$$

These dynamics for the exchange rate satisfy the no arbitrage condition. Indeed, assume that the exchange rate has dynamics of a geometric Brownian motion:

$$dX_{t}^{ab} = X_{t}^{ab} \mu dt + X_{t}^{ab} \sigma dW_{t}.$$
and we are modelling from the point of view of an investor in country \( b \), so that \( B_t \)
refers to a Brownian motion under \( \mathbb{P}_b^* \), the risk neutral probability in country \( b \).
Let \( A_t \) and \( B_t \) denote the share price of the money market in country \( a \) and \( b \)
respectively, then:
\[
A_t = \exp(r_a t),
\]
and
\[
B_t = \exp(r_b t).
\]
So the share price of the money market in country \( a \) at time \( t \) expressed in country
\( b \) currency and after solving the stochastic differential equation is:
\[
A_t X_{at}^{ab} = X_{0}^{ab} \exp((r_a + \mu)t - \frac{1}{2} \sigma^2 t + \sigma W_t).
\]
Because of the absence of arbitrage opportunities, the discounted price of the money
market in country \( a \) in terms of country \( b \) currency is a martingale under \( \mathbb{P}_b^* \). On
the other hand the discounted share price is:
\[
\exp(-r_b t) X_{t}^{ab} = X_{0}^{ab} \exp \left( (\mu + r_a - r_b)t - \frac{\sigma^2}{2} t + \sigma W_t \right).
\]
Since
\[
\mathcal{E} = \exp \left( -\frac{\sigma^2}{2} t + \sigma W_t \right)
\]
is martingale,
we must have \( \mu = r_b - r_a \).

### 2.4.7 Bond dynamics

In this section we shall find the dynamics of a zero-coupon contract. This allows us
to find the dynamics of the forward rate.

Before we proceed, we recall the following result about inverse matrices whose proof
is given in Van der Hoek and Elliott [62].

**Lemma 8.** Let \( \Phi(t, u) \) be the unique solution of the following system of Ordinary
Differential Equations,
\[
\frac{\partial}{\partial u} \Phi(t, u) = \Gamma_u \Phi(t, u) \quad \text{and} \quad \Phi(t, t) = I.
\]  \hfill (2.4.25)
CHAPTER 2. THE STOCHASTIC DISCOUNT FACTOR

Then as a function of \(t\), \(\Phi(t, u)\) is the unique solution of the following system of Ordinary Differential Equations:

\[
\frac{\partial}{\partial t} \Phi(t, u) = -\Gamma_t \Phi(t, u) \quad \Phi(t, t) = I. \tag{2.4.26}
\]

**Proof.** Similarly, the system of Ordinary Differential Equations

\[
\frac{\partial}{\partial u} \Psi(t, u) = -\Psi(t, u) \Gamma_u \quad \Phi(t, t) = I \tag{2.4.27}
\]

has unique solution. Then

\[
\frac{\partial}{\partial u} (\Psi(t, u) \Phi(t, u)) = -\Psi(t, u) \Gamma_u \Phi(t, u) \Psi(t, u) \Gamma_u \Phi(t, u) = 0.
\]

So \(\Psi(t, u) \Phi(t, u)\) is constant in \(u\), that is:

\[
\Psi(t, u) \Phi(t, u) = \Psi(t, t) \Phi(t, t) = I.
\]

So

\[
\Psi(t, u)^{-1} = \Phi(t, u).
\]

Now consider \(Z(t, T) = \Psi(t, u) \Phi(u, T)\). Then \(Z(t, t) = I\) and \(Z\) is independent of \(u\), so:

\[
\frac{\partial Z(t, T)}{\partial T} = -\Psi \Gamma_u \Phi(u, T) + \Psi(t, u) \frac{\partial}{\partial u} \Phi(u, T) = 0,
\]

and as \(\Psi(t, u)\) is a non-singular matrix,

\[
\frac{\partial}{\partial u} \Phi(u, T) = -\Gamma_u \Phi(u, T). \tag{2.4.28}
\]

\[\square\]
By definition, we have:

\[ P(t,T) = \mathbb{E} \left[ \frac{\pi_T}{\pi_t} \middle| \mathcal{F}_t \right] \]

\[ = \mathbb{E} [\pi_{t,T} | \mathcal{F}_t] \]

\[ = \mathbb{E} [\langle H_{t,T}, 1 \rangle | \mathcal{F}_t] \]

where \( H_{t,T} = \langle \hat{H}_{t,T}, 1 \rangle \)

\[ = \langle \Psi(t,T), X, 1 \rangle \]

(2.4.29)

So

\[ \frac{d}{dt} P(t,T) = \langle \partial_t \Psi(t,T)X_t, 1 \rangle dt + \langle \Psi(t,T)AX_t, 1 \rangle dt + \langle \Psi(t,T)dM_t, 1 \rangle \]

Since \( \Psi(t,T) \) satisfies (2.2.11) \( = \langle -\Psi(t,T)KX_t, 1 \rangle dt + \langle \Psi(t,T)AX_t, 1 \rangle dt + \langle \Psi(t,T)dM_t, 1 \rangle \)

\[ = \langle \Psi(t,T)(A - K)X_t, 1 \rangle dt + \langle \Psi(t,T)dM_t, 1 \rangle . \]

Recall that

\[ r_t = -\langle KX_t, 1 \rangle = -\langle \text{diag}(\theta + \frac{1}{2}\gamma^2)X_t, 1 \rangle , \]

and from (2.2.10),

\[ (A - K) = -\text{diag}(\theta + \frac{1}{2}\gamma^2) . \]

Therefore

\[ \langle \Psi(t,T)(A - K)X_t, 1 \rangle = -\langle \Psi(t,T)\text{diag}(\theta + \frac{1}{2}\gamma^2)X_t, 1 \rangle . \]

We also notice that

\[ -\text{diag}(\theta + \frac{1}{2}\gamma^2)X_t = r_tX_t \]

The dynamics of \( P(t,T) \) become:

\[ \frac{d}{dt} P(t,T) = r_t \langle \Psi(t,T)X_t, 1 \rangle dt + \langle \Psi(t,T)dM_t, 1 \rangle \]

(2.4.30)

Finally

\[ \frac{d}{dt} P(t,T) = r_tP(t,T)dt + \langle \Psi(t,T)dM_t, 1 \rangle . \]
CHAPTER 2. THE STOCHASTIC DISCOUNT FACTOR

Now that we have obtained the dynamics of \( P(t, T) \), we can derive the dynamics of the forward rate \( f(t, T) \). We know that

\[
P(t, T) = \exp(-\int_t^T f(t, u)du).
\]

Therefore,

\[
f(t, T) = -\frac{\partial}{\partial T} (\log(P(t, T))) = -\frac{\partial}{\partial T} P(t, T).
\]

So from (2.4.29) and Lemma 8, we have

\[
f(t, T) = -\frac{\langle K \Psi(t, T) X_t, 1 \rangle}{\langle \Psi(t, T) X_t, 1 \rangle} = -\frac{\langle K \exp(K(T - t)) X_t, 1 \rangle}{\langle \exp(K(T - t)) X_t, 1 \rangle}.
\]

(2.4.32)

Putting

\[
\Phi^i(t, T) = -\frac{\langle K \exp(K(T - t)) e_i, 1 \rangle}{\langle \exp(K(T - t)) e_i, 1 \rangle},
\]

and

\[
\Phi(t, T) = (\Phi^1(t, T), \Phi^2(t, T), \ldots, \Phi^N(t, T)),
\]

we have

\[
f(t, T) = \langle \Phi(t, T), X_t \rangle.
\]

(2.4.33)

Having obtained an expression for forward rates \( f(t, T) \), we now determine their dynamics.

**Theorem 4.** The forward interest rate has the following dynamics

\[
\frac{\partial}{\partial t} f(t, T) = \left( r_t - \frac{\langle \Psi(t, T), X_t \rangle}{P(t, T)} \right) f(t, T) + \langle \Phi(t, T), dM_t^* \rangle.
\]

(2.4.34)

where \( M^* \) is a \( \mathbb{R}^N \)-martingale.

**Proof.** From (2.4.33) we have:

\[
\frac{\partial}{\partial t} f(t, T) = \left\langle \frac{\partial}{\partial t} \Phi(t, T), X_t \right\rangle dt + \langle \Phi(t, T), AX_t \rangle dt + \langle \Phi(t, T), dM_t^* \rangle
\]

\[
= \left\langle \frac{\partial}{\partial t} \Phi(t, T) + A^* \Phi(t, T), X_t \right\rangle dt + \langle \Phi(t, T), dM_t^* \rangle.
\]

We know

\[
\Phi^i(t, T) = -\frac{\partial}{\partial T} \log(\Psi(t, T)e_i, 1).
\]
Therefore

\[
\frac{\partial}{\partial t} \Phi^i(t, T) = \frac{\partial}{\partial t} \left[ -\frac{\partial}{\partial T} \log \langle \Psi(t, T)e_i, 1 \rangle \right]
= -\frac{\partial}{\partial T} \left[ \frac{\partial}{\partial t} \log \langle \Psi(t, T)e_i, 1 \rangle \right]
= -\frac{\partial}{\partial T} \left[ -\langle \Psi(t, T)Ke_i, 1 \rangle \langle \Psi(t, T)e_i, 1 \rangle \right]
\]
from Lemma 8

\[
= \langle \Psi(t, T)e_i, 1 \rangle - 2 \left[ \langle \Psi(t, T)e_i, 1 \rangle \langle \Psi(t, T)Ke_i, 1 \rangle \right] - \langle \Psi(t, T)Ke_i, 1 \rangle \langle \Psi(t, T)e_i, 1 \rangle .
\]

(2.4.35)

However

\[
(A^*\Phi(t, T))^i = -\sum_{j=1}^{N} A_{ji} \frac{\langle K\Psi(t, T)e_j, 1 \rangle}{\langle \Psi(t, T)e_j, 1 \rangle}
= -\sum_{j=1}^{N} \langle \text{diag}(1'\Psi(t, T))^{-1} K \Psi(t, T)e_j, 1 \rangle A_{ji}
= -(1'(\text{diag}(1'\Psi(t, T))^{-1}) K \Psi(t, T)A)^i.
\]

Therefore,

\[
\langle A^*\Phi(t, T), X_t \rangle = -(1'(\text{diag}(1'\Psi(t, T))^{-1}) K \Psi(t, T)A)X_t
= -(1'diag(1'\Psi(t, T))^{-1} K \Psi(t, T)(K - diag(\theta + \gamma^2 2)))X_t
= -(1'diag(1'\Psi(t, T))^{-1} K \Psi(t, T)(K)X_t
+ (1'diag(1'\Psi(t, T))^{-1} K \Psi(t, T)diag(\theta + \gamma^2 2))X_t .
\]

Also, from (2.2.15) we have

\[
diag(\theta + \gamma^2 2)X_t = -r_t X_t .
\]

Therefore,

\[
\langle A^*\Phi(t, T), X_t \rangle = -\frac{1'K\Psi(t, T)KX_t}{1'\Psi(t, T)X_t} + \frac{1'K\Psi(t, T)X_t}{1'\Psi(t, T)X_t} r_t .
\]

(2.4.36)
So from (2.4.36) and (2.4.35), we have:
\[
\langle \frac{\partial}{\partial t} \Phi(t, T) + A^* \Phi(t, T), X_t \rangle = r_t \langle K \Psi(t, T) X_t, 1 \rangle - \langle K \Psi(t, T) X_t, 1 \rangle - \langle \Psi(t, T) X_t, 1 \rangle.
\]
\[
= r_t f(t, T) - f(t, T) \frac{\langle \Psi(t, T) K X_t, 1 \rangle}{P(t, T)}.
\]

\[\square\]

### 2.4.8 The forward measure

Suppose the forward measure \( \mathbb{P}^T \) is defined by:
\[
\frac{d\mathbb{P}^T}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} = \frac{\pi_T}{\mathbb{E}[\pi_T]}.
\]  
(2.4.37)

Then for any \( \mathcal{F}_T \)-measurable variable, we have from Bayes’ theorem:
\[
\mathbb{E}^T[Y|\mathcal{F}_t] = \mathbb{E}[\pi_T Y|\mathcal{F}_t] \frac{\mathbb{E}[\pi_T]}{\mathbb{E}[\pi_T]} = \mathbb{E}[\pi_{t,T} Y|\mathcal{F}_t] \frac{\mathbb{E}[\pi_{t,T}]}{\mathbb{E}[\pi_{t,T}]} = \mathbb{E}[\pi_{t,T} Y|\mathcal{F}_t] \frac{P(t,T)}{\mathbb{E}[\pi_{t,T}]}.
\]

**Lemma 9.**
\[
f(t, T) = \mathbb{E}^T[r_T|\mathcal{F}_t].
\]  
(2.4.38)

**Proof.** From (2.4.32) we have
\[
f(t, T) = -\frac{\langle K \Psi(t, T) X_t, 1 \rangle}{\langle \Psi(t, T) X_t, 1 \rangle} - \frac{\langle K \mathbb{E}[H_{t,T} X_t, 1] \rangle}{\langle H_{t,T} X_t, 1 \rangle} - \frac{\mathbb{E}[\langle K \mathbb{E}[\pi_{t,T} X_t, 1] \rangle|\mathcal{F}_t]}{P(t,T)} - \frac{\mathbb{E}[\pi_{t,T} \langle K X_T, 1 \rangle|\mathcal{F}_t]}{P(t,T)}.
\]
\[
= \frac{\mathbb{E}[\pi_{t,T} r_T|\mathcal{F}_t]}{P(t,T)}, \text{ from (2.2.15)}
\]
\[
= \mathbb{E}^T[r_T|\mathcal{F}_t].
\]
2.5 Results summary

In this section, we summarise the main results of this chapter.

2.5.1 Risk neutral measure

As defined in Section 2.2.2, the new probability measure $\mathbb{P}^*$ defined by:

$$\frac{d\mathbb{P}^*}{d\mathbb{P}}|_{\mathcal{F}_T} = \frac{\beta_T \pi_T}{\mathbb{E}[\beta_T \pi_T]}.$$

is a probability under which the discounted price of an asset is a martingale.

2.5.2 Stock price

We assume that the stock price is determined by all future dividend payments. These are modelled by a Geometric Brownian motion $\delta_t$. Several expressions for the dynamics of the stock price were found depending on the application. Most simply the stock price is:

$$S_t = \langle Q_t, X_t \rangle \quad \text{where } Q_t = \delta_t (P^{-1})' \text{diag}\left(\frac{-1}{\lambda_1}, \frac{-1}{\lambda_2}, \ldots, \frac{-1}{\lambda_N}\right) P' 1.$$

2.5.3 European stock option

The option price was expressed in terms of the Fourier transform of the occupation times of the Markov chain and the occupation probability:

$$c_t = \sum_{i=1}^{N} \mathbb{E}[H_i^* (J_{i,T}^1, J_{i,T}^2, \ldots, J_{i,T}^N)|\mathcal{F} \vee X_T = e_i] \mathbb{P}(X_T = e_i).$$

Here

$$\mathbb{E}[H_i^* (Y_1, \ldots, Y_N)|\mathcal{F} \vee X_T] = \frac{1}{2\pi} \int_{\mathbb{R}^N} \hat{H}_i(y_1, \ldots, y_N) \hat{\Phi}_i(y_1, \ldots, y_N) dy_1 \ldots dy_N.$$
and
\[ \widehat{\Phi}_i = \langle Z_{t,T}, e_i \rangle = \langle \exp \left( (A + \text{diag}(-iu))(T - t) \right) X_t, e_i \rangle. \]
is the conditional characteristic function of the occupation time of the Markov chain.\[ \widehat{H}_i \]is the Fourier transform of the function (2.4.9) and
\[
\mathbb{P}(X_T = e_i | X_t) = \langle e^{A(T-t)}X_t, e_i \rangle.
\]
is the probability for the Markov chain to be in state \( i \).

### 2.5.4 Bond prices and dynamics

The bond with maturity \( T \) has a price at \( t \leq T \).

\[
P(t, T) = \langle \exp(K(T - t)), X_t, 1 \rangle,
\]

\[
\frac{d}{dt}P(t, T) = rtP(t, T)dt + \langle \Psi(t, T), M_t, 1 \rangle.
\]

Here \( M_t \) is a martingale under \( \mathbb{P} \) and \( \mathbb{P}^* \).

### 2.5.5 Exchange rates

If \( X_t^{ab} \) represents the value of one unit of currency in country \( a \) in terms of currency in country \( b \), then:

\[
dX_t^{ab} = X_t^{ab}(r^b - r^a) + X_t^{ab} \langle \gamma^{ab}, X_t \rangle dW_t^{*b}.
\]

Here \( W_t^{*b} \) is risk neutral Brownian motion under \( \mathbb{P}^*_b \).

### 2.5.6 Forward rates

If \( f(t, T) \) represents the forward rate, then:

\[
\frac{\partial}{\partial t} f(t, T) = \left( rt - \frac{\langle \Psi(t, T), X_t \rangle}{P(t, T)} \right) f(t, T) + \langle \Phi(t, T), dM_t^* \rangle.
\]

Here \( M \) is a martingale under \( \mathbb{P} \) and \( \mathbb{P}^* \).
2.5.7 Forward measure

With a forward measure defined by:

\[
\frac{dP_T}{dP} = \frac{\pi_T}{\mathbb{E}[\pi_T]},
\]

where \( \pi_T = \pi_{0,T} \), we obtain a change of measure for the forward rates, under which the forward rate is a martingale.

2.5.8 Forward rates

As a result of the forward measure, we have the martingale property satisfied for the forward rates. Recall that \( r_T = f(T, T) \). Then we have:

\[
f(t, T) = \mathbb{E}^T[r_T | \mathcal{F}_t].
\]

2.6 Conclusion and further research

The first section of this chapter is an overview of the prior literature discussing derivative pricing using a stochastic discount factor, as well as the hypothesis that share prices can be modelled as the net present value of future dividend payments. The second section gives a model for the Stochastic Discount Factor. We assume the latter has two sources of randomness, the first comes from a Wiener process describing the short term randomness of the market prices, resulting from the market equilibrium between the offer and the demand, and the second is modelled by using a continuous time Markov jump process, which models possible medium and long term changes in the state of the economy. This model for the SDF has the dynamics of a regime switching geometric Brownian motion; it is then exploited to price the basic financial derivatives such as the zero-coupon bonds, forwards and the short rate. From the SDF model we construct a risk neutral measure framework, in which we define the new dynamics for the Markov jump process as well as a new Brownian motion.
In section 3 we consider a model for stock pricing. A regime switching Black and Scholes model is used to describe the randomness of the share dividend payments, while the stock price, the underlying, is defined as the expectation of the discounted value of all future dividends. We first derive an implicit formula for the stock price as a function of the state of economy and the level of dividends, beside the SDF model parameters. An assumption on the sign of the model parameters allows us to obtain a more explicit second expression for the stock price.

In section 4 this stock price model is used to derive the European call option price. The model includes two sources of randomness, which requires as a first step considering a larger filtration where the information generated by the Markov jump process is known. Later the option price is obtained using the conditional characteristic function of the random vector of occupation times. Finally, the call option price is derived using Parseval’s theorem, the Fourier transform of the stock price density and the conditional characteristic function. A subsection addresses the $\delta$-hedging strategy for the call, using a self-financed portfolio created from a risk-less asset and traded instruments with different maturities. The section concludes with further results produced by the model such as the currency exchange rate and the forward curve dynamics.

Section 5 summarises the main results and formulae of the chapter.

Further work can focus on numerical testing of the model on real market data. For a small dividends data sample, methods like interpolation can be used.
Chapter 3

Stochastic discount factor in power market

3.1 Introduction

In this chapter we shall discuss a new model for power prices. In the last two decades energy markets around the world were, and still are, converting to a deregulated trade economy. Financial engineering has become an important tool used to hedge the risk resulting from the volatile price of the underlying.

Commodities in general are a different class of assets from the usual purely financial traded instruments, such as stocks and bonds. The nature of the commodity itself is a major factor in defining the behaviour of its prices, and also prices of the wide range of contracts issued on the spot or forward price.

One important feature of the power market is that electricity is expensive and difficult to store; on the other hand the market operators must keep an equilibrium between the power generators’ output and the consumers’ demand on a continuous basis, using trading strategies to keep the market balanced. See [30, 31]. However, the non-flexibility of the offer and demand for the commodity causes the prices to have a particular patterns. Four features can describe the general behaviour of elec-
Electricity prices: seasonality, mean reversion, jumps or spikes, and negative prices. Because the traded price is the equilibrium price, a seasonality in power demand, mainly driven by the weather, is behind the seasonality affecting the spot level. Small changes in the quantity of power generated, or an expected rise in demand, can cause very steep variations in the spot price. These extreme changes appear as spikes in the spot curve. Unlike stocks, power prices are not allowed to evolve freely in the long term, but even with high jumps, the prices gravitate around a long term average. This is reflected in the long term return of the power price to the cost of production. This typical pattern is modelled in previous literature using an Ornstein Uhlenbeck process, (See [60]), to simulate the mean reversion behaviour. The spikes are modelled as jump processes. (See [12, 46]). In more recent works Markov jump processes are used. (See [34], [27], [45]).

The seasonality in the price can be modelled using a deterministic component. Some authors use constant piece-wise linear functions. (See [47]). Sinusoidal functions are also used to model seasonality as in Pilipovic [57]. In this work we follow an approach which combines both a stylised function, as in Lucia and Schwartz [50], which models the seasonality as a deterministic function with a sinusoidal component in addition to a piece-wise linear function.

In some volatile electricity markets, such as the Australian case, an imbalance in the offer demand equilibrium may result in a very sharp drop in the power price giving negative values for a short period of time. More details and techniques used in this case are discussed in Chapter 5.

Combining the above processes, we shall propose a model for the electricity spot price. With the latter as the underlying we shall derive the price of a forward contract, defined to be the discounted value of the future spot price using a non arbitrage argument. Here we use results for the Stochastic Discount Factor from Chapter 2. Prices of other derivatives, such as a call on forwards and swaps, are also derived.
3.2 Presenting the model

Suppose \((\Omega, \mathcal{F}, \{\mathcal{F}\}, \mathbb{P})\) is a complete filtered probability space, where \(\mathbb{P}\) is the real world probability measure. Suppose the short term randomness of the price of power is modelled by the following regime switching mean reverting diffusion process \(X_t\) having the dynamics:

\[
dX_t = \kappa(\langle \mu; Z_t \rangle - X_t)dt + \sigma dW_t. \tag{3.2.1}
\]

Here \(Z\) is a finite state continuous time Markov chain. As before, without loss of generality we may assume that \(Z\) takes values in \(\{e_1, \ldots, e_N\}\), the canonical basis of the Euclidian space \(\mathbb{R}^N\). The parameter \(\sigma\) is the volatility of the process and \(\sigma \in \mathbb{R}\); the parameter \(\mu = (\mu_1, \mu_2, \ldots, \mu_N) \in \mathbb{R}^N\) is a long term mean of the process which switches from one state to another according to the Markov process \(Z\). The parameter \(\kappa\) is the mean reversion rate, or, as some authors prefer to say, the speed of return, and \(\kappa \in \mathbb{R}\). See [48] for more on OU processes. \(W\) is a Brownian motion which is independent of \(Z\).

The above model for the electricity spot log-price is a regime switching Ornstein-Uhlenbeck process. In the classical pricing of financial instruments, where geometric Brownian motions are widely used to model the underlying dynamics, it is assumed that price changes are independent. In the power market, the mean reversion is thought to model the empirical observations that electricity prices usually fluctuate around a long term mean, which is very close to the cost of production. For more about using the mean reverting processes in energy market, see [7].

Further, we have the following dynamics for the Markov chain \(Z\):

\[
Z_t = Z_0 + \int_0^t AZ_s ds + \int_0^t dM_s. \tag{3.2.2}
\]

Here \(A\) is the transition rate matrix of \(Z\) and \(M\) is an \(\mathbb{R}^N\) martingale.

Taking \(X_0 = 0\), the solution of the above stochastic differential equation (3.2.1) is:

\[
X_t = \int_0^t e^{\kappa(u-t)}\langle \kappa \mu; Z_u \rangle du + \int_0^t \sigma e^{\kappa(u-t)} dW_u. \tag{3.2.3}
\]
We then assume the spot price of the commodity at time \( t \) is:

\[
S_t = f_t \exp(X_t)\langle \alpha, Z_t \rangle.
\] (3.2.4)

Here \( \alpha = (\alpha_1, \ldots, \alpha_N) \in \mathbb{R}^N \) models shock jumps in spot price. \( f_t \) is a deterministic function modelling the seasonality of the spot price and is given by:

\[
f_t = \epsilon + \beta t + dD_{\text{day}} + \delta \sin(\frac{2\pi}{365} t + \varphi),
\] (3.2.5)

where \( \epsilon, \beta, d, \delta \) and \( \varphi \) are constant parameters and \( D_{\text{day}} \) is a dummy variable taking values in \( \{1, \ldots, 7\} \). The function \( f_t \) is chosen to have this particular form in order to replicate the periodic behaviour of power prices, since there is a yearly, weekly seasonality. We assume here that we have a daily average price of electricity, which allows us to neglect the daily seasonality effect.

We also consider a pricing Kernel process or a stochastic discount factor \( \pi \), such that for a liquid enough tradable asset with price process \( A \), we have for time indices \( s \leq t \):

\[
\mathbb{E}[\pi_t A_t | \mathcal{F}_s] = \pi_s A_s.
\] (3.2.6)

Here the expectation is taken with respect to the historical measure \( \mathbb{P} \). We assume the process \( \pi \) is a regime switching geometric Brownian motion given by:

\[
\pi_t = \exp \left( \int_0^t \langle \theta, Z_u \rangle du + \int_0^t \langle \gamma, Z_u \rangle dW_u \right).
\] (3.2.7)

Here \( \theta = (\theta_1, \ldots, \theta_N) \) and \( \gamma = (\gamma_1, \ldots, \gamma_N) \) are regime switching parameter vectors. Our processes are defined on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \), where \( \mathcal{F} = \{ \mathcal{F}_t, t \geq 0 \} \) is the right continuous, complete filtration generated by the Brownian motion \( W \) and the Markov chain \( Z \). We also write \( \mathcal{F}^W = \{ \mathcal{F}_t^W, t \geq 0 \} \) and \( \mathcal{F}^Z = \{ \mathcal{F}_t^Z, t \geq 0 \} \) for the right continuous, complete filtrations generated by the Brownian motion and the Markov chain respectively.
3.3 Pricing of the forward

Power market investors, whether suppliers or producers, decide on the quantity of electricity to buy or sell from empirical estimation of the energy needed to satisfy the consumers’ need. Because the spot price is extremely volatile and can change by a factor of 10 in the space of few hours, it is mandatory for investors to hedge against possible losses.

The most traded power derivatives in the market are the forward contracts. These enable the holder to buy the commodity at a future time at a price fixed today. In other words the forward contract is based on how the investor expects the spot market to change by the delivery time.

Using the above models for the spot price and the pricing kernel, we shall calculate the forward price for a contract agreed at time $s$ for delivery time $t \geq s$. This forward price $f(s, t)$ is given by:

$$\pi_s f(s, t) = \mathbb{E}[\pi_t S_t | \mathcal{F}_s].$$

**Remark 7.** Here we note that $f(t, t) = S_t$, as the forward price converges to the spot price.

If we use the notation $\pi_{s,t} = \frac{\pi_t}{\pi_s}$, we have:

$$f(s, t) = \mathbb{E}[\pi_{s,t} S_t | \mathcal{F}_s].$$ (3.3.1)

**Lemma 10.**

$$f(s, t) = f_t \exp \left( \int_0^s e^{\kappa(u-t)} \langle \kappa \mu, Z_u \rangle du + \int_0^s \sigma e^{\kappa(u-t)} dW_u \right) \exp \left( \frac{1}{2} \int_s^t \sigma^2 e^{2\kappa(u-t)} du \right) \mathbb{E}[\Gamma_{s,t} \langle \alpha, Z_t \rangle | \mathcal{F}_s].$$ (3.3.2)

**Proof.** Using repeated conditioning:

$$f(s, t) = \mathbb{E}[\pi_{s,t} S_t | \mathcal{F}_s^W \vee \mathcal{F}_s^Z]$$
$$= \mathbb{E}[\mathbb{E}[\pi_{s,t} S_t | \mathcal{F}_s^W \vee \mathcal{F}_t^Z] | \mathcal{F}_s^W \vee \mathcal{F}_s^Z].$$
Using the repeated conditioning, we assume initially that we know the history of $Z$ up to time $t$. We then condition back to time $s$. Now:

$$\pi_{s,t}S_t = f_t \exp \left( \int_0^s e^{\kappa(u-t)} \langle \kappa \mu, Z_u \rangle du + \int_0^s \sigma e^{\kappa(u-t)} dW_u \right) \exp \left( \int_s^t (\theta + \kappa \mu e^{\kappa(u-t)} , Z_u) du + \int_s^t (\sigma e^{\kappa(u-t)} + \langle \gamma, Z_u \rangle) dW_u \right) \langle \alpha, Z_t \rangle.$$ 

Write

$$D(0, s, t) = f_t \exp \left( \int_0^s e^{\kappa(u-t)} \langle \kappa \mu, Z_u \rangle du + \int_0^s \sigma e^{\kappa(u-t)} dW_u \right) \exp \left( \int_s^t (\theta + \kappa \mu e^{\kappa(u-t)} , Z_u) du \right) \langle \alpha, Z_t \rangle. \quad (3.3.3)$$

Then

$$E[\pi_{s,t}S_t | F^{W}_s \lor F^{Z}_t] = D(0, s, t)E \left[ \exp \left( \int_s^t (\sigma e^{\kappa(u-t)} + \langle \gamma, Z_u \rangle) dW_u \right) | F^{W}_s \lor F^{Z}_t \right]. \quad (3.3.4)$$

Now

$$\int_s^t (\sigma e^{\kappa(u-t)} + \langle \gamma, Z_u \rangle) dW_u | F^{W}_s \lor F^{Z}_t \sim \mathcal{N} \left( 0, \int_s^t (\sigma e^{\kappa(u-t)} + \langle \gamma, Z_u \rangle)^2 ds \right),$$

where $\mathcal{N}(a, b)$ represents the normal probability distribution with mean $a$ and variance $b$.

Therefore,

$$E \left[ \exp \left( \int_s^t (\sigma e^{\kappa(u-t)} + \langle \gamma, Z_u \rangle) dW_u \right) | F^{W}_s \lor F^{Z}_t \right] = \exp \left( \frac{1}{2} \int_s^t (\sigma e^{\kappa(u-t)} + \langle \gamma, Z_u \rangle)^2 ds \right).$$

Write:

$$\Gamma_{s,t} = \exp \left( \int_s^t (\theta + e^{\kappa(u-t)} (\kappa \mu + \sigma \gamma) + \frac{1}{2} \gamma^2, Z_u) du \right)$$

$$= \exp \left( \int_s^t \langle W_u, Z_u \rangle du \right)$$

where

$$W_u = \theta + e^{\kappa(u-t)} (\kappa \mu + \sigma \gamma) + \frac{1}{2} \gamma^2. \quad (3.3.5)$$
Then:

\[
\begin{align*}
  f(s,t) &= f_t \exp \left( \int_0^s e^{\kappa(u-t)} (\kappa \mu, Z_u) \, du + \int_0^s \sigma e^{\kappa(u-t)} \, dW_u \right) \\
    & \quad \times \exp \left( \frac{1}{2} \int_s^t \sigma^2 e^{2\kappa(u-t)} \, du \right) \mathbb{E}[\Gamma_{s,t}(\alpha, Z_t) \mid \mathcal{F}_s] .
\end{align*}
\]  

\text{(3.3.6)}

Lemma 11. Write:

\[
\mathbb{W}_u = \text{diag}(\theta + e^{\kappa(u-t)}(\kappa \mu + \sigma \gamma) + \frac{1}{2} \gamma^2),
\]

a diagonal deterministic matrix. Recall \( A \) is the transition rate matrix of \( Z \). Suppose \( \Phi(s,t) \) is the matrix solution of the differential equation

\[
\frac{d\Phi(s,u)}{du} = (\mathbb{W}_u + A)\Phi(s,u)
\]

\[\Phi(s,s) = I_N .\]

Then

\[\mathbb{E}[\Gamma_{s,t}(\alpha, Z_t) \mid \mathcal{F}_s] = (\Phi(s,t) Z_s, \alpha) .\]

Proof. Consider

\[Y_u = \Gamma_{s,u} Z_u .\]

So

\[dY_u = d\Gamma_{s,u} Z_u + \Gamma_{s,u} dZ_u .\]

We have:

\[dZ_u = AZ_u + dM_u .\]

Also:

\[d\Gamma_{s,u} = (\mathbb{W}_u, Z_u) \Gamma_{s,u} du .\]

Then

\[dY_u = (\mathbb{W}_u, Z_u) Y_u du + AY_u du + \Gamma_{s,u} dM_u ,\]
and
\[ Y_t = Y_s + \int_s^t (W_u, Z_u) Y_u du + \int_s^t AY_u du + \int_s^t \Gamma_{s,u} dM_u. \]

Since the last integral is a martingale,
\[ \mathbb{E}[Y_t|\mathcal{F}_s] = Y_s + \mathbb{E}[\int_s^t (W_u, Z_u) Y_u du|\mathcal{F}_s] + \mathbb{E}[\int_s^t AY_u du|\mathcal{F}_s] \]
\[ = Z_s + \int_s^t (W_u + A) \mathbb{E}[Y_u|\mathcal{F}_s] du. \]

where \( W_u \) is the diagonal matrix we obtain from having the vector \( W_u \) on the diagonal.

Consequently
\[ \mathbb{E}[Y_t|\mathcal{F}_s] = \Phi(s,t)Z_s. \]

Then,
\[ \mathbb{E}[\Gamma_{s,t}\langle \alpha, Z_t \rangle|\mathcal{F}_s] = \mathbb{E}[\langle \alpha, Y_t \rangle|\mathcal{F}_s] \]
\[ = \langle \alpha, \mathbb{E}[Y_t|\mathcal{F}_s] \rangle \]
\[ = \langle \Phi(s,t)Z_s, \alpha \rangle. \]

Write:
\[ V^2(s,t) = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa(t-s)}). \]

Then the forward price is:
\[ f(s,t) = f_t \exp \left( \int_0^s e^{\kappa(u-t)} \langle \kappa \mu, Z_u \rangle du + \int_0^s \sigma e^{\kappa(u-t)} dW_u + \frac{V^2(s,t)}{2} \right) \langle \Phi(s,t)Z_s, \alpha \rangle \]
\[ = f_t \exp \left( e^{\kappa(s-t)} X_s + \frac{V^2(s,t)}{2} \right) \langle \Phi(s,t)Z_s, \alpha \rangle. \]

(3.3.7)

Note the above expression depends only on the model parameters and the spot level.
3.4 Pricing of options on forward

An option on a forward is a contract which allows the holder to enter the forward at a future date for a later delivery. Consider a forward price process entered at time \( s \) for delivery at time \( t \), where \( s \leq t \). We know from (3.3.7) that the forward price at \( s \) is:

\[
f(s, t) = G(s, t, Z) \exp\left(\int_0^s \sigma e^{-\kappa(t-u)} dW_u \right).
\]  

(3.4.1)

Here

\[
G(s, t, Z) = f_t \exp\left(\int_0^s e^{\kappa(u-t)} \langle \kappa \mu, Z_u \rangle du + \frac{\mathcal{V}^2(s, t)}{2} \right) \langle \Phi(s, t) Z_s, \alpha \rangle.
\]

For \( 0 \leq s \leq t \) consider a call option at \( \tau = 0 \) to enter a forward at \( \tau = s \) with delivery at time \( \tau = t \), with a strike price \( K \). Using the pricing kernel, the call option price is:

\[
C(0, s, t) = \mathbb{E}\left[\pi_s (f(s, t) - K)^+ | \mathcal{F}_0 \right].
\]  

(3.4.2)

Recall that:

\[
\pi_s = \exp\left(\int_0^s \langle \theta, Z_u \rangle du + \int_0^s \langle \gamma, Z_u \rangle dW_u \right).
\]

We shall use repeated conditioning as for the forward price. Denote by \( \mathcal{F}_0^W \vee \mathcal{F}_s^Z \) the sigma algebra containing the history of \( Z \) up to time \( s \).

Then

\[
C(0, s, t) = \mathbb{E}[\mathbb{E}[\pi_s (f(s, t) - K)^+ | \mathcal{F}_0^W \vee \mathcal{F}_s^Z] | \mathcal{F}_0].
\]

Proposition 3. Write

\[
C(0, s, t; Z) = \mathbb{E}[\pi_s (f(s, t) - K)^+ | \mathcal{F}_0^W \vee \mathcal{F}_s^Z].
\]  

(3.4.3)

and define the quantities

\[
d_2 = -\log\left(\frac{G^{**}(s, t, Z)}{G^{**}(s, t, Z)}\right) + \frac{1}{2} \int_0^s \sigma^2 e^{2\kappa(u-t)} du \right) \right] \right) + \int_0^s \sigma^2 e^{-2\kappa(t-u)} du \right)
\]

(3.4.4)

where

\[
G^{**}(s, t, Z) = \exp\left(\int_0^s \sigma e^{\kappa(u-t)} \langle \gamma, Z_u \rangle du \right) \exp\left(\frac{1}{2} \int_0^s \sigma^2 e^{2\kappa(u-t)} du \right) G(s, t, Z)
\]  

(3.4.5)
CHAPTER 3. STOCHASTIC DISCOUNT FACTOR IN POWER MARKET 60

Then

\[ C(0, s, t; Z) = \exp \left( \int_0^s \langle \theta + \frac{1}{2} \gamma^2, Z_u \rangle du \right) \left[ G^{**}(s, t, Z) \mathcal{N}(d_1) - K \mathcal{N}(d_2) \right]. \tag{3.4.6} \]

Proof. Recall:

\[ f(s, t) = G(s, t, Z) \exp \left( \int_0^s \sigma e^{-\kappa(t-u)} dW_u \right). \]

So

\[ C(0, s, t; Z) = \exp \left( \int_0^s \langle \theta, Z_u \rangle du \right) \mathbb{E} \left[ \exp \left( \int_0^s \langle \gamma, Z_u \rangle dW_u \right) (f(s, t) - K)^+ | \mathcal{F}_0^W \vee \mathcal{F}_s^Z \right]. \]

Consider the process \( \Lambda \) such that:

\[ \Lambda_s = \exp \left( \int_0^s \langle \gamma, Z_u \rangle dW_u - \frac{1}{2} \int_0^s \langle \gamma^2, Z_u \rangle du \right). \tag{3.4.7} \]

Using the process \( \Lambda \) we define \( \mathbb{P}^\gamma \), an equivalent probability to the historical measure \( \mathbb{P} \). That is, we define a new measure \( \mathbb{P}^\gamma \) by setting:

\[ \frac{d\mathbb{P}^\gamma}{d\mathbb{P}} |_{\mathcal{F}_s} = \Lambda_s. \tag{3.4.8} \]

By Girsanov’s theorem, under \( \mathbb{P}^\gamma \), the process \( W^\gamma \) defined by:

\[ W^\gamma_u = W_u - \int_0^u \langle \gamma, Z_{\tau} \rangle d\tau \]

is a Brownian motion. Then:

\[ C(0, s, t; Z) = \exp \left( \int_0^s \langle \theta + \frac{1}{2} \gamma^2, Z_u \rangle du \right) \mathbb{E} \left[ \Lambda_s (f(s, t) - K)^+ | \mathcal{F}_0^W \vee \mathcal{F}_s^Z \right] \]

\[ = \exp \left( \int_0^s \langle \theta + \frac{1}{2} \gamma^2, Z_u \rangle du \right) \mathbb{E}^\gamma \left[ (f(s, t) - K)^+ | \mathcal{F}_0^W \vee \mathcal{F}_s^Z \right] \]

\[ = \exp \left( \int_0^s \langle \theta + \frac{1}{2} \gamma^2, Z_u \rangle du \right) \mathbb{E}^\gamma \left[ G^*(s, t, Z) \exp \left( \int_0^s \sigma e^{-\kappa(t-u)} dW^\gamma_u \right) - K \right]^+ | \mathcal{F}_0^W \vee \mathcal{F}_t^Z \right]. \]

Here

\[ G^*(s, t, Z) = G(s, t, Z) \exp \left( \int_0^s \sigma e^{\kappa(t-u)} \langle \gamma, Z_u \rangle du \right). \]
CHAPTER 3. STOCHASTIC DISCOUNT FACTOR IN POWER MARKET

Then
\[ C(0, s, t; Z) = \exp \left( \int_0^s \langle \theta + \frac{1}{2} \gamma^2, Z_u \rangle du \right) \times \]
\[ \mathbb{E}^\gamma \left[ \left( \frac{d}{dZ} \right)^* (s, t, Z) \exp \left( \int_0^s \sigma e^{-\kappa(t-u)} dW^\gamma_u - \frac{1}{2} \int_0^s \sigma^2 e^{2\kappa(u-t)} du \right) - K \right]^+ |\mathcal{F}_0^W \lor \mathcal{F}_s^Z]. \]

Note that
\[ \exp \left( \int_0^s \sigma e^{-\kappa(t-u)} dW^\gamma_u - \frac{1}{2} \int_0^s \sigma^2 e^{2\kappa(u-t)} du \right) \]
is an exponential martingale. Then we have:
\[ C(0, s, t; Z) = \exp \left( \int_0^s \langle \theta + \frac{1}{2} \gamma^2, Z_u \rangle du \right) \left[ \frac{d}{dZ} (s, t, Z) \mathcal{N}(d_1) - K \mathcal{N}(d_2) \right]. \]

Consider the following random processes:
\[ \rho_s = \int_0^s \langle \theta + \frac{1}{2} \gamma^2, Z_u \rangle du \]
and
\[ \chi_s = \int_0^s e^{\kappa(u-t)} \langle \mu + \sigma \gamma, Z_u \rangle du. \]

We have
\[ C(0, s, t; Z) = W(\rho, \chi, Z_s) \]
where \( W \) is a \( C^{2,1} \) function. Let \( \lambda(\rho, \chi, e_j) \) be the joint density function of \( \rho, \chi | Z_s = e_j \).

Then
\[
C(0, s, t) = \mathbb{E}[W(\rho, \chi, Z_s)|\mathcal{F}_0^W,Z]
= \sum_{1 \leq j \leq N} \mathbb{E}[W(\rho, \chi, e_j)|\mathcal{F}_0^W,Z]P(Z_s = e_j)
= \sum_{1 \leq j \leq N} P(Z_s = e_j) \int_{\mathbb{R}} \int_{\mathbb{R}} W(\rho, \chi, e_j) \lambda(\rho, \chi, e_j) d\rho d\chi.
\]

By abuse of notation here, \( \rho \) and \( \chi \) stand for both the random variable and the observations. For some \((u, v) \in \mathbb{R}^2\), let \( \hat{\lambda}(u, v, e_j) \) be the Fourier transform of
\( \hat{\lambda}(u,v,e_j) \), that is:

\[
\hat{\lambda}(u,v,e_j) = \mathbb{E} \left[ \frac{1}{\sqrt{2\pi}} \exp \left( iu\rho_s + iv\chi_s \langle Z_s, e_j \rangle \right) \right]
= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{(iu\rho_s + iv\chi_s)} \lambda(\rho, \chi, e_j) d\rho d\chi.
\]

To calculate \( \hat{\lambda}(u,v,e_j) \) for \( 1 \leq j \leq N \), we introduce the random vector \( Y \) defined as:

for \( s \in \mathbb{R}^+ \):

\[
Y_s = \exp(iu\rho_s + iv\chi_s)Z_s.
\] (3.4.10)

**Lemma 12.** Consider the diagonal matrix \( F \), such that:

\[
F_s = iu \text{diag}(\theta + \frac{1}{2}\gamma^2) + iv\kappa(s-t) \text{diag}(\mu + \sigma\gamma)
\] (3.4.11)

and let the matrix \( \Psi \) be the solution of the system of differential equations:

\[
\frac{d\Psi(u)}{du} = (A + F_s)\Psi(u)
\]

\[
\Psi(0) = I_N.
\] (3.4.12)

Then

\[
\forall s \geq 0 \quad \mathbb{E}[Y_s] = \Psi(s)Z_0.
\] (3.4.13)

**Proof.** From the definition:

\[
Y_s = \exp(iu\rho_s + iv\chi_s)Z_s
\]

and

\[
dZ_s = AZ_s + dM_s.
\]

So

\[
dY_s = \exp(iu\rho + iv\chi)dZ_s + Y_s \left[ iu\langle \theta + \frac{1}{2}\gamma^2, Z_s \rangle + iv\langle \mu + \sigma\gamma; Z_s \rangle \right] ds
= (A + F_s)Y_s ds + \exp(iu\rho + iv\chi)dM_s,
\]
and, since the process $M$ is a martingale under $\mathbb{P}$,

$$
\mathbb{E}[Y_s] = Z_0 + \int_0^s (A + F_\tau)\mathbb{E}[Y_\tau]d\tau.
$$

Therefore,

$$
\mathbb{E}[Y_s] = \Psi(s)Z_0.
$$

From the previous lemma, we conclude that:

$$
\hat{\lambda}(u,v,e_j) = \mathbb{E}\left[\frac{1}{\sqrt{2\pi}} \exp(iup_s + iv\chi_s)\langle Z_s, e_j \rangle\right]
$$

$$
= \frac{1}{\sqrt{2\pi}} \mathbb{E}[\langle Y_s, e_j \rangle]
$$

$$
= \frac{1}{\sqrt{2\pi}} \langle \mathbb{E}[Y_s], e_j \rangle
$$

$$
= \frac{1}{\sqrt{2\pi}} \langle \Psi(s)Z_0, e_j \rangle.
$$

Using Parseval’s theorem:

$$
C(0,s,t) = \sum_{1 \leq j \leq N} P(Z_s = e_j) \int_\mathbb{R} \int_\mathbb{R} W(\rho, \chi, e_j)\lambda(\rho, \chi, e_j)d\rho d\chi
$$

$$
= \sum_{1 \leq j \leq N} P(Z_s = e_j) \int_\mathbb{R} \int_\mathbb{R} \hat{W}(\rho, \chi, e_j)\hat{\lambda}(\rho, \chi, e_j)d\rho d\chi
$$

(3.4.14)

where

$$
\hat{W}(\rho, \chi, Z_s) = \frac{1}{\sqrt{2\pi}} \int_\mathbb{R} \int_\mathbb{R} e^{iup_s+iv\chi_s} W(\rho, \chi, e_j)d\rho d\chi.
$$

### 3.5 Conclusion

In this chapter, we extend the previous chapter results on the stochastic discount factor to electricity and power derivative pricing. Here we use a regime switching mean reverting Ornstein-Uhlenbeck with a jump process to model the power price. So we are in the presence of two sources of randomness, a Brownian motion to model
the short term changes in the price, and a jump process to model the peak-load prices or spikes, which requires double conditioning with respect to larger filtrations. Defining the forward price as the discounted value of the future spot price, we derive an explicit formula for the forward in terms of the current spot price and the model parameters. The price of a call on forward is also calculated using Parseval’s theorem to obtain a Black and Scholes-like formula.
Chapter 4

Power pricing using compensated jump processes

4.1 Introduction

In this chapter we present a new model for power pricing where the market randomness is modelled by a compensated jump process. Jump processes were introduced in pricing theory at an early stage, with Merton’s jump diffusion model [52], to simulate assets with prices having spike-like patterns. This type of model was much used later in commodity and energy markets for spot based modelling due to the specific price distribution with fat tails and extreme variation in a short time interval.

Several papers have adopted the jump diffusion models for electricity markets, see [11, 53, 2], but the mean reversion is always chosen as a more realistic model for energy prices, see [33]. As shown in Carr et al. [9] it is possible not to include a diffusion process if we allow the jump term to have an infinite activity. In this chapter we follow this argument and we restrict to a compensated jump process to model the market randomness. The compensation allows to have a martingale, and consequently independent increment property for the price, see [18].

In this chapter we start from a mean reverting Ornstein-Uhlenbeck-like process
where the diffusion term is replaced with a compensated jump process. We develop a model framework to price a forward and swap contract with the spot electricity price being the underlying. The prices formulae are obtained in terms of general expression of the jump process and its compensator, and can be adapted to any chosen jump process like poisson, variance gamma or inverse gaussian.

4.2 Presenting the model

Suppose \((\Omega, \mathcal{F}, \{\mathcal{F}\}, \mathbb{P})\) is a complete filtered probability space, where \(\mathbb{P}\) is the real world probability measure. Suppose the short term randomness of the price of power is modelled by a mean reverting jump process \(X_t\) with the dynamics:

\[
    dX_t = \kappa(\mu - X_t)dt + \sigma dM_t. \tag{4.2.1}
\]

Here \(M_t\) is a compensated pure jump process, such that:

\[
    M_t = \int_0^t \int_{\mathbb{R}} x(\gamma - \tilde{\gamma}(x))(dx, du),
\]

where \(\tilde{\gamma}\) is the compensator or the predictable projection of the jump measure, that is, \(M_t\) is a local martingale, (see Appendix A.5). It is possible to include a diffusion process in the spot model, but as shown in Carr et al. [9], it is not necessary if the model allows for infinite activity of the jumps. This condition is satisfied when:

\[
    \int_{\mathbb{R}} \tilde{\gamma}(x)dx = \infty. \tag{4.2.2}
\]

Also we can allow for small activity when moving away from the origin, that is, the occurrence of jumps becomes less when there size becomes larger.

By definition, a compensated jump process is a martingale which is constructed by subtracting the compensator, or dual predictable projection, of the pure jump process. The features of the power price, precisely the spikes, suggest the inclusion of a jump or Lévy process in the pricing of electricity derivatives. Processes such as a compound Poisson and Variance-Gamma, see [26], are widely used. In the literature
we can find explicit expressions for compensators of various processes; see [56, 26] for example. In the following we suppose the compensator or the Lévy measure of the jump process is of the form:

\[ \tilde{\gamma} = \tilde{k}(x). \]  

(4.2.3)

Since the asset price is a positive quantity we model the asset price \( S_t \) as:

\[ S_t = S_0 f_t \exp(X_t), \]  

(4.2.4)

where \( f_t \) is the deterministic seasonal function defined in (3.2.5). The process \( X_t \) with the dynamics given in (4.2.1) satisfies for \( 0 \leq s \leq t \):

\[ X_t = e^{-\kappa(t-s)}X_s + \int_s^t \kappa \mu e^{-\kappa(t-u)} du + \int_s^t \sigma e^{-\kappa(t-u)} dM_u. \]

So

\[ S_t = (S_s)^{e^{-\kappa(t-s)}} \exp \left( \int_s^t \kappa \mu e^{-\kappa(t-u)} du + \int_s^t \sigma e^{-\kappa(t-u)} dM_u \right). \]  

(4.2.5)

Without loss of generality, we assume \( S_0 = 1 \), so:

\[ S_t = \exp \left( \int_0^t \kappa \mu e^{-\kappa(t-u)} du + \int_0^t \sigma e^{-\kappa(t-u)} dM_u \right). \]  

(4.2.6)

4.3 Power derivatives pricing

In this section we shall develop a framework to price some of the most traded electricity financial securities, the forwards and the swap contracts.

4.3.1 Risk-neutral measure

If a compensated jump process is used to model the electricity spot price, the usual change of probability to a risk neutral using the classical Girsanov’s theorem, as in the Brownian motion case, is no longer possible.

Consequently we present a Girsanov’s theorem for pure jump processes; see [18]. Let \( h : \mathbb{R} \to [0, \infty) \) be a function such that the process \( \Lambda \) is a martingale where:
\[ \Lambda_t = 1 + \int_0^t \Lambda_s^- \left( \int_{-\infty}^{+\infty} (h(x) - 1) \gamma(dx, ds) - \bar{k}(x) dx ds \right) \]  
\[ = \exp \left( - \int_0^t \int_{-\infty}^{+\infty} (h(x) - 1) \bar{k}(x) dx ds + \int_0^t \int_{-\infty}^{+\infty} \log(h(x)) \gamma(dx, ds) \right). \]  

Let \( \{\mathcal{F}_t\} \) denote the right continuous complete filtration generated by \( X \), and define a probability measure \( \mathbb{P}^h \) equivalent to \( \mathbb{P} \) by the Radon-Nykodim derivative:

\[ \frac{d\mathbb{P}^h}{d\mathbb{P}} |_{\mathcal{F}_t} = \Lambda_t. \]  

Then if \( M \) is a \( \mathbb{P} \)-martingale, we have:

\[ M_t - \int_0^t \int_{-\infty}^{+\infty} x (h(x) - 1) \bar{k}(x) dx ds. \]

is a \( \mathbb{P}^h \)-martingale.

More generally, if \( \tilde{\gamma} \) is the compensator of \( \gamma \) under the original probability measure \( \mathbb{P} \), then \( h \tilde{\gamma} \) is the compensator under the new probability measure \( \mathbb{P}^h \).

From arbitrage theory, (see [5]), the a risk neutral probability \( \mathbb{P}^h \) is a probability being equivalent to \( \mathbb{P} \) such that for tradable assets in the market, the discounted price is martingale. So we are looking for an \( h : \mathbb{R} \to [0, \infty) \) such that the process \( (e^{-rt}S_t)_{t \geq 0} \) is martingale under \( \mathbb{P}^h \). Using the differentiation rule for jump processes, see [18]:

\[ F(X_t) = F(X_0) + \int_0^t F'(X_{s^-}) dX_s + \sum_{0 \leq s \leq t} (F(X_s) - F(X_{s^-}) - F'(X_{s^-}) \Delta X_s), \]

Here \( \Delta X_s = X_s - X_{s^-} \).

**Proposition 4.** If the electricity spot price model is given by (4.2.5), and \( \mathbb{P}^h \) is an equivalent probability measure to \( \mathbb{P} \), then \( \mathbb{P}^h \) is a risk neutral measure if and only if the function \( h \) introduced in (4.3.1) satisfies:

\[ r = \int_{\mathbb{R}} (e^{ax} - 1) h(x) \bar{k}(x) dx + \kappa (\mu - X_t) - \int_{\mathbb{R}} \sigma x \bar{k}(x) dx. \]  

\[ (4.3.3) \]
CHAPTER 4. POWER PRICING USING COMPENSATED JUMP PROCESSES

Proof. We obtain the following dynamics for price process:

\[ dS_t = \exp(X_{t^-})dX_t + (\exp(X_t) - \exp(X_{t^-}) - \exp(X_{t^-})\Delta X_t) \]

\[ = \exp(X_{t^-}) \left( \kappa(\mu - X_t)dt + \int_{\mathbb{R}} \sigma x(\gamma - \hat{\gamma})(dx, dt) \right) \]

\[ + \exp(X_{t^-}) \left[ \int_{\mathbb{R}} (e^{\sigma x} - 1)\gamma(dx, dt) - \sigma \int_{\mathbb{R}} x\gamma(dx, dt) \right] \]

\[ = \exp(X_{t^-}) \left( \kappa(\mu - X_t)dt - \int_{\mathbb{R}} \sigma x\hat{\gamma}(dx, dt) \right) + \exp(X_{t^-}) \int_{\mathbb{R}} (e^{\sigma x} - 1)\gamma(dx, dt). \]

The appearance of the factor \((e^x - 1)\) motivates us to consider the process:

\[ Z_t = S_t \exp(Y_t), \]

where

\[ \exp(Y_t) = \exp \left( -\int_0^t \int_{\mathbb{R}} (e^{\sigma x} - 1)h(x)\hat{\gamma}(dx, dt) - \int_0^t \kappa(\mu - X_s)ds + \int_0^t \int_{\mathbb{R}} \sigma x\hat{\gamma}(dx, dt) \right). \]

Since the processes \(X\) and \(Y\) do not have common jumps, the dynamics of the process \(Z\) are:

\[ dZ_t = d(\exp(X_t)) \exp(Y_t) + \exp(X_t)d(\exp(Y_t)). \]

Then

\[ dZ_t = Z_t \left[ \kappa(\mu - X_t)dt - \sigma \int_{\mathbb{R}} \sigma x\hat{\gamma}(dx, dt) + \int_{\mathbb{R}} (e^{\sigma x} - 1)\gamma(dx, dt) \right] \]

\[ + Z_t \left[ -\int_{\mathbb{R}} (e^{\sigma x} - 1)h(x)\hat{\gamma}(dx, dt) - \kappa(\mu - X_t)dt + \int_{\mathbb{R}} \sigma x\hat{\gamma}(dx, dt) \right] \]

\[ = Z_t \int_{\mathbb{R}} (e^{\sigma x} - 1) (\gamma - h(x)\hat{\gamma})(dx, dt). \]

and the process \(Z\) is a martingale under \(\mathbb{P}^h\).

Consequently the discounted price process is a martingale under \(\mathbb{P}^h\) if and only if for all \(t \geq 0\):

\[ Y_t = -rt. \]

So \(\mathbb{P}^h\) is a risk-neutral measure if \(h\) satisfies:

\[ r = \int_{\mathbb{R}} (e^{\sigma x} - 1)h(x)k(x)dx + \kappa(\mu - X_t) - \int_{\mathbb{R}} \sigma x\bar{k}(x)dx. \]
CHAPTER 4. POWER PRICING USING COMPENSATED JUMP PROCESSES

Remark 8. Determining the risk neutral probability measure is subject to defining a function $h$ satisfying (4.3.3), however, there are many functions that can be chosen. We can choose for instance $h$ to be independent of $x$ directly, that is from (4.3.3):

$$h = \frac{r + \int_{\mathbb{R}} \sigma x k(x) dx - \kappa (\mu - \log(\frac{S_T}{S_0}))}{\int_{\mathbb{R}} (e^{\sigma x} - 1) k(x) dx}.$$ (4.3.4)

4.3.2 Pricing of forwards

Consider a forward contract purchased at time $t$ and promising future delivery at $T > t$.

By definition, if $\mathbb{E}^h$ is a risk neutral measure, and $\{\mathcal{F}_t\}_{t \geq 0}$ is the filtration generated by the jump process $M$, then the risk-neutral forward price, see [2, 32], is:

$$f(t, T) = \mathbb{E}^h [S(T)|\mathcal{F}_t].$$

Proposition 5. If the current electricity spot price is $S_t$, and $h$ is a function that defines the equivalent probability measure given in (4.3.2), then the risk neutral price of a forward contract is:

$$f(t, T) = (S_t)^{e^{-\kappa(T-t)}} \exp \left( \int_{t}^{T} \int_{-\infty}^{+\infty} (e^{\sigma x e^{-\kappa(T-u)} + \log(h(x))} + x \sigma e^{-\kappa(T-u)} - h(x)) k(x) dx du \right) \times \exp \left( \int_{t}^{T} \kappa \mu e^{-\kappa(T-u)} du \right).$$ (4.3.5)

Proof. Recall from (4.2.5) that:

$$S_t = (S_s)^{e^{-\kappa(t-s)}} \exp \left( \int_{s}^{t} \kappa \mu e^{-\kappa(t-u)} du + \int_{s}^{t} \sigma e^{-\kappa(t-u)} dM_u \right).$$
By application of Bayes’ rule, see [61], we can write:

\[ f(t, T) = \mathbb{E} \left[ \frac{\Lambda_T}{\Lambda_t} S(T) | \mathcal{F}_t \right] \]

\[ = (S_t)^{e^{-\kappa(T-t)}} \mathbb{E} \left[ \exp \left( - \int_t^T \int_{-\infty}^{+\infty} (h(x) - 1) \bar{k}(x) dx ds \right. \right. \]

\[ + \int_t^T \int_{-\infty}^{+\infty} \log(h(x)) \gamma(dx, ds) + \int_t^T \kappa \mu e^{-\kappa(T-u)} du + \int_t^T \sigma \bar{x} e^{-\kappa(T-u)} dM_u \bigg| \mathcal{F}_t \bigg] \]

\[ = (S_t)^{e^{-\kappa(T-t)}} \exp \left( - \int_t^T \int_{-\infty}^{+\infty} (h(x) - 1) \bar{k}(x) dx ds + \int_t^T \kappa \mu e^{-\kappa(T-u)} du \right) \times \]

\[ \mathbb{E} \left[ \exp \left( \int_t^T \int_{-\infty}^{+\infty} \log(h(x)) \gamma(dx, ds) + \int_t^T \sigma \bar{x} e^{-\kappa(T-u)} dM_u \bigg| \mathcal{F}_t \bigg] \right] \]

\[ = (S_t)^{e^{-\kappa(T-t)}} \exp \left( - \int_t^T \int_{-\infty}^{+\infty} (h(x) - 1 + \sigma \bar{x} e^{-\kappa(T-u)}) \bar{k}(x) dx du + \int_t^T \kappa \mu e^{-\kappa(T-u)} du \right) \times \]

\[ \mathbb{E} \left[ \exp \left( \int_t^T \left( \int_{-\infty}^{+\infty} \log(h(x)) + \sigma \bar{x} e^{-\kappa(T-u)} \right) \gamma(dx, du) \right| \mathcal{F}_t \right] . \]

(4.3.6)

Write

\[ H_t = \int_t^T \int_{\mathbb{R}} \left( \sigma \bar{x} e^{-\kappa(T-u)} + \log(h(x)) \right) \gamma(dx, du). \]

Now

\[ \Delta H_t = H_t - H_{t^-} = - \int_{\mathbb{R}} \left( \sigma \bar{x} e^{-\kappa(T-t)} + \log(h(x)) \right) \gamma(dx, dt). \]

Using the Itô differentiation rule for jump processes,

\[ de^{H_t} = e^{H_{t^-}} \int_{\mathbb{R}} \left( e^{\sigma \bar{x} e^{-\kappa(T-t)} + \log(h(x))} - 1 \right) \gamma(dx, dt). \]

Writing the integral form and using Fubini’s theorem:

\[ e^{H_t} = 1 + \int_t^T e^{H_{u^-}} \int_{\mathbb{R}} \left( 1 - e^{\sigma \bar{x} e^{-\kappa(T-u)} + \log(h(x))} \right) (\gamma(dx, du) - \bar{k}(x) dx du) \]

\[ + \int_t^T e^{H_{u^-}} \int_{\mathbb{R}} \left( e^{\sigma \bar{x} e^{-\kappa(T-u)} + \log(h(x))} - 1 \right) \bar{k}(x) dx du. \]

Since the first integral is a \( \mathbb{P} \)-martingale,

\[ \mathbb{E}[e^{H_t} | \mathcal{F}_t] = \int_t^T \mathbb{E}[e^{H_{u^-}} | \mathcal{F}_t] \int_{\mathbb{R}} \left( e^{\sigma \bar{x} e^{-\kappa(T-u)} + \log(h(x))} - 1 \right) \bar{k}(x) dx du \]
CHAPTER 4. POWER PRICING USING COMPENSATED JUMP PROCESSES

and by solving the ordinary differential equation,

\[ \mathbb{E}[e^{H_t}|\mathcal{F}_t] = \exp \left( \int_t^T \int_{\mathbb{R}} \left( e^{\sigma xe^{-\kappa(T-u)}+\log(h(x))} - 1 \right) \bar{k}(x)dx du \right). \]

So, the risk-neutral forward price is:

\[ f(t,T) = (S_t)e^{-\kappa(T-t)} \exp \left( -\int_t^T \int_{-\infty}^{+\infty} (h(x)-1+x\sigma e^{-\kappa(T-u)})\bar{k}(x)dxdu + \int_t^T \kappa\mu e^{-\kappa(T-u)}du \right) \times \exp \left( \int_t^T \int_{\mathbb{R}} \left( e^{\sigma xe^{-\kappa(T-u)}+\log(h(x))} - 1 \right) \bar{k}(x)dx du \right). \] (4.3.7)

4.3.3 The forward price dynamics

We shall now calculate the dynamics for the forward contract dynamics, the later are significant to determine the price of calls on forwards, swaps, as well as in forecasting the forward curve.

**Proposition 6.** If the price of the forward contract is given by (4.3.5), then the forward dynamics are:

\[ \frac{df(t,T)}{f(t-,T)} = \int_{\mathbb{R}} \left( e^{\sigma xe^{-\kappa(T-u)}+\log(h(x))} - 1 \right) \gamma(dx,dt) - h(x)\bar{k}(x)dxdt. \] (4.3.8)

**Proof.** Write:

\[ D_t = \exp \left( -\int_t^T \int_{-\infty}^{+\infty} (h(x)-1+x\sigma e^{-\kappa(T-u)})\bar{k}(x)dxdu + \int_t^T \kappa\mu e^{-\kappa(T-u)}du \right) \times \exp \left( \int_t^T \int_{\mathbb{R}} \left( e^{\sigma xe^{-\kappa(T-u)}+\log(h(x))} - 1 \right) \bar{k}(x)dx du \right). \]

Then the forward price is:

\[ f(t,T) = \exp(X_t e^{-\kappa(T-t)})D_t. \]

Since \( D \) is a deterministic quantity,

\[ df(t,T) = d\left( \exp(X_t e^{-\kappa(T-t)}) \right) D_t + \exp(X_t e^{-\kappa(T-t)})dD_t, \]
CHAPTER 4. POWER PRICING USING COMPENSATED JUMP

PROCESSES

By Itô’s differentiation:

\[ d\left( \exp(X_t e^{-\kappa(T-t)}) \right) = d\left( X_t e^{-\kappa(T-t)} \right) \exp(X_t e^{-\kappa(T-t)}) + \exp(X_t e^{-\kappa(T-t)}) - \exp(X_t e^{-\kappa(T-t)}) - \exp(X_t e^{-\kappa(T-t)}) \Delta(X_t e^{-\kappa(T-t)}) . \]

On the other hand,

\[ d\left( X_t e^{-\kappa(T-t)} \right) = (\kappa - X_t) dt + \sigma dM_t e^{-\kappa(T-t)} + \kappa X_t e^{-\kappa(T-t)} dt \]

\[ = \left( \kappa \mu dt + \int_R \sigma x (\gamma - \bar{\gamma})(dx, dt) \right) e^{-\kappa(T-t)} \]

Also

\[ \Delta(X_t e^{-\kappa(T-t)}) = \left( \int_R \sigma x \gamma(dx, dt) \right) e^{-\kappa(T-t)} \]

and

\[ \exp(X_t e^{-\kappa(T-t)}) - \exp(X_t e^{-\kappa(T-t)}) = \exp(X_t e^{-\kappa(T-t)}) \int_R (e^{\sigma x e^{-\kappa(T-t)}} - 1) \gamma(dx, dt) . \]

So

\[ d\left( \exp(X_t e^{-\kappa(T-t)}) \right) = \exp(X_t e^{-\kappa(T-t)}) \left[ \left( \kappa \mu dt + \int_R \sigma x (\gamma - \bar{\gamma})(dx, dt) \right) e^{-\kappa(T-t)} \right. \]
\[ + \int_R (e^{\sigma x e^{-\kappa(T-t)}} - 1) \gamma(dx, dt) - \left( \int_R \sigma x \gamma(dx, dt) \right) e^{-\kappa(T-t)} \]
\[ = \exp(X_t e^{-\kappa(T-t)}) \left[ \left( \kappa \mu dt - \int_R \sigma x \bar{k}(x) dxdt \right) e^{-\kappa(T-t)} \right. \]
\[ + \int_R (e^{\sigma x e^{-\kappa(T-t)}} - 1) \gamma(dx, dt) \] .

On the other hand, since \( D_t \) has only deterministic terms,

\[ dD_t = D_t \left[ \int_{-\infty}^{+\infty} \left( h(x) + x \sigma e^{-\kappa(T-t)} - e^{\sigma x e^{-\kappa(T-t)}} + \log(h(x)) \right) \bar{k}(x) dx - \kappa \mu e^{-\kappa(T-t)} \right] dt . \]

So the dynamics of the forward price are:

\[ \frac{df(t,T)}{f(t-,T)} = \int_R (e^{\sigma x e^{-\kappa(T-t)}} - 1) \gamma(dx, dt) + \int_R (-e^{\sigma x e^{-\kappa(T-t)}} + 1) h(x) \bar{k}(x) dx dt \]
\[ = \int_R (e^{\sigma x e^{-\kappa(T-t)}} - 1) \gamma(dx, dt) - h(x) \bar{k}(x) dx dt . \] 

\[ (4.3.9) \]
Remark 9. From (4.3.9) the forward dynamics can be written:

\[
\frac{df(t,T)}{f(t-,T)} = \int_\mathbb{R} (e^{\sigma xe^{-\kappa(T-t)}} - 1)\nu^h(dx,dt) = \int_\mathbb{R} \phi(t,T,x)\nu^h(dx,dt).
\]

(4.3.10)

Here \(\nu^h\) is a compensated random measure under \(P^h\) and \(\phi\) is the volatility term with respect to the compensated jump measure. These latter dynamics agree with the definition for the forward price as a martingale under the the risk-neutral measure \(P^h\), see [4, 2].

### 4.3.4 Electricity swap price

In electricity markets, similarly to other energy markets, the forward contract is usually written for a commodity delivery over a period of time, see [32, 65, 2]. Consider an electricity swap with delivery period \([T_1, T_2]\) for \(T_1 < T_2\), and let \(F(t, T_1, T_2)\) be the swap price for \(t \leq T_1\). As a result of the non arbitrage condition (see [3, 2]), if \(f(t,u)\) is the price of a forward contract on delivery at time \(u \in [T_1, T_2]\), the relation between the Swap and forward prices satisfies, see [2]:

\[
F(t, T_1, T_2) = \int_{T_1}^{T_2} \tilde{w}(u, T_1, T_2)f(t,u)du
\]

(4.3.11)

where

\[
\tilde{w}(u, s, t) = \frac{w(u)}{\int_s^t w(\tau)d\tau} \quad \text{for } 0 \leq u \leq s < t.
\]

That means that holding a swap contract is the same in terms of payoff as holding a weighted continuous stream of forward contracts. If the swap settles at maturity, then \(w(u) = 1\). Otherwise, \(w(u) = \exp(-ru)\) when the settlement is made continuously over the period \([T_1, T_2]\). So if we have a model for the forward price, determining the swap price is by calculating the deterministic integral (4.3.11). Now assume that the swap price has the following dynamics, see [2, 12]:

\[
dF(t, T_1, T_2) = \int_\mathbb{R} \Phi(t, T_1, T_2, x)\nu^h(dx, dt).
\]

(4.3.12)

Suppose we wish to express the swap volatility \(\Phi\) in terms of the forward parameters.
CHAPTER 4. POWER PRICING USING COMPENSATED JUMP PROCESSES

Proposition 7. The swap volatility can be expressed in terms of the forward volatility as:

\[ \Phi(t, T_1, T_2, x) = \int_{T_1}^{T_2} \tilde{w}(u, T_1, T_2) f(t^-, u) \phi(t, u, x) du. \]  

(4.3.13)

Proof. From (4.3.9) we write for \( u \in [0, t] \):

\[ f(t, u) = f(0, u) + \int_0^t f(s^-, u) \int_{\mathbb{R}} \phi(t, u, x) \nu^h(dx, ds) du \]

So by replacing in (4.3.11) and applying stochastic Fubini theorem, see Appendix A.6:

\[ F(t, T_1, T_2) = \int_{T_1}^{T_2} \tilde{w}(u, T_1, T_2) f(t, u) du \]
\[ = \int_{T_1}^{T_2} \tilde{w}(u, T_1, T_2) f(0, u) du + \int_{T_1}^{T_2} \int_{0}^{t} \int_{\mathbb{R}} \tilde{w}(u, T_1, T_2) f(s^-, u) \phi(t, u, x) \nu^h(dx, ds) du \]
\[ = \int_{T_1}^{T_2} \tilde{w}(u, T_1, T_2) f(0, u) du + \int_{0}^{t} \int_{T_1}^{T_2} \tilde{w}(u, T_1, T_2) f(s^-, u) \phi(t, u, x) \nu^h(dx, ds) du \]

So

\[ dF(t, T_1, T_2) = \int_{T_1}^{T_2} \int_{\mathbb{R}} \tilde{w}(u, T_1, T_2) f(t^-, u) \phi(t, u, x) \nu^h(dx, dt) du \]
\[ = \int_{\mathbb{R}} \left( \int_{T_1}^{T_2} \tilde{w}(u, T_1, T_2) f(t^-, u) \phi(t, u, x) du \right) \nu^h(dx, dt). \]

\[ \square \]

4.4 Conclusion

In this chapter a new model for power pricing is suggested and explored, based on a mean reverting jump process where the jumps are modelled by a compensated jump process to gain the martingale property.

A mathematical setting is presented, a risk neutral measure is defined and an electricity forward contract is priced. The dynamics of forward are then used to derive the swap option price.
Chapter 5

Power pricing using filtering theory

5.1 Introduction

The deregulation of electricity markets during the later decades of the last century in different markets around the world has led to a considerable increase in the risk involved in power trading. Because it is extremely expensive, and usually difficult to store power, electricity as an underlying does not have elasticity in the offer price. This can cause extreme spot price fluctuations and lead to critical financial problems for companies that buy power in the wholesale market.

Power market agents must hedge against highly volatile price risk, even if purely speculative strategies are still widely used among investors. They often resort to financial derivatives such as forwards, options and swaps. There are two main approaches to modelling forward prices, depending on market data availability. To price forwards, a Heath-Jarrow-Morton approach is used in [42] where only the contracts traded on the market are modelled. This approach starts by assuming a model for the forward price dynamics, and then determines the price of related derivatives such as swaps and swaptions, see [3, 4]. Of course, if the forward curve is known,
CHAPTER 5. POWER PRICING USING FILTERING THEORY

one can deduce the spot price since the two quantities coincide when time converges to maturity. This approach is mainly used when a relation between the spot and forward price is not needed; it allows one to neglect the complexity of the spot price dynamics.

A second approach is based on first modelling the spot price dynamics, and then establishing the link between the forward and spot price. Unlike stocks and other commodities, electricity spot prices feature a variation around production cost so, in the short term, price spreads are observed, while in the long run the price follows the cost of production. This explains the attractiveness of the mean reverting model in pricing electricity. The paper by Schwartz [60] is one of the earliest works that contributes to modelling energy and power prices. He uses a mean reverting Ornstein-Uhlenbeck process to describe power prices. In Lucia and Schwartz [50] the seasonal effect was taken into account, (see [54] and [11]). Using spot price models allows one to price derivatives written on both forward and spot prices, so for parameter calibration the lack of data on forward prices is no more a limitation. Usually, following this approach involves introducing a convenience yield, or risk premium, to determine a risk neutral measure in the presence of an incomplete market, see [32].

In this chapter we shall introduce a new model for commodity prices, in particular for the price of power. We model the behaviour of the market using mean reverting dynamics which include a hidden Markov switching drift. This models long term changes in power price trends due to systemic changes in the industry. The off-peak prices seen as spikes are modelled with a Markov regime switching coefficient, and the seasonality effect is included using a bias from deterministic factors.

5.2 Presenting the model

Suppose all processes are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The short term randomness of the logarithm of the price of power is modelled using the following
regime switching, mean reverting diffusion process $X_t$, having dynamics:

$$dX_u = \kappa(\langle \mu, Z_u \rangle - X_u)du + \sigma dW_u. \quad (5.2.1)$$

Here $\kappa$ is the speed of return of the process to the long term mean. $Z$ is a finite state, non-observable, continuous time Markov chain. Without loss of generality we may assume that $Z$ takes values in $\{e_1, \ldots, e_N\}$, the canonical basis for the Euclidian space $\mathbb{R}^N$. Then $\mu = (\mu_1, \ldots, \mu_N) \in \mathbb{R}^N$ is a parameter vector such that the long term mean for the process $X$ switches depending on the value of $Z$, that is $\mu_u = \langle \mu, Z_u \rangle$. $W$ is a 1-dimensional standard Brownian motion.

Suppose $A = (a_{ji}, 1 \leq i, j \leq N)$ is the transition rate matrix for the Markov chain $Z$, then, from [20], $Z$ has dynamics:

$$Z_t = Z_0 + \int_0^t AZ_s ds + \int_0^t dM_s \quad (5.2.2)$$

where $M$ is an $\mathbb{R}^N$-martingale.

Taking $X_0 = 0$, the solution of the above stochastic differential equation is:

$$X_t = \int_0^t e^{\kappa(u-t)}(\kappa \mu, Z_u)du + \int_0^t \sigma e^{\kappa(u-t)}dW_u. \quad (5.2.3)$$

Suppose $J$ is a second finite state continuous time Markov chain taking values in $\{e_1, e_2, \ldots, e_M\}$ and independent of $X$.

If $K = (K_{ij}, 1 \leq i, j \leq N)$ is the transition rate matrix for $J$, then $J$ has dynamics:

$$J_t = J_0 + \int_0^t KJs ds + \int_0^t dV_s \quad (5.2.4)$$

where $K$ is the transition rate matrix of $J$ and $V$ is an $\mathbb{R}^M$-martingale.

We assume that the jumps in the spot price are observed and the jump sizes take values from a set $\{\alpha_1, \alpha_2, \ldots, \alpha_M\}$. Write $\alpha = (\alpha_1, \ldots, \alpha_M) \in \mathbb{R}^M$.

$f_t$ is a deterministic function which models the yearly, daily and weekly seasonality of the spot price, respectively. We suppose it is given by:

$$f_t = \epsilon + \beta t + dD_{day} + \delta_y \sin\left(\frac{2\pi}{365 \times 48} t + \varphi_y\right) + \delta_d \sin\left(\frac{2\pi}{48} t + \varphi_d\right), \quad (5.2.5)$$
where $\epsilon, \beta, d$ are constant parameters. Here, unlike in Chapter 3, we assume that the electricity price data is half hourly, so we introduce two sinusoidal functions with different amplitudes $\delta_y, \delta_d$ and phases $\varphi_y, \varphi_d$ for the yearly and daily periodicity respectively. We then assume the spot price of the commodity at time $t \geq 0$ is:

$$S_t = S_0 f_t \exp(X_t) \langle \alpha, J_t \rangle. \quad (5.2.6)$$

### 5.3 Forward pricing

From (5.2.3) and (5.2.6):

$$S_t = S_0 f_t \exp \left( \int_0^t \kappa e^{\kappa(u-t)} \langle \mu, Z_u \rangle du + \int_0^t \sigma e^{\kappa(u-t)} dW_u \right) \langle \alpha, J_t \rangle. \quad (5.3.1)$$

Here, the price jumps are modelled by the factor $\langle \alpha, J_t \rangle$. We think this is better than jump diffusion dynamics where the return to base load after a spike is difficult to model.

On a series of price data, one can pinpoint when the price jumps occur, consequently, the process defined by:

$$\log \left( \frac{S_t}{\langle \alpha, J_t \rangle} \right) = \log(S_0) + \log(f_t) + X_t. \quad (5.3.2)$$

is observed. However, $S_0$ is observed, so the process $y$ is observed, where

$$y_t = e^{\kappa t} (\ln(f_t) + X_t)$$

$$= e^{\kappa t} \ln(f_t) + \kappa \int_0^t e^{\kappa u} \langle \mu, Z_u \rangle du + \sigma \int_0^t e^{\kappa u} dW_u$$

$$= \int_0^t a_u du + \int_0^t b_u\langle \mu, Z_u \rangle du + \sigma \int_0^t e^{\kappa u} dW_u, \quad (5.3.3)$$

where

$$a_u = \frac{d}{du}(e^{\kappa u} \ln(f_u))$$

and

$$b_u = \kappa e^{\kappa u}.$$ 

Consider the filtration $\{\mathcal{Y}_t, t \geq 0\}$ where:

$$\mathcal{Y}_t = \sigma \{y_u : 0 \leq u \leq t\}. \quad (5.3.4)$$
We write $\mathcal{F}^W$, $\mathcal{F}^J$ and $\mathcal{F}^Z$ for the right continuous, complete filtrations generated by the Brownian motion and the Markov chains $J$ and $Z$ respectively. Suppose $\{\mathcal{F}_t\}$ is the filtration generated by the process $y$ and the Markov chain $J$, so $\mathcal{F} = \mathcal{Y} \vee \mathcal{F}^J$ represents the total observed information from the spot price.

We shall calculate the forward price for a contract entered at time $s$ for delivery at time $t \geq s$. It is costless to enter a forward contract, so if $f(s,t)$ is the forward price negotiated at time $s$ to be paid for the commodity at time $t$, then by the fundamental pricing theorem, see [16]:

$$e^{-r(t-s)}\mathbb{E}^\lambda[f(s,t) - S(t)|\mathcal{F}_s] = 0,$$

where $r$ is the risk free interest rate. The pricing is performed in an arbitrage-free framework, so if $\mathbb{E}^\lambda$ is a risk neutral measure, then the forward price is given by:

$$f(s,t) = \mathbb{E}^\lambda[S_t|\mathcal{F}_s] = \mathbb{E}^\lambda[S_t|\mathcal{Y}_s \vee \mathcal{F}^J_s].$$

Note that $f(t,t) = S_t$ as the forward price converges to the spot price.

### 5.3.1 Risk premium and change of measure

We first model the price of risk in our dynamics. Pricing in the power market is performed under the no arbitrage assumption. From Harrison and Pliska [40], if a market is complete, any contingent claim can be replicated with a self financing strategy and there exists a unique risk-neutral martingale measure. However in the electricity market, and because of the non storability of the underlying, a derivative can not be replicated with a portfolio consisting of the underlying and a financing debt account, so any equivalent risk-neutral measure is not unique.

Some authors model the power market assuming it is already under the risk neutral measure and perform pricing under the historical probability measure; this approach assumes that we are calibrating the model through implied parameters.

A more suitable way to proceed is to use the concept of risk premium as in [32, 2, 64], which is defined as the reward to the investor for investing in the risky asset instead
of a risk-free one. The introduction of the risk premium in the spot model allows us to price electricity derivatives consistently with the prices of forward contracts quoted in the market. The calibration is based on the initial yield curve. We shall use the approach of Lucia and Schwartz [50], where the risk premium is captured in the model through an adjustment in the drift of the asset return, known as the market price of risk. The concept of risk premium can be seen as an adjustment of the price drift to compensate bearing the risk of buying the asset. So we chose a martingale measure which is consistent with the forward prices quoted on the market. Let \( \hat{\lambda} \) be the market price of risk, where:

\[
\hat{\lambda} = \lambda \frac{\sigma}{\kappa}.
\]

(5.3.4)

Here \( \lambda \) is the price of risk per unit linked to the process \( X \). From the above expression \( \hat{\lambda} \) is positively correlated to the spot volatility and negatively to the speed of return of the process \( X \) to the long term mean.

Consider a new probability measure \( \mathbb{P}^{\lambda} \) defined by:

\[
\frac{d\mathbb{P}^{\lambda}}{d\mathbb{P}}|_{\mathcal{F}_t} = \exp \left( - \int_0^t \frac{\kappa \hat{\lambda}}{\sigma} dW_u - \frac{1}{2} \int_0^t \frac{(\kappa \hat{\lambda})^2}{\sigma^2} du \right)
\]

\[
= \exp \left( - \int_0^t \lambda dW_u - \int_0^t \frac{\lambda^2}{2} du \right).
\]

(5.3.5)

By Girsanov’s theorem, under \( \mathbb{P}^{\lambda} \), the dynamics of the process \( X \) is then:

\[
dX_t = \kappa(\langle \hat{\mu}, Z_u \rangle - X_u)du + \sigma dW^\lambda_u.
\]

(5.3.6)

where

\[
\forall j \in \{1, \ldots, N\} : \hat{\mu}_j = \mu_j - \hat{\lambda}
\]

and

\( W^\lambda_t = W_t + \hat{\lambda}t \) is a Brownian motion under \( \mathbb{P}^{\lambda} \).

Also, under \( \mathbb{P}^{\lambda} \):

\[
y_t = \int_0^t a_u du + \int_0^t b_u \langle \hat{\mu}, Z_u \rangle du + \sigma \int_0^t e^{\kappa u} dW^\lambda_u
\]

(5.3.7)
Under $\mathbb{P}^\lambda$, the spot price is given by:

$$S_t = S_0 f_t \exp \left( \int_0^t \kappa e^{\kappa(u-t)} \langle \hat{\mu}; Z_u \rangle du + \int_0^t \sigma e^{\kappa(u-t)} dW_u^\lambda \right) \langle \alpha, J_t \rangle,$$

and the forward price is:

$$f(s, t) = \mathbb{E}^\lambda [S_t | \mathcal{Y}_s \lor \mathcal{F}_s^J].$$

Write

$$A_t := S_0 f_t \exp \left( e^{-\kappa t} \int_0^t \sigma e^{\kappa u} dW_u^\lambda \right).$$

$$B_t := \exp \left( e^{-\kappa t} \int_0^t e^{\kappa u} \langle \kappa \hat{\mu}; Z_u \rangle du \right).$$

$$C_t := \langle \alpha, J_t \rangle.$$

The forward price is then:

$$f(s, t) = \mathbb{E}^\lambda [S_t | \mathcal{Y}_s \lor \mathcal{F}_s^J] = \mathbb{E}^\lambda [A_t B_t C_t | \mathcal{Y}_s \lor \mathcal{F}_s^J].$$

and because $W$, $Z$ and $J$ are independent,

$$f(s, t) = \mathbb{E}^\lambda [A_t | \mathcal{Y}_s] \mathbb{E}^\lambda [B_t | \mathcal{Y}_s] \mathbb{E}^\lambda [C_t | \mathcal{Y}_s \lor \mathcal{F}_s^J].$$

These three expectations will now be calculated.

**Notation 3.** To simplify the notation, we shall now write $\mathbb{E}$ for $\mathbb{E}^\lambda$.

The final expectation is straightforward from the dynamics of the Markov chain $J$, and given in the following proposition:

**Proposition 8.** let $J$ be a continuous time Markov process, with dynamics:

$$J_t = J_0 + \int_0^t K J_s ds + \int_0^t dV_s.$$ 

where $K$ is a transition rate matrix and $V$ is an $\mathbb{R}^M$ martingale vector. Then:

$$\mathbb{E}[C_t | \mathcal{Y}_s \lor \mathcal{F}_s^J] = \mathbb{E}[\langle \alpha, J_t \rangle | \mathcal{Y}_s \lor \mathcal{F}_s^J] = \langle \alpha, e^{K(t-s)} J_s \rangle.$$  

(5.3.13)
Proof. It is straightforward by introducing the change of variable: $Y_t = e^{-Kt} J_t$.

Using Itô differentiation, we obtain:

$$dY_t = e^{-Kt} dV_t.$$ 

so

$$\mathbb{E}[Y_t | \mathcal{Y}_s \vee \mathcal{F}^J_s] = Y_s + \mathbb{E}\left[\int_s^t e^{-Ku} dV_u | \mathcal{Y}_s \vee \mathcal{F}^J_s\right]$$

and because $V$ is an $\mathbb{R}^M$ martingale vector, the integral term is a martingale and:

$$\mathbb{E}[J_t | \mathcal{Y}_s \vee \mathcal{F}^J_s] = e^{K(t-s)} J_s. \quad (5.3.14)$$

For the first term we have:

$$\mathbb{E}[A_t | \mathcal{Y}_s \vee \mathcal{F}^J_s] = S_0 f_t \mathbb{E}\left[e^{-\kappa t} \int_0^t \sigma e^{\kappa u} dW^\lambda_u | \mathcal{Y}_s \vee \mathcal{F}^J_s\right]$$

$$= S_0 f_t \mathbb{E}\left[e^{-\kappa t} \int_0^s \sigma e^{\kappa u} dW^\lambda_u\right] \mathbb{E}\left[e^{-\kappa t} \int_s^t \sigma e^{\kappa u} dW^\lambda_u | \mathcal{Y}_s \vee \mathcal{F}^J_s \vee \mathcal{F}^W_s\right]$$

$$= S_0 f_t \exp\left(\frac{1}{2} e^{-2\kappa t} \int_0^t \sigma^2 e^{2\kappa u} du\right) \mathbb{E}\left[e^{-\kappa t} \int_0^s \sigma e^{\kappa u} dW^\lambda_u | \mathcal{Y}_s \vee \mathcal{F}^J_s\right].$$

Consequently, we wish to calculate:

$$\mathbb{E}\left[e^{-\kappa t} \int_0^s \sigma e^{\kappa u} dW^\lambda_u | \mathcal{Y}_s \vee \mathcal{F}^J_s\right] = \mathbb{E}\left[e^{-\kappa t} \int_0^s \sigma e^{\kappa u} dW^\lambda_u | \mathcal{Y}_s\right].$$

However, the Brownian motion $W^\lambda$ is not observable, so from (5.3.2):

$$\mathbb{E}\left[e^{-\kappa t} \int_0^s \sigma e^{\kappa u} dW^\lambda_u | \mathcal{Y}_s\right] = \mathbb{E}\left[e^{-\kappa t} \left(y_s - \int_0^s a_u du - \int_0^s b_u (\hat{\mu}, Z_u) du\right) | \mathcal{Y}_s\right]$$

$$= \exp\left(-\kappa e^{-\kappa t} \int_0^s b_u (\hat{\mu}, Z_u) du\right) \mathbb{E}\left[e^{-\kappa t} \left(y_s - \int_0^s a_u du\right) | \mathcal{Y}_s\right]$$

$$= \exp\left(-\kappa e^{-\kappa t} \int_0^s e^{\kappa u} (\hat{\mu}, Z_u) du\right) \mathbb{E}\left[e^{-\kappa t} \left(y_s - \int_0^s a_u du\right) | \mathcal{Y}_s\right].$$

Consequently, we wish to determine:

$$\mathbb{E}\left[e^{-\kappa e^{-\kappa t} \int_0^s e^{\kappa u} (\hat{\mu}, Z_u) du} | \mathcal{Y}_s\right]. \quad (5.3.15)$$
Write
\[ \alpha = -\kappa e^{-\kappa t}. \]

We shall calculate:
\[ \mathbb{E}\left[ \exp\left( \alpha \int_0^s e^{\kappa u} \langle \hat{\mu}, Z_u \rangle du \right) \mid \mathcal{Y}_s \right]. \]

The process \( Z \) is not observable, so we first develop a filter for \( Z \) in the following section:

### 5.3.2 Recursive filter for hidden Markov process

Below we determine the filter estimate of the process \( Z \), conditioned on the observations \( \mathcal{Y} \), as introduced in [19].

**Lemma 13.** Let \( Z \) be the Markov chain defined by (5.2.2) and \( \mathcal{Y} \) the filtration defined in (5.3.3). Then we have:
\[ \mathbb{E}[Z_t | \mathcal{Y}_t] = \frac{q_t}{\langle q_t, 1 \rangle}. \]

where \( q_t \) is the vector solution of the following stochastic differential equation:
\[ q_t = q_0 + \int_0^t \frac{a_u}{\sigma^2 e^{2\kappa u}} q_u dy_u + \int_0^t \frac{\kappa}{\sigma^2 e^{\kappa u}} \text{diag}(\hat{\mu}) q_u dy_u + \int_0^t A q_u du. \quad (5.3.16) \]

**Proof.** Consider a probability measure \( \mathbb{P}^\lambda \) under which:
\[ \Phi_t := \int_0^t \frac{dy_u}{\sigma e^{\kappa u}} = \frac{1}{\sigma} \int_0^t e^{-\kappa u} dy_u \]

is a standard Brownian motion.

Write
\[ \theta_u = \frac{a_u + \kappa e^{\kappa u} \langle \hat{\mu}, Z_u \rangle}{\sigma e^{\kappa u}}, \quad (5.3.18) \]

and consider the process \( \Lambda \) where :
\[ \Lambda_t = \exp \left( \int_0^t \theta_u d\Phi_u - \frac{1}{2} \int_0^t \theta_u^2 du \right). \quad (5.3.19) \]

Define the probability \( \mathbb{P}^\lambda \) by setting:
\[ \frac{d\mathbb{P}^\lambda}{d\mathbb{P}} \mid_{\mathcal{F}_t} = \Lambda_t. \quad (5.3.20) \]
From Girsanov’s theorem, under $P^\lambda$

$$W_t^\lambda = \Phi_t - \int_0^t \theta_u du$$

is a Brownian motion. That is

$$dW_t^\lambda = \frac{dy_t}{\sigma e^{\kappa t}} - \theta_t dt,$$ (5.3.21)

so under $\mathbb{P}$:

$$dy_t = a_t dt + b_t \langle \hat{\mu}, Z_t \rangle dt + \sigma e^{\kappa t} dW_t^\lambda$$

That is, $y$ has the required dynamics (5.3.2) under $P^\lambda$.

We wish to calculate $E[Z_t | \mathcal{Y}_t]$. By Bayes’ rule, (see [20]) :

$$E[Z_t | \mathcal{Y}_t] = \frac{E[\Lambda_t Z_t | \mathcal{Y}_t]}{E[\Lambda_t | \mathcal{Y}_t]}.$$ 

Write :

$$q_t = E[\Lambda_t Z_t | \mathcal{Y}_t].$$ (5.3.22)

We first calculate $\Lambda_t Z_t$:

$$d(\Lambda_t Z_t) = \Lambda_t (dZ_t) + (d\Lambda_t) Z_t.$$ 

Now

$$d\Lambda_t = \theta_t \Lambda_t d\Phi_t$$

$$= \frac{\theta_t \Lambda_t}{\sigma e^{\kappa t}} dy_t.$$ (5.3.23)

and

$$dZ_t = AZ_t dt + dM_t \in \mathbb{R}^N.$$ 

Therefore,

$$d(\Lambda_t Z_t) = A\Lambda_t Z_t dt + \Lambda_t dM_t + \theta_t \Lambda_t Z_t d\Phi_t.$$ 

So, in integral form:

$$\Lambda_t Z_t = Z_0 + \int_0^t A\Lambda_u Z_u du + \int_0^t \Lambda_u dM_u + \int_0^t \theta_u \Lambda_u Z_u d\Phi_u.$$
Consequently, by the stochastic Fubini theorem, see [66]:

\[ q_t = \bar{E}[\Lambda_t Z_t | \mathcal{Y}_t] = q_0 + \int_0^t A \bar{E}[\Lambda_u Z_u | \mathcal{Y}_u] du + \int_0^t \theta_u \bar{E}[\Lambda_u Z_u | \mathcal{Y}_u] d\Phi_u \]

\[ = q_0 + \int_0^t A q_u du + \int_0^t \theta_u q_u d\Phi_u \]

\[ = q_0 + \int_0^t \frac{a_u}{\sigma^2 e^{2\kappa u}} q_u dy_u + \int_0^t \frac{\kappa}{\sigma^2 e^{2\kappa u}} \text{diag}(\hat{\mu}) q_u dy_u + \int_0^t A q_u du. \]

(5.3.24)

Note that

\[ \langle q_t, 1 \rangle = \langle \bar{E}[\Lambda_t Z_t | \mathcal{Y}_t], 1 \rangle = \bar{E}[\Lambda_t \langle Z_t, 1 \rangle | \mathcal{Y}_t] = \bar{E}[\Lambda_t | \mathcal{Y}_t]. \]

So, if \( q \) is the solution of (5.3.24), then

\[ \bar{E}[\Lambda_t | \mathcal{Y}_t] = \langle q_t, 1 \rangle. \] (5.3.25)

and

\[ \bar{E}[Z_t | \mathcal{Y}_t] = \frac{q_t}{\langle q_t, 1 \rangle}. \]

To obtain the forward price we shall now calculate:

\[ \bar{E}[\exp(\alpha \int_0^t e^{\kappa u} \langle \hat{\mu}, Z_u \rangle du) | \mathcal{Y}_t] \]

for any constant \( \alpha \in \mathbb{R} \).

**Proposition 9.** Using the previous notation, we have the following equality:

\[ \bar{E}[\exp(\alpha \int_0^t e^{\kappa u} \langle \hat{\mu}, Z_u \rangle du) | \mathcal{Y}_t] = \frac{\langle \rho_t, 1 \rangle}{\langle q_t, 1 \rangle} \]

where \( q \) is the solution of (5.3.16) and \( \rho \) is the vector solution of the following stochastic differential equation:

\[ \rho_t = \rho_0 + \int_0^t (A + \alpha e^{\kappa u} \text{diag}(\hat{\mu})) \rho_u du + \int_0^t \left( \frac{a_u}{\sigma^2 e^{2\kappa u}} I_N + \frac{\kappa}{\sigma^2 e^{2\kappa u}} \text{diag}(\hat{\mu}) \right) \rho_u dy_u. \]

(5.3.26)
Proof. Again consider $\overline{P}$ under which:

$$\Phi_t := \int_0^t \frac{dy_u}{\sigma e^{\kappa u}} = \frac{1}{\sigma} \int_0^t e^{-\kappa u} dy_u$$

is a Brownian motion. With $\theta$ as defined in (5.3.18), consider $\Lambda_t$ with:

$$d\Lambda_t = \frac{\theta_t \Lambda_t dy_u}{\sigma e^{\kappa u}}. \quad (5.3.27)$$

and again, define $P$ by:

$$dP = \frac{d\overline{P}}{\Lambda_t}. \quad (5.3.28)$$

Write

$$Q_t = Q_t(\alpha) = \exp \left( \alpha \int_0^t e^{\kappa u} \langle \hat{\mu}, Z_u \rangle \, du \right). \quad (5.3.29)$$

Then

$$dQ_t = Q_t(\alpha e^{\kappa t} \langle \hat{\mu}, Z_t \rangle) \, dt.$$ 

Consider,

$$R_t = Q_t Z_t. \quad (5.3.30)$$

Then

$$dR_t = d(Q_t Z_t) = Q_t dZ_t + dQ_t Z_t$$

$$= Q_t(AZ_t dt + dM_t) + Q_t(\alpha e^{\kappa t} \langle \hat{\mu}, Z_t \rangle)Z_t dt$$

$$= AR_t dt + Q_t dM_t + \alpha e^{\kappa t} \text{diag}(\hat{\mu}) R_t dt.$$ 

Since

$$d\Lambda_t = \frac{\theta_t \Lambda_t dy_u}{\sigma e^{\kappa u}},$$

we have

$$d(\Lambda_t R_t) = \Lambda_t dR_t + d\Lambda_t R_t.$$ 

So

$$\Lambda_t R_t = Z_0 + \int_0^t \Lambda_u AR_u du + \int_0^t \Lambda_u Q_u dM_u + \int_0^t \Lambda_u \alpha e^{\kappa u} \text{diag}(\hat{\mu}) R_u du + \int_0^t R_u \Lambda_u \theta_u \frac{dy_u}{\sigma e^{\kappa u}}. \quad (5.3.31)$$
Now
\[ \theta_u = \frac{a_u + \kappa e^{\kappa u} \langle \hat{\mu}, Z_u \rangle}{\sigma e^{\kappa u}}, \]
so
\[
\Lambda_t R_t = Z_0 + \int_0^t \Lambda_u A R_u du + \int_0^t \Lambda_u Q_u dM_u + \int_0^t \Lambda_u \alpha e^{\kappa u} \text{diag}(\hat{\mu}) R_u du \\
+ \int_0^t \frac{a_u}{\sigma^2 e^{2\kappa u}} R_u \Lambda_u dy_u + \int_0^t \frac{\kappa}{\sigma^2 e^{\kappa u}} \text{diag}(\hat{\mu}) \Lambda_u R_u dy_u.
\]
Write
\[
\rho_t = \mathbb{E}[\Lambda_t R_t | \mathcal{Y}_t] \in \mathbb{R}^M
\]
Then, using again the stochastic Fubini theorem, we have that \( \rho \) is the solution of the following vector stochastic differential equation:
\[
\rho_t = \rho_0 + \int_0^t A \rho_u du + \int_0^t \alpha e^{\kappa u} \text{diag}(\hat{\mu}) \rho_u du + \int_0^t \frac{a_u}{\sigma^2 e^{2\kappa u}} \rho_u dy_u + \int_0^t \frac{\kappa}{\sigma^2 e^{\kappa u}} \text{diag}(\hat{\mu}) \rho_u dy_u. \tag{5.3.32}
\]
Now,
\[
\rho_t = \mathbb{E}[\Lambda_t R_t | \mathcal{Y}_t] \\
= \mathbb{E}[\Lambda_t Q_t Z_t | \mathcal{Y}_t] \\
= \rho_t(\alpha).
\]
so,
\[
\langle \rho_t, 1 \rangle = \mathbb{E}[\Lambda_t Q_t | \mathcal{Y}_t]. \tag{5.3.33}
\]
The quantity we wish to calculate is:
\[
\mathbb{E}[Q_t | \mathcal{Y}_t] = \frac{\mathbb{E}[\Lambda_t Q_t | \mathcal{Y}_t]}{\mathbb{E}[\Lambda_t | \mathcal{Y}_t]}.
\]
Consequently, equations (5.3.25) and (5.3.33) give an expression for:
\[
\mathbb{E}[Q_t | \mathcal{Y}_t] = \mathbb{E}[\exp(\alpha \int_0^t e^{\kappa(u-t)} \kappa(\mu, Z_u) du) | \mathcal{Y}_t] = \frac{\langle \rho_t(\alpha), 1 \rangle}{\langle q_t, 1 \rangle}
\]
for any \( \alpha \in \mathbb{R} \). \(\Box\)
Corollary 3. We can take $\alpha = -\kappa e^{-\kappa t}$ in Proposition 9 and so determine:

$$
E[\exp \left( -\kappa e^{-\kappa t} \int_{0}^{s} e^{\kappa u} \langle \hat{\mu}, Z_{u} \rangle du \right) | Y_{s}] = \frac{\langle \rho_{s}(-\kappa e^{-\kappa t}), 1 \rangle}{\langle q_{s}, 1 \rangle}.
$$

Therefore:

$$
E[A_{t} | Y_{s} \vee F_{s}^{J}] = S_{0} f_{t} \exp \left( \frac{1}{2} e^{-2\kappa t} \int_{s}^{t} \sigma^{2} e^{2\kappa u} du \right) \exp \left( e^{-\kappa t} \left( y_{s} - \int_{0}^{s} a_{u} du \right) \right) \frac{\langle \rho_{s}(-\kappa e^{-\kappa t}), 1 \rangle}{\langle q_{s}, 1 \rangle}.
$$

The calculation of the final term is partly similar. The expectation value is given in the following proposition:

Proposition 10.

$$
E[B_{t} | Y_{s} \vee F_{s}^{J}] = E[\exp \left( \kappa e^{-\kappa t} \int_{0}^{t} e^{\kappa u} \langle \hat{\mu}; Z_{u} \rangle du \right) | Y_{s}] = \langle \Psi(s, t) \frac{\rho_{s}(\kappa e^{-\kappa t})}{\langle q_{s}, 1 \rangle}, 1 \rangle,
$$

where $\Psi(s, t)$ is the solution of the system of ordinary differential equations:

$$
\frac{d\Psi(s, t)}{dt} = (\alpha e^{\kappa t} \text{diag}(\hat{\mu}) + A) \Psi(s, t),
$$

$$
\Psi(s, s) = I_{N_{2}}.
$$

As before, $\rho_{s}(\alpha)$ and $q_{s}$ are, respectively, the solutions of the differential equations:

$$
\rho_{s} = \rho_{0} + \int_{0}^{s} (A + \alpha e^{\kappa u} \text{diag}(\hat{\mu})) \rho_{u} du + \int_{0}^{s} \frac{a_{u}}{\sigma^{2} e^{2\kappa u}} \rho_{u} dy_{u} + \int_{0}^{s} \frac{\kappa}{\sigma^{2} e^{2\kappa u}} \text{diag}(\hat{\mu}) \rho_{u} dy_{u},
$$

$$
q_{s} = Z_{0} + \int_{0}^{s} \frac{a_{u}}{\sigma^{2} e^{2\kappa u}} q_{u} dy_{u} + \int_{0}^{s} \frac{\kappa}{\sigma^{2} e^{2\kappa u}} \text{diag}(\hat{\mu}) q_{u} dy_{u} + \int_{0}^{s} A q_{u} du.
$$

Proof. Again consider:

$$
V_{0,t} = Q_{0,t} Z_{t}
$$

where

$$
Q_{0,t} = \exp \left( \alpha \int_{0}^{t} e^{\kappa u} \langle \hat{\mu}; Z_{u} \rangle du \right)
$$
for a constant $\alpha$. Then
\[
\mathbb{E}[Q_t | \mathcal{Y}_s] = \mathbb{E}\left[\exp \left(\alpha \int_0^s e^{\kappa u} \langle \hat{\mu}; Z_u \rangle du \right) \exp \left(\alpha \int_s^t e^{\kappa u} \langle \hat{\mu}; Z_u \rangle du \right) | \mathcal{Y}_s \right]
\]
\[
= \mathbb{E}[Q_{0,s} Q_{s,t} | \mathcal{Y}_s]
\]
\[
= \mathbb{E} \left[ Q_{0,s} \mathbb{E} \left[ \exp \left(\alpha \int_s^t e^{\kappa u} \langle \hat{\mu}; Z_u \rangle du \right) | \mathcal{Y}_s \cap \mathcal{F}_s^Z \right] | \mathcal{Y}_s \right]
\]
\[
= \mathbb{E} \left[ Q_{0,s} \mathbb{E} \left[ Q_{s,t} | \mathcal{Y}_s \cap \mathcal{F}_s^Z \right] | \mathcal{Y}_s \right].
\]

We first calculate $\mathbb{E} \left[ Q_{s,t} | \mathcal{Y}_s \cap \mathcal{F}_s^Z \right]$. Write
\[
V_{s,t} = Q_{s,t}(\alpha) Z_t.
\]
\[
V_{s,s} = Z_s.
\]

The dynamics of $V_{s,t}$ are:
\[
dV_{s,t} = d(Q_{s,t}(\alpha)) Z_t + Q_{s,t}(\alpha) dZ_t
\]
\[
= \alpha e^{\kappa t} \langle \hat{\mu}; Z_t \rangle Q_{s,t}(\alpha) Z_t dt + Q_{s,t}(\alpha) (AZ_t dt + dM_t)
\]
\[
= (\alpha e^{\kappa t} \text{diag}(\hat{\mu}) + A) Q_{s,t}(\alpha) Z_t dt + Q_{s,t}(\alpha) dM_t.
\]

So
\[
V_{s,t} = V_{s,s} + \int_s^t (\alpha e^{\kappa u} \text{diag}(\hat{\mu}) + A) V_u du + \int_s^t Q_u(\alpha) dM_u,
\]
then,
\[
\mathbb{E}[V_t | \mathcal{Y}_s \cap \mathcal{F}_s^Z] = V_{s,s} + \int_s^t (\alpha e^{\kappa u} \text{diag}(\hat{\mu}) + A) \mathbb{E}[V_u | \mathcal{Y}_s \cap \mathcal{F}_s^Z] du.
\]

Suppose $\Psi(s, t)$ is the matrix solution of the system of ordinary differential equations:
\[
\frac{d\Psi(s, t)}{dt} = (\alpha e^{\kappa t} \text{diag}(\hat{\mu}) + A) \Psi(s, t)
\]
\[
\Psi(s, s) = I_{N_2}.
\]

Then
\[
\mathbb{E}[V_{s,t} | \mathcal{Y}_s \cap \mathcal{F}_s^Z] = \Psi(s, t) Z_s
\]
and
\[
\mathbb{E} \left[ Q_{s,t} | \mathcal{Y}_s \cap \mathcal{F}_s^Z \right] = (\Psi(s, t) Z_s, 1).
\]
Therefore,
\[
\mathbb{E}[Q_{0,t}|\mathcal{Y}_s] = \mathbb{E}[Q_{0,s}\langle \Psi(s,t)Z_s, 1 \rangle|\mathcal{Y}_s]
\]
\[
= \mathbb{E}[\exp \left( \alpha \int_0^s e^{\kappa u} \langle \hat{\mu}, Z_u \rangle du \right) \langle \Psi(s,t)Z_s, 1 \rangle|\mathcal{Y}_s]
\]
\[
= \langle \Psi(s,t) \mathbb{E}[\exp \left( \alpha \int_0^s e^{\kappa u} \langle \hat{\mu}, Z_u \rangle du \right) Z_s|\mathcal{Y}_s], 1 \rangle.
\]

We now calculate:
\[
\mathbb{E}[\exp \left( \alpha \int_0^s e^{\kappa u} \langle \hat{\mu}, Z_u \rangle du \right) Z_s|\mathcal{Y}_s].
\]

However, from (5.3.25) and (5.3.33), we have:
\[
\mathbb{E}[\exp \left( \kappa \int_0^s e^{\kappa(u-t)} \langle \hat{\mu}, Z_u \rangle du \right) Z_s|\mathcal{Y}_s] = \frac{\rho_s(ke^{-\kappa t})}{\langle q_s, 1 \rangle},
\]
where \(\rho_s(\alpha)\) is the solution of the system of differential equations:
\[
\rho_s = \rho_0 + \int_0^s A \rho_u du + \alpha \int_0^s e^{\kappa u} \text{diag}(\hat{\mu}) \rho_u du + \int_0^s \frac{a_u}{\sigma^2 e^{2\kappa u}} \rho_u dy_u + \int_0^s \frac{\kappa}{\sigma^2 e^{2\kappa u}} \text{diag}(\hat{\mu}) \rho_u dy_u
\]
and \(q_t\) is the solution of the differential equation:
\[
q_t = Z_0 + \int_0^t \frac{a_u}{\sigma^2 e^{2\kappa u}} q_u dy_u + \int_0^t \frac{\kappa}{\sigma^2 e^{2\kappa u}} \text{diag}(\hat{\mu}) q_u dy_u + \int_0^t A q_u du.
\]

Therefore:
\[
\mathbb{E}[\exp \left( ke^{-\kappa t} \int_0^t e^{\kappa u} \langle \hat{\mu}, Z_u \rangle du \right) |\mathcal{Y}_s] = \langle \Psi(s,t) \frac{\rho_s(ke^{-\kappa t})}{\langle q_s, 1 \rangle}, 1 \rangle.
\]

Therefore, from (5.3.34), proposition 10 and proposition 8 we have an expression for the forward price:
\[
f(s,t) = S_0 f_t \exp \left( \frac{\sigma^2}{4\kappa} (1 - e^{-2\kappa(t-s)}) \right) \langle \alpha, e^{K(t-s)} J_s \rangle
\]
\[
\exp \left( e^{-\kappa t} \left( y_s - \int_0^s a_u du \right) \right) \frac{\langle \rho_s(-ke^{-\kappa t}), 1 \rangle}{\langle q_s, 1 \rangle} \langle \Psi(s,t) \frac{\rho_s(ke^{-\kappa t})}{\langle q_s, 1 \rangle}, 1 \rangle
\]
\[
= S_0 f_t \exp \left( \frac{\sigma^2}{4\kappa} (1 - e^{-2\kappa(t-s)}) \right) \langle \alpha, e^{K(t-s)} J_s \rangle
\]
\[
\exp \left( e^{-\kappa t} \left( \int_0^s b_u \langle \hat{\mu}, Z_u \rangle du + \sigma \int_0^s e^{\kappa u} dW_u^\lambda \right) \right) \frac{\langle \rho_s(-ke^{-\kappa t}), 1 \rangle}{\langle q_s, 1 \rangle} \langle \Psi(s,t) \frac{\rho_s(ke^{-\kappa t})}{\langle q_s, 1 \rangle}, 1 \rangle.
\]
The above forward expression will be used to price options on forwards, but a more explicit expression for the forward price in terms of the spot $S_s$ is given by:

$$f(s, t) = S_s f_t \exp \left( \frac{\sigma^2}{4\kappa} (1 - e^{-2\kappa(t-s)}) \right) \langle \alpha, \epsilon^\kappa(t-s) J_s \rangle \frac{\rho_s(-\kappa e^{-\kappa t})}{\langle q_s, 1 \rangle} \langle \Psi(s, t) \rho_s(\kappa e^{-\kappa t}) \rangle \langle q_s, 1 \rangle.$$ (5.3.41)

### 5.4 Pricing of options on forward

In this section we shall derive the price of a call option on forward contract, but firstly the expression for the forward price (5.3.40) needs to be simplified so that a call option price can be obtained.

#### 5.4.1 Expectation based approximation

The following results are for approximating the quantities $q_s$ and $\rho_s(\alpha)$ in the forward price (5.3.40).

**Proposition 11.** Let $H$ be the matrix solution of the following system of differential equations:

$$\frac{dH_{0,t}}{dt} = \left[ A + \theta_u \left( \frac{a_u}{\sigma e^u} I_N + \frac{\kappa}{\sigma} \text{diag}(\hat{\mu}) \right) \right] H_{0,t}$$

$$H_{0,0} = I_N.$$

Then:

$$\mathbb{E}[q_s | \mathcal{F}_0] = H_{0,s} q_0. \quad (5.4.1)$$

**Proof.** By definition of $\tilde{P}$ and using the conditional Bayes’ theorem, we have:

$$\mathbb{E}[q_s | \mathcal{F}_0] = \frac{\mathbb{E}[\Lambda_s q_s | \mathcal{F}_0]}{\mathbb{E}[\Lambda_s | \mathcal{F}_0]}.$$ (5.4.2)

Since $\Lambda$ is a martingale under the real world probability measure:

$$\mathbb{E}[\Lambda_s | \mathcal{F}_0] = \Lambda_0 = 1.$$
Consequently, we shall evaluate the numerator in (5.4.2).

From (5.3.16) and (5.3.17), write:

\[ q_s = q_0 + \int_0^s \frac{a_u}{\sigma e^{\kappa u}} q_u d\Phi_u + \int_0^s \frac{\kappa}{\sigma} diag(\hat{\mu}) q_u d\Phi_u + \int_0^s A q_u d\mu \]

so that

\[ dq_s = \left( \frac{a_s}{\sigma e^{\kappa s}} + \frac{\kappa}{\sigma} diag(\hat{\mu}) \right) q_s d\Phi_s + A q_s ds. \tag{5.4.3} \]

Also, from (5.3.23):

\[ d\Lambda_s = \theta_s \Lambda_s d\Phi_s \]

Using Itô’s lemma:

\[ d(\Lambda_s q_s) = d\Lambda_s q_s + \Lambda_s dq_s + d\langle \Lambda, q \rangle_s \]

\[ = \theta_s \Lambda_s q_s d\Phi_s + \left( \frac{a_s}{\sigma e^{\kappa s}} I_N + \frac{\kappa}{\sigma} diag(\hat{\mu}) \right) \Lambda_s q_s d\Phi_s + A \Lambda_s q_s ds \]

\[ + \theta_s \left( \frac{a_s}{\sigma e^{\kappa s}} I_N + \frac{\kappa}{\sigma} diag(\hat{\mu}) \right) \Lambda_s q_s ds. \]

So,

\[ \Lambda_s q_s = \Lambda_0 q_0 + \int_0^s \left( \theta_u I_N + \frac{a_u}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} diag(\hat{\mu}) \right) \Lambda_u q_u d\Phi_u \]

\[ + \int_0^s \left[ A + \theta_u \left( \frac{a_u}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} diag(\hat{\mu}) \right) \right] \Lambda_u q_u d\mu. \tag{5.4.4} \]

\( \Phi \) is a Brownian motion under \( \bar{P} \) so by taking the expected value under \( \bar{P} \):

\[ \bar{E}[\Lambda_s q_s | F_0] = \Lambda_0 q_0 + \int_0^s \left[ A + \theta_u \left( \frac{a_u}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} diag(\hat{\mu}) \right) \right] \bar{E}[\Lambda_u q_u | F_0] d\mu. \tag{5.4.5} \]

Now let \( H \) be the matrix solution of the following system of differential equations:

\[ \frac{dH_{0,t}}{dt} = \left[ A + \theta_u \left( \frac{a_u}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} diag(\hat{\mu}) \right) \right] H_{0,t} \]

\[ H_{0,0} = I_N. \]

Then:

\[ \bar{E}[q_s | F_0] = \bar{E}[\Lambda_s q_s | F_0] = H_{0,s} q_0. \tag{5.4.6} \]
We now approximate the quantities $\rho_s(\kappa e^{-kt})$ and $\rho_s(-\kappa e^{-kt})$.

We firstly calculate $\mathbb{E}[\rho_s(\alpha)|\mathcal{F}_0]$ where $\rho$ is the solution of (5.3.26).

**Proposition 12.** Write $G$ for the matrix solution of the system of differential equations:

$$
\frac{dG_{0,u}(\alpha)}{du} = \left[ A + \alpha e^{\kappa u} \text{diag}(\hat{\mu}) + \theta \left( \frac{a_u}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} \text{diag}(\hat{\mu}) \right) \right] G_{0,u}
$$

$$
G_{0,0}(\alpha) = I_N .
$$

Then

$$
\mathbb{E}[\rho_s(\alpha)|\mathcal{F}_0] = \mathbb{E}[\Lambda_s \rho_s(\alpha)|\mathcal{F}_0] = G_{0,s} \rho_0(\alpha).
$$

**Proof.** From the conditional Baye’s rule:

$$
\mathbb{E}[\rho_s(\alpha)|\mathcal{F}_0] = \mathbb{E}[\Lambda_s \rho_s(\alpha)|\mathcal{F}_0].
$$

Now we shall calculate $\mathbb{E}[\Lambda_s \rho_s|\mathcal{F}_0]$. Since

$$
\frac{dy_u}{\sigma e^{\kappa u}} = d\Phi_u ,
$$

We can write the dynamics of $\rho$ from (5.3.26) as:

$$
d\rho_s = (A + \alpha e^{\kappa u} \text{diag}(\hat{\mu})) \rho_s ds + \left( \frac{a_s}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} \text{diag}(\hat{\mu}) \right) \rho_s d\Phi_s .
$$

Also

$$
d\Lambda_s = \theta_s \Lambda_s d\Phi_s .
$$

So from Itô’s lemma:

$$
d(\Lambda_s \rho_s) = d\Lambda_s \rho_s + \Lambda_s d\rho_s + d\langle \Lambda, \rho \rangle_s
$$

$$
= (A + \alpha e^{\kappa u} \text{diag}(\hat{\mu})) \Lambda_s \rho_s ds + \left( \frac{a_s}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} \text{diag}(\hat{\mu}) \right) \Lambda_s \rho_s d\Phi_s + \theta_s \Lambda_s \rho_s d\Phi_s
$$

$$
+ \left( \frac{a_s}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} \text{diag}(\hat{\mu}) \right) \theta_s \rho_s \Lambda_s ds .
$$

Then

$$
\Lambda_s \rho_s = \int_0^s \left[ A + \alpha e^{\kappa u} \text{diag}(\hat{\mu}) + \theta \left( \frac{a_u}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} \text{diag}(\hat{\mu}) \right) \right] \Lambda_u \rho_u du +
$$

$$
\int_0^s \left( \theta_u I_N + \frac{a_u}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} \text{diag}(\hat{\mu}) \right) \Lambda_u \rho_u d\Phi_u . \quad (5.4.7)
$$
and by taking the expected value under \( \overline{P} \),
\[
\mathbb{E}[\Lambda_s \rho_s | \mathcal{F}_0] = \int_0^s \left[ A + \alpha e^{\kappa u} \text{diag}(\mu) + \theta \left( \frac{a_u}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} \text{diag}(\hat{\mu}) \right) \right] \mathbb{E}[\Lambda_u \rho_u | \mathcal{F}_0] du.
\]

Write \( G \) for the matrix solution of the system of differential equations:
\[
\frac{dG_{0,u}}{du} = \left[ A + \alpha e^{\kappa u} \text{diag}(\mu) + \theta \left( \frac{a_u}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} \text{diag}(\hat{\mu}) \right) \right] G_{0,u} \quad G_{0,0} = I_N.
\]

Then
\[
\mathbb{E}[\rho_s | \mathcal{F}_0] = \mathbb{E}[\Lambda_s \rho_s | \mathcal{F}_0] = G_{0,s} \rho_0.
\]

We now give approximations for the quantities \( \rho_s(\kappa e^{-\kappa t}) \) and \( \rho_s(-\kappa e^{-\kappa t}) \).

**Proposition 13.** The forward price given in (5.3.40) can be approximated with:
\[
f(s, t) = C_{0,s,t} \exp \left( e^{-\kappa t} \sigma \int_0^s e^{\kappa u} dW_u \right), \tag{5.4.8}
\]
where
\[
C_{0,s,t} = S_0 f_t \exp \left( \frac{\sigma^2}{4\kappa} (1 - e^{-2\kappa(t-s)}) \right) \left\langle \alpha, e^{\kappa t} J_0 \right\rangle \left\langle G_{0,s}(\kappa e^{-\kappa t})\rho_0, 1 \right\rangle \left\langle H_{0,s}, 1 \right\rangle \left\langle \Psi(s, t) \frac{\left\langle G_{0,s}(-\kappa e^{\kappa t})\rho_0, 1 \right\rangle}{\left\langle H_{0,s} q_0, 1 \right\rangle^2} G_{0,s}(\kappa e^{\kappa t})\rho_0, 1 \right\rangle. \tag{5.4.9}
\]

**Proof.** We shall approximate \( \rho \) by its expected value:
\[
\rho_s(-\kappa e^{-\kappa t}) \approx \mathbb{E}[\rho_s(-\kappa e^{-\kappa t}) | \mathcal{F}_0] = G_{0,s}(-\kappa e^{-\kappa t}) \rho_0.
\]
Here \( G_{0,t}(-\kappa e^{-\kappa t}) \) is the solution of the system of differential equations:
\[
\frac{dG_{0,u}(-\kappa e^{-\kappa t})}{du} = \left[ A - \kappa e^{\kappa(u-t)} \text{diag}(\hat{\mu}) + \theta \left( \frac{a_u}{\sigma e^{\kappa u}} I_N + \frac{\kappa}{\sigma} \text{diag}(\hat{\mu}) \right) \right] G_{0,u}(-\kappa e^{-\kappa t}) \quad G_{0,0}(-\kappa e^{-\kappa t}) = I_N.
\]
Similarly, we assume that:

\[
\rho_s(\kappa e^{-kt}) \approx \mathbb{E}[\rho_s(\kappa e^{-kt})|\mathcal{F}_0] = G_{0,t}(\kappa e^{-kt}) \rho_0
\]

Where \( G_{0,t}(\kappa e^{-kt}) \) is the solution of the system of differential equations

\[
\frac{dG_{0,u}(\kappa e^{-kt})}{du} = \left[ A + \kappa e^{(u-t)} \text{diag}(\hat{\mu}) + \theta \left( \frac{a_u}{\sigma e^u} I_N + \frac{\kappa}{\sigma} \text{diag}(\hat{\mu}) \right) \right] G_{0,u}(\kappa e^{-kt})
\]

\[ G_{0,0}(\kappa e^{-kt}) = I_N. \]

We also simplify the forward price by approximating

\[
\exp \left( e^{-kt} \left( \int_0^s b_u(\hat{\mu}, Z_u) du \right) \right) = \exp \left( \kappa e^{-kt} \left( \int_0^s e^{\kappa u} \langle \hat{\mu}, Z_u \rangle du \right) \right)
\]

by

\[
\mathbb{E}[\exp \left( \kappa e^{-kt} \left( \int_0^s e^{\kappa u} \langle \hat{\mu}, Z_u \rangle du \right) \right) |\mathcal{F}_0].
\]

From Proposition 9, with \( \alpha = \kappa e^{-kt} \).

\[
\mathbb{E}[\exp \left( \kappa e^{-kt} \left( \int_0^s e^{\kappa u} \langle \hat{\mu}, Z_u \rangle du \right) \right) |\mathcal{F}_0] = \frac{\langle \rho_s(\kappa e^{-kt}), 1 \rangle}{\langle q_s, 1 \rangle}
\]

where \( \rho(\kappa e^{-kt}) \) is the solution of the stochastic differential equation:

\[
\rho_s = \rho_0 + \int_0^s (A + \kappa e^{(u-t)} \text{diag}(\hat{\mu})) \rho_u du + \int_0^s \left( \frac{a_u}{\sigma^2 e^{2u}} I_N + \frac{\kappa}{\sigma^2 e^{\kappa u}} \text{diag}(\hat{\mu}) \right) \rho_u dy_u
\]

and \( q_s \) is solution of (5.3.37). Therefore an approximation for the forward price is:

\[
f(s, t) = C_{0,s,t} \exp \left( e^{-\kappa t} \int_0^s e^{\kappa u} dW_u^\lambda \right),
\]

where

\[
C_{0,s,t} = S_0 f_t \exp \left( \frac{\sigma^2}{4\kappa} (1 - e^{-2\kappa(t-s)}) \right) \left( \langle \alpha, e^{\kappa t} J_0 \rangle \frac{\langle G_{0,s}(\kappa e^{-kt}) \rho_0, 1 \rangle}{\langle H_{0,s}, 1 \rangle} \right)
\]

\[
\left( \Psi(s, t) \frac{\langle G_{0,s}(\kappa e^{-kt}) \rho_0, 1 \rangle}{H_{0,s} \rho_0, 1} \right)^2 G_{0,s}(\kappa e^{-kt}) \rho_0, 1 \right).
\]
5.4.2 Deriving the call price

We now calculate the price of a call option written at time 0 on a forward to be entered at time \( s \geq 0 \), with a strike price \( K > 0 \). Recall that
\[
 f(s, t) = C_{0,s,t} \exp \left( e^{-\kappa t} \sigma \int_0^s e^{\kappa u}dW_u^\lambda \right),
\]
and consider \( r \) to be the risk free interest rate. Note that we are calculating under the arbitrage free probability measure, so from [39], the call price at time 0 is given by:
\[
 V(0, s, t) = \mathbb{E}[e^{-rs}(f(s, t) - K)^+|\mathcal{F}_0]. 
\] (5.4.10)

**Proposition 14.** Define the quantities:
\[
 \nu^2 = \int_0^s e^{2\kappa u}du, \quad d_1 = \frac{e^{\kappa t}}{\nu\sigma} \ln \left( \frac{K}{C_{0,s,t}} \right) - e^{-\kappa t} \sigma \nu, \quad d_2 = d_1 + e^{-\kappa t} \sigma \nu.
\]
Then the call option price is given by:
\[
 V(0, s, t) = e^{-rs}C_{0,s,t} \exp \left( \frac{1}{2} e^{-2\kappa t} \sigma^2 \nu^2 \right) \mathcal{N}(-d_1) - e^{-rs}K \mathcal{N}(-d_2) \quad (5.4.11)
\]

**Proof.** From the call price definition (5.4.10):
\[
 V(0, s, t) = \mathbb{E}[e^{-rs}(f(s, t) - K)^+|\mathcal{F}_0]
 = \mathbb{E}[e^{-rs}(f(s, t) - K)1_{(f(s,t)\geq K)}|\mathcal{F}_0]
 = \mathbb{E}[e^{-rs}f(s, t)1_{(f(s,t)\geq K)}|\mathcal{F}_0] - \mathbb{E}[e^{-rs}K1_{(f(s,t)\geq K)}|\mathcal{F}_0].
\]

Under the risk neutral probability, \( W^\lambda \) is a Brownian motion, so
\[
 z = \int_0^s e^{\kappa u}dW_u^\lambda \sim \mathcal{N}(0, \int_0^s e^{2\kappa u}du),
\]
where \( \mathcal{N} \) refers to the standard normal distribution.

write
\[
 d_1 = \frac{e^{\kappa t}}{\nu\sigma} \ln \left( \frac{K}{C_{0,s,t}} \right) - e^{-\kappa t} \sigma \nu,
\]
and
\[
 \nu^2 = \int_0^s e^{2\kappa u}du.
\]
then
\[
E\left[e^{-rs}f(s,t)\mathbb{1}_{\{f(s,t)\geq K\}}|\mathcal{F}_0\right] = e^{-rs}C_{0,s,t} \int_{\exp\left(\frac{K}{C_{0,s,t}}\right)}^{\infty} \exp\left(e^{-\kappa t}z\right) \frac{1}{\nu \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{z^2}{\nu^2}\right) dz
\]
\[
= e^{-rs}C_{0,s,t} \exp\left(\frac{1}{2} e^{-2\kappa t} \sigma^2 \nu^2\right) \int_{\exp\left(\frac{K}{C_{0,s,t}}\right)}^{\infty} \frac{1}{\nu \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(z - e^{-\kappa t} \sigma\nu)^2}{\nu}\right) dz
\]
\[
= e^{-rs}C_{0,s,t} \exp\left(\frac{1}{2} e^{-2\kappa t} \sigma^2 \nu^2\right) \int_{d_1}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} x^2\right) dx
\]
\[
= e^{-rs}C_{0,s,t} \exp\left(\frac{1}{2} e^{-2\kappa t} \sigma^2 \nu^2\right) \mathcal{N}(-d_1).
\]
Similarly, for the second expectation term,
\[
E\left[e^{-rs}K\mathbb{1}_{\{f(s,t)\geq K\}}|\mathcal{F}_0\right] = e^{-rs}K \int_{\exp\left(\frac{K}{C_{0,s,t}}\right)}^{\infty} \frac{1}{\nu \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(z - e^{-\kappa t} \sigma\nu)^2}{\nu}\right) dz
\]
\[
= e^{-rs}K \int_{d_2}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} x^2\right) dx
\]
\[
= e^{-rs}K \mathcal{N}(-d_2).
\]
Here
\[
d_2 = \frac{e^{\kappa t}}{\nu \sigma} \ln\left(\frac{K}{C_{0,s,t}}\right) = d_1 + e^{-\kappa t} \sigma \nu.
\]
Therefore, the call price is approximately given by:
\[
V(0,s,t) = e^{-rs}C_{0,s,t} \exp\left(\frac{1}{2} e^{-2\kappa t} \sigma^2 \nu^2\right) \mathcal{N}(-d_1) - e^{-rs}K \mathcal{N}(-d_2).
\]

5.5 Parameter estimation

In this section we shall develop the estimators for the spot price model parameters. In the presence of a hidden Markov process in the model, the usual maximum likelihood techniques are not immediately applicable, so we need to adapt a recursive filter to a discrete dataset. We apply exact adaptive recursive filters for Markov chains developed in [18].
5.5.1 Discretisation of the observed process \( y_t \)

Recall from (5.3.2) that we observe:
\[
y_t = \int_0^t a_u du + \int_0^t b_u (\mu, Z_u) du + \sigma \int_0^t e^{\kappa u} dW^\lambda_u,
\]
where
\[
a_u = \frac{d}{du} \left( e^{\kappa u} \ln(f_u) \right),
\]
and
\[
b_u = \kappa e^{\kappa u}.
\]

Suppose in practice we observe \( y \) at the end of time intervals of size \( h \). Then for a small time steps \( h \), we approximately observe the increments:
\[
Y_t = y_{t+h} - y_t \approx \left( e^{\kappa(t+h)} \ln(f_{t+h}) - e^{\kappa t} \ln(f_t) \right) + b_t (\mu, Z_t) h + \sigma e^{\kappa t} (W^\lambda_{t+h} - W^\lambda_t). \quad (5.5.1)
\]

Because \( W^\lambda \) is a \( \mathbb{P} \)-Brownian motion:
\[
W^\lambda_{t+h} - W^\lambda_t \sim \mathcal{N}(0, \sqrt{h}).
\]

To simplify the notation, suppose that \( t \) is the \( n^{th} \) time step so \( t = nh \); then write:
\[
Y_n = H_n + D_{n-1} (\mu, Z_{n-1}) + K_{n-1} B_n,
\]
where
\[
\begin{aligned}
H_n &= e^{\kappa(n+1)h} \ln(f_{(n+1)h} g_{(n+1)h}) - e^{\kappa nh} \ln(f_{nh} g_{nh}) \\
D_n &= b_{nh} h = \kappa e^{\kappa nh} \\
K_n &= \sigma e^{\kappa nh} \sqrt{h},
\end{aligned} \quad (5.5.3)
\]
and \( B_n \) are i.i.d standard normal variables.

We assume the sequence \( \{Y_n, n = 1, 2, \ldots\} \) are our observations. We shall modify the filtering theory for Markov chains observed in Gaussian noise developed in [19].

Write \( \phi \) for the density of the standard normal distribution, so \( \phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) \)
and define the quantities \( \lambda_n \) and \( \Lambda_n \) by:
\[
\lambda_n = \frac{\phi(B_n)}{\phi(Y_n) K_{n-1}} = \frac{\phi \left( \frac{Y_n - H_n - D_{n-1} (\mu, Z_{n-1})}{K_{n-1}} \right)}{\phi(Y_n) K_{n-1}} \quad (5.5.4)
\]
and
\[ \Lambda_n = \prod_{l=1}^{n} \lambda_l = \prod_{l=1}^{n} \frac{\phi \left( \frac{Y_l - H_l - D_{l-1}(\mu, Z_{l-1})}{K_{l-1}} \right)}{\phi(Y_l)K_{l-1}}. \] (5.5.5)

Consider the quantities \( \tilde{\lambda} \) and \( \tilde{\Lambda} \), such that:
\[ \forall n \geq 1 : \tilde{\lambda}_n = \frac{1}{\lambda_n} \]
and
\[ \forall n \geq 1 : \tilde{\Lambda}_n = \frac{1}{\Lambda_n}. \]

**Lemma 14.** Define the probability measure \( \tilde{\mathbb{P}} \) by setting the restriction of the Radon-Nikodym derivative to \( \mathcal{Y}_n \) equal to \( \tilde{\Lambda}_n \):
\[ \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} |_{\mathcal{Y}_n} = \tilde{\Lambda}_n. \] (5.5.6)

Then under \( \tilde{\mathbb{P}} \), \( Z \) has the same dynamics and the \( Y_n \) are i.i.d \( \mathcal{N}(0, 1) \) random variables.

**Proof.** We shall prove that under \( \tilde{\mathbb{P}} \) the random variables \( Y_n \) are iid standard normally distributed.

For any random variable \( \chi \), we shall write \( I(\chi \leq x) \) for the indicator function of the set \( \{ \omega : \chi(\omega) \leq x \} \). Recall that \( \mathcal{F}^Z \) refers to the filtration generated by the hidden Markov process \( Z \), so \( \mathcal{Y} \vee \mathcal{F}^Z \) will denote the complete filtration generated by \( y \) and \( Z \). Using the conditional version of Bayes’ theorem we have for any \( x \in \mathbb{R} \):
\[ \tilde{\mathbb{P}}(Y_n < x | \mathcal{Y}_{n-1}) = \mathbb{E}[I(Y_n < x) | \mathcal{Y}_{n-1} \vee \mathcal{F}^Z_{n-1}] \]
\[ = \frac{\mathbb{E}[\tilde{\Lambda}_n I(Y_n < x) | \mathcal{Y}_{n-1} \vee \mathcal{F}^Z_{n-1}]}{\mathbb{E}[\tilde{\Lambda}_n | \mathcal{Y}_{n-1} \vee \mathcal{F}^Z_{n-1}]} \]
\[ = \frac{\tilde{\Lambda}_n^{-1} \mathbb{E}[\tilde{\lambda}_n I(Y_n < x) | \mathcal{Y}_{n-1} \vee \mathcal{F}^Z_{n-1}]}{\mathbb{E}[\tilde{\Lambda}_n | \mathcal{Y}_{n-1} \vee \mathcal{F}^Z_{n-1}]} \].

We use the conditioning with respect to \( \mathcal{Y}_{n-1} \vee \mathcal{F}^Z_{n-1} \) since the terms \( Y_n \) contain the hidden Markov chain \( Z \). Note that if :
\[ \Gamma(x) = \int_{-\infty}^{x} \frac{\phi(Y_n)K_{n-1}}{\phi(B_n)} \phi(B_n) dB_n \]
then
\[ \Gamma(x) = \int_{-\infty}^{x} \phi(Y_n) dY_n. \]

So
\[ \mathbb{E}[\bar{\lambda}_n | \mathcal{Y}_{n-1} \cup \mathcal{F}_{n-1}^Z] = 1. \]

Therefore,
\[ \mathbb{P}(Y_n < x | \mathcal{Y}_{n-1} \cup \mathcal{F}_{n-1}^Z) = \int_{-\infty}^{\infty} \frac{\phi(Y_n) K_{n-1}}{\phi(B_n)} \phi(B_n) I(Y_n < x) dB_n = \Gamma(x) = \mathbb{P}(Y_n < x). \]
(5.5.7)

This proves the independence and the normal distribution of the observation increments \(Y_n\).

We now prove that the dynamics of the Markov chain \(Z\) are the same under \(\mathbb{P}\) and \(\mathbb{\bar{P}}\). Suppose that under \(\mathbb{P}\) the dynamics of \(Z\) are approximately:
\[ Z_{n+1} = \Pi Z_n + \Delta M_{n+1} \]
with
\[ \Pi = I + hA, \]
and here \(\Delta M_{n+1} = M_{n+1} - M_n\) is a vector martingale increment.

To show that the dynamics stay the same under \(\mathbb{\bar{P}}\), we must show that under \(\mathbb{\bar{P}}\):
\[ \mathbb{\bar{E}}[\Delta M_{n+1} | Z_0, \ldots, Z_n] = \mathbb{\bar{E}}[Z_{n+1} - \Pi Z_n | Z_0, \ldots, Z_n] = 0. \]

This means that \(Z\) has the same martingale decomposition under \(\mathbb{\bar{P}}\). Because \(\Pi Z_n\) is \(\sigma(Z_0, \ldots, Z_n)\)-measurable:
\[ \mathbb{\bar{E}}[Z_{n+1} - (I + Ah)Z_n | Z_0, \ldots, Z_n] = \mathbb{\bar{E}}[Z_{n+1} | Z_0, \ldots, Z_n] - \Pi Z_n. \]

We have:
\[ \mathbb{\bar{E}}[Z_{n+1} | Z_0, \ldots, Z_n] = \mathbb{\bar{E}}[\mathbb{\bar{E}}[Z_{n+1} | Z_0, \ldots, Z_n, Y_0, \ldots, Y_n] | Z_0, \ldots, Z_n]. \]

Working on the inner expectation and for ease of notation calling the sigma algebra generated by \(Z\) and \(Y\):
\[ \sigma(Z_0, \ldots, Z_n, Y_0, \ldots, Y_n) = \mathcal{L}_n. \]
Now, using Bayes’ rule for conditional expectation, and using the fact that \( \Lambda_n \) is \( \mathcal{L}_n \)-measurable:

\[
\bar{E}[Z_{n+1}|\mathcal{L}_n] = \frac{\bar{E}[\Lambda_{n+1}Z_{n+1}|\mathcal{L}_n]}{\bar{E}[\Lambda_{n+1}|\mathcal{L}_n]} = \frac{\bar{E}[\Lambda_{n+1}Z_{n+1}|\mathcal{L}_n]}{\bar{E}[\Lambda_{n+1}|\mathcal{L}_n]}
\]

Now to calculate the numerator term, we have:

\[
\bar{E}[\Lambda_{n+1}Z_{n+1}|\mathcal{L}_n] = \bar{E}\left[\frac{\phi(Y_{n+1})K_n}{\phi(B_{n+1})}Z_{n+1}|\mathcal{L}_n\right]
\]

\[
= \bar{E}\left[\frac{1}{\sqrt{2\pi}} \exp\left(\left(\frac{1}{2}H_{n+1} + D_n\langle \mu, Z_n \rangle + K_nB_{n+1}\right)^2\right) \frac{K_n}{\phi(B_{n+1})}Z_{n+1}|\mathcal{L}_n\right]
\]

\[
= \bar{E}\left[\frac{1}{\sqrt{2\pi}} \exp\left(\left(\frac{1}{2}H_{n+1} + D_n\langle \mu, Z_n \rangle + K_nB_{n+1}\right)^2\right) \frac{K_n}{\phi(B_{n+1})}(\Pi Z_n + \Delta M_{n+1})|\mathcal{L}_n\right]
\]

\[
= \bar{E}\left[\frac{1}{\sqrt{2\pi}} \exp\left(\left(\frac{1}{2}H_{n+1} + D_n\langle \mu, Z_n \rangle + K_nB_{n+1}\right)^2\right) \frac{K_n}{\phi(B_{n+1})}\Pi Z_n|\mathcal{L}_n\right]
\]

\[
= \bar{E}\left[\sum_{i=1}^{N} \langle Z_n, e_i \rangle \frac{1}{\sqrt{2\pi}} \exp\left(\left(\frac{1}{2}H_{n+1} + D_n\langle \mu, e_i \rangle + K_nB_{n+1}\right)^2\right) \frac{K_n}{\phi(B_{n+1})}\Pi e_i|\mathcal{L}_n\right]
\]

\[
= \bar{E}\left[\sum_{i=1}^{N} \langle Z_n, e_i \rangle \frac{1}{\sqrt{2\pi}} \exp\left(\left(\frac{1}{2}H_{n+1} + D_n\langle \mu, e_i \rangle + K_nB_{n+1}\right)^2\right) \frac{K_n}{\phi(B_{n+1})}\Pi e_i|\mathcal{L}_n\right]
\]

\[
= \sum_{i=1}^{N} \bar{E}[\langle Z_n, e_i \rangle|\mathcal{L}_n] \bar{E}\left[\frac{1}{\sqrt{2\pi}} \exp\left(\left(\frac{1}{2}H_{n+1} + D_n\langle \mu, e_i \rangle + K_nB_{n+1}\right)^2\right) \frac{K_n}{\phi(B_{n+1})}\Pi e_i|\mathcal{L}_n\right]
\]

(5.5.8)

The last equality is a result of the fact that under \( \mathbb{P} \) the random variables \( B_{n+1} \) are i.i.d standard normal random variables, and independent of \( Z_{n+1} \), so if:

\[
\mathcal{R}_n = (H_{n+1} + D_n\langle \mu, e_i \rangle + K_nB_{n+1})^2
\]

then

\[
\bar{E}\left[\frac{1}{\sqrt{2\pi}} \exp(\mathcal{R}_n) \frac{K_n}{\phi(B_{n+1})}\Pi e_i|\mathcal{L}_n\right] = \int_{\mathbb{R}} \frac{\phi(Y_{n+1})K_n}{\phi(B_{n+1})} \phi(B_{n+1}) d(B_{n+1}) \Pi e_i
\]

\[
= \int_{\mathbb{R}} \phi(Y_{n+1}) d(Y_{n+1}) \Pi e_i
\]

\[
= \Pi e_i.
\]
That is:
\[
\mathbb{E}[\lambda_{n+1}Z_{n+1}|\mathcal{L}_n] = \sum_{i=1}^{N} \mathbb{E}[\langle Z_n, e_i \rangle | Z_n] \Pi e_i \\
= \sum_{i=1}^{N} \langle \mathbb{E}[Z_{n+1}|Z_n], e_i \rangle \Pi e_i \\
= \sum_{i=1}^{N} \langle Z_n, e_i \rangle \Pi e_i \\
= \Pi Z_n.
\]

Similar to the calculation performed in (5.5.8), we can prove that:
\[
\mathbb{E}[\bar{\lambda}_{n+1}|\mathcal{L}_n] = 1.
\]

That means that under \(\bar{\mathbb{P}}\):
\[
\mathbb{E}[\Delta M_{n+1}|Z_0, \ldots, Z_n] = \mathbb{E}[Z_{n+1} - (I + Ah)Z_n|Z_0, \ldots, Z_n] = 0
\]
and so the Markov chain has the same dynamics under \(\mathbb{P}\) and \(\bar{\mathbb{P}}\).

5.5.2 Estimation

In the discrete model (5.5.2) we observe the quantity \(Y\) while the random Markov chain \(Z\) is unobservable. We shall estimate the parameter vector \(\xi = (\kappa, \sigma, \hat{\mu})\) recursively using the Expectation Maximisation (EM) algorithm.

The EM algorithm starts from some prior values \(\xi_0 = (\kappa_0, \sigma_0, \hat{\mu}_0)\) and re-estimates each of these parameters, one at a time, recursively given the data.

At each step \(p\) the new parameter estimate \(\hat{\xi}_{p+1}\) is chosen in a way to maximise the expected log-likelihood function given the observations \(Y_1, \ldots, Y_n\). Suppose that after \(p\) iterations, we have the estimates \(\xi_p = (\kappa_p, \sigma_p, \hat{\mu}_p)\).
From the likelihood function given in (5.5.5) and (5.5.6), we obtain:

\[
\log \Lambda_n = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \sum_{l=1}^{n} \left( \frac{Y_l - H_l - D_{l-1}\langle \mu, Z_{l-1} \rangle}{K_{l-1}} \right)^2 \\
+ \frac{N}{2} \log(2\pi) + \frac{1}{2} \sum_{l=1}^{n} Y_l^2
\]

\[
= -\sum_{l=1}^{n} \log(K_{l-1}) - \frac{1}{2} \sum_{l=1}^{n} \left( \frac{Y_l - H_l - D_{l-1}\langle \mu, Z_{l-1} \rangle}{K_{l-1}} \right)^2 + \frac{1}{2} \sum_{l=1}^{n} Y_l^2.
\]

\[(5.5.9)\]

### 5.5.3 Estimation of the drift parameter \( \mu \)

The parameter vector \( \mu \) represents the long term mean in the continuous part of our power model defined in (5.2.1), and switches from a value to another according to the state of the Hidden Markov process \( Z \). We can write:

\[
\log \Lambda_n = -\sum_{l=1}^{n} \log(K_{l-1}) - \frac{1}{2} \sum_{l=1}^{n} \sum_{i=1}^{N} \left( \frac{Y_l - H_l - D_{l-1}\langle \mu, e_i \rangle}{K_{l-1}} \right)^2 \langle Z_{l-1}, e_i \rangle + \frac{1}{2} \sum_{l=1}^{n} Y_l^2.
\]

We differentiate in the log-likelihood function with respect to the parameters \( \mu_1, \mu_2, \ldots, \mu_N \).

\[
\frac{\partial \log \Lambda_n}{\partial \mu_j} = \sum_{l=1}^{n} \frac{D_{l-1}}{K_{l-1}} \left( \frac{Y_l - H_l - D_{l-1}\mu_j}{K_{l-1}} \right) \langle Z_{l-1}, e_j \rangle.
\]

At each step \( p \), we update the estimate of the parameter vector \( \mu \) using the value that maximises the likelihood function. Consider the equation:

\[
\frac{\partial \log \Lambda_n}{\partial \mu_j} = 0.
\]

That is:

\[
\sum_{l=1}^{n} \frac{D_{l-1}}{K_{l-1}^2} Y_l \langle Z_{l-1}, e_j \rangle - \sum_{l=1}^{n} \frac{D_{l-1}}{K_{l-1}^2} H_l \langle Z_{l-1}, e_j \rangle = \mu_j \sum_{l=1}^{n} \frac{D_{l-1}^2}{K_{l-1}^2} \langle Z_{l-1}, e_j \rangle.
\]

As we do not observe the state variable \( Z \), we evaluate the above sums by taking their expected value and using a filter. That is, we consider:

\[
E[\sum_{l=1}^{n} \frac{D_{l-1}Y_l}{K_{l-1}^2} \langle Z_{l-1}, e_j \rangle | \mathcal{Y}_n] - E[\sum_{l=1}^{n} \frac{D_{l-1}H_l}{K_{l-1}^2} \langle Z_{l-1}, e_j \rangle | \mathcal{Y}_n] = \mu_j E[\sum_{l=1}^{n} \frac{D_{l-1}^2}{K_{l-1}^2} \langle Z_{l-1}, e_j \rangle | \mathcal{Y}_n].
\]

\[(5.5.10)\]
Write:

\[
\begin{align*}
\hat{A}(n) &= \mathbb{E}\left[ \sum_{l=1}^{n} \frac{D_{l-1}Y_l}{K_{l-1}^2} (Z_{l-1}, e_j) | \mathcal{Y}_n \right] \\
\hat{B}(n) &= \mathbb{E}\left[ \sum_{l=1}^{n} \frac{D_{l-1}H_l}{K_{l-1}^2} (Z_{l-1}, e_j) | \mathcal{Y}_n \right] \\
\hat{C}(n) &= \mathbb{E}\left[ \sum_{l=1}^{n} \frac{D_{l-1}^{2}}{K_{l-1}^2} (Z_{l-1}, e_j) | \mathcal{Y}_n \right].
\end{align*}
\]

Then the new value for the estimate of \( \mu_j \) is given by:

\[
\hat{A}(n) - \hat{B}(n) = \mu_j \hat{C}(n) .
\] (5.5.11)

We first calculate \( \hat{A}(n) \).

**Proposition 15.** Write

\[
A_n = \sum_{l=1}^{n} \frac{D_{l-1}Y_l}{K_{l-1}^2} (Z_{l-1}, e_j).
\]

If we consider again the probability measure \( \bar{\mathbb{P}} \) and note for any process \( X, \sigma_n(X) = \mathbb{E}[\Lambda_n X] \), then a recursive filter for \( A_n \) is given by:

\[
\mathbb{E}[A_n | \mathcal{Y}_n] = \frac{\langle \sigma_n(A_n Z_n) \rangle}{\sigma_n(1)} ,
\]

such that

\[
\sigma_n(A_n Z_n) = \Pi diag(\Gamma^n) \sigma_{n-1}(A_{n-1} Z_{n-1}) + \Gamma^n_j (\sigma_{n-1}(Z_{n-1}), e_j) \Pi e_j
\] (5.5.12)

and

\[
\sigma_n(1) = \langle \sigma_n(Z_n), 1 \rangle = (\Pi diag(\Gamma) \sigma_{n-1}(Z_{n-1}), 1).
\] (5.5.13)

**Proof.** Write

\[
A_n = \sum_{l=1}^{n} \frac{D_{l-1}Y_l}{K_{l-1}^2} (Z_{l-1}, e_j).
\]

So

\[
A_n = A_{n-1} + \frac{D_{n-1}Y_n}{K_{n-1}^2} (Z_{n-1}, e_j).
\]
From the dynamics of the Markov process $Z$,

$$Z_n = \Pi Z_{n-1} + M_n.$$  

As in [19], we consider $A_n Z_n$, that is we calculate:

$$\mathbb{E}[A_n Z_n | \mathcal{Y}_n].$$

We would then have:

$$\hat{A}(n) = \mathbb{E}[A_n | \mathcal{Y}_n] = \mathbb{E}[A_n (Z_n, 1) | \mathcal{Y}_n] = (\mathbb{E}[A_n Z_n | \mathcal{Y}_n], 1).$$

Using the probability measure $\bar{\mathbb{P}}$, and the conditional version of Baye’s theorem we have:

$$\mathbb{E}[A_n Z_n | \mathcal{Y}_n] = \frac{\mathbb{E}[\Lambda_n A_n Z_n | \mathcal{Y}_n]}{\mathbb{E}[\Lambda_n | \mathcal{Y}_n]]. \quad (5.5.14)$$

For the numerator term, write

$$\sigma_n(A_n Z_n) = \mathbb{E}[\Lambda_n A_n Z_n | \mathcal{Y}_n].$$

We shall find a relation that allows us to recursively calculate $\sigma_n(A_n Z_n)$.

Now,

$$\mathbb{E}[\Lambda_n A_n Z_n | \mathcal{Y}_n] = \mathbb{E}[\Lambda_n \left(A_{n-1} + \frac{D_{n-1} Y_n}{K_{n-1}^2} \langle Z_{n-1}, e_j \rangle \right) Z_n | \mathcal{Y}_n]
\quad (5.5.15)$$

Consider the first term in the right hand side of the above equality. Note that for all $n \geq 1$:

$$\sum_{i=1}^{N} \langle Z_{n-1}, e_i \rangle = 1.$$

Write

$$\Gamma_n^i = 1 = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} \left( \frac{Y_n - H_n - D_{n-1} \langle \hat{\mu}, e_i \rangle}{K_{n-1}} \right)^2 \right).$$
\[ \mathbb{E}[\Lambda_n A_{n-1} Z_n | \mathcal{Y}_n] = \mathbb{E}[\Lambda_{n-1} \lambda_n A_{n-1} (\Pi Z_{n-1} + M_n) (\sum_{i=1}^{N} \langle Z_{n-1}, e_i \rangle) | \mathcal{Y}_n] \]

\[ = \sum_{i=1}^{N} \mathbb{E}[\Lambda_{n-1} \lambda_n A_{n-1} \Pi Z_{n-1} \langle Z_{n-1}, e_i \rangle | \mathcal{Y}_n] \]

\[ = \sum_{i=1}^{N} \Gamma_i^n \mathbb{E}[\Lambda_{n-1} A_{n-1} \langle Z_{n-1}, e_i \rangle | \mathcal{Y}_n] \Pi e_i \]

\[ = \sum_{i=1}^{N} \Gamma_i^n (\mathbb{E}[\Lambda_{n-1} A_{n-1} | \mathcal{Y}_{n-1}], e_i) \Pi e_i \]

\[ = \prod \text{diag}(\Gamma^n) \sigma_{n-1}(A_{n-1} Z_{n-1}) . \]

In the second equality, we used the fact that \( M \) is a martingale vector and:

\[ \mathbb{E}[M_n | \mathcal{Y}_n] = 0 \]

We also used the fact that under \( \tilde{\mathbb{P}} \), \( y_n \) is an \( i.i.d \) process.

Now the second term:

\[ \mathbb{E}[\Lambda_n \langle Z_{n-1}, e_j \rangle Z_n | \mathcal{Y}_n] \frac{D_{n-1} Y_n}{K_{n-1}^2} . \]

Consider:

\[ \mathbb{E}[\Lambda_n \langle Z_{n-1}, e_j \rangle Z_n | \mathcal{Y}_n] = \mathbb{E}[\Lambda_n \langle Z_{n-1}, e_j \rangle (\Pi Z_{n-1} + M_n) | \mathcal{Y}_n] \]

\[ = \mathbb{E}[\Lambda_{n-1} \lambda_n \langle Z_{n-1}, e_j \rangle \Pi e_j | \mathcal{Y}_{n-1}] \]

\[ = \Gamma_j^n \mathbb{E}[\Lambda_{n-1} A_{n-1} | \mathcal{Y}_{n-1}], e_j) \Pi e_j \]

\[ = \Gamma_j^n (\sigma_{n-1}(Z_{n-1}), e_j) \Pi e_j . \]

(5.5.17)
We can then derive a recursive filter for $\sigma_n(Z_n)$:

$$
\sigma_n(Z_n) = \mathbb{E}[\Lambda_n Z_n | \mathcal{Y}_n] = \mathbb{E}[\Lambda_{n-1} \lambda_n (\Pi Z_{n-1} + M_n) | \mathcal{Y}_{n-1}]
$$

$$
= \mathbb{E}[\Lambda_{n-1} \lambda_n \Pi Z_{n-1} (\sum_{i=1}^{N} \langle Z_{n-1}, 1 \rangle) | \mathcal{Y}_{n-1}]
$$

$$
= \sum_{i=1}^{N} \Gamma^n_i \mathbb{E}[\Lambda_{n-1} \Pi e_i \langle Z_{n-1}, e_i \rangle | \mathcal{Y}_{n-1}]
$$

$$
= \sum_{i=1}^{N} \Gamma^n_i \langle \mathbb{E}[\Lambda_{n-1} Z_{n-1} | \mathcal{Y}_{n-1}], e_i \rangle \Pi e_i
$$

$$
= \sum_{i=1}^{N} \Gamma^n_i \langle \sigma_{n-1}(Z_{n-1}), e_i \rangle \Pi e_i
$$

$$
= \Pi diag(\Gamma) \sigma_{n-1}(Z_{n-1}).
$$

That is, we can calculate the numerator of (5.5.14) given in (5.5.15) as a recursive filter. To calculate the denominator, we have:

$$
\mathbb{E}[\Lambda_n \langle Z_n, 1 \rangle | \mathcal{Y}_n] = \langle \mathbb{E}[\Lambda_n Z_n | \mathcal{Y}_n], 1 \rangle.
$$

(5.5.19)

To obtain the estimate of $\mu_j$, we need to evaluate the two remaining expectations $\hat{B}(n)$ and $\hat{C}(n)$ in equation (5.5.11). In a similar way to the calculations for $\hat{A}(n)$, we develop a recursive filter for $\hat{B}(n)$ and $\hat{C}(n)$.

We can use the calculation results from the $\hat{A}(n)$ term to derive expressions for $\hat{B}(n)$ and $\hat{C}(n)$.

**Corollary 4.** Write

$$
B_n = \sum_{l=1}^{n} \frac{D_{l-1} H_l}{K_{l-1}^2} \langle Z_{l-1}, e_j \rangle \text{ and } C_n = \sum_{l=1}^{n} \frac{D_{l-1}^2}{K_{l-1}^2} \langle Z_{l-1}, e_j \rangle.
$$

Then

$$
\hat{B}(n) = \langle \frac{\sigma_n(B_n Z_n)}{\sigma_n(1)}, 1 \rangle \text{ and } \hat{C}(n) = \langle \frac{\sigma_n(C_n Z_n)}{\sigma_n(1)}, 1 \rangle,
$$

(5.5.20)

where

$$
\sigma_n(B_n Z_n) = \Pi diag(\Gamma^n) \sigma_{n-1}(\Lambda_{n-1} Z_{n-1}) + \frac{D_{n-1} H_n}{K_{n-1}^2} \Gamma^n_j \langle \sigma_{n-1}(Z_{n-1}), e_j \rangle \Pi e_j,
$$

(5.5.21)
and
\[
\sigma_n(C_n Z_n) = \Pi diag(\Gamma^n) \sigma_{n-1}(C_{n-1} Z_{n-1}) + \frac{D_{n-1}^2 H_n}{K_{n-1}^2} \Gamma_j^n \langle \sigma_{n-1}(Z_{n-1}), e_j \rangle \Pi e_j. \tag{5.5.22}
\]

**Proof.** We write:
\[
B_n = B_{n-1} + \frac{D_{n-1} H_n}{K_{n-1}^2} \langle Z_{n-1}, e_j \rangle.
\]
We have:
\[
\hat{B}(n) = \mathbb{E}[B_n | \mathcal{Y}_n] = \langle \mathbb{E}[B_n Z_n | \mathcal{Y}_n], 1 \rangle.
\]
Using the Baye’s theorem:
\[
\mathbb{E}[B_n Z_n | \mathcal{Y}_n] = \frac{\mathbb{E}[\Lambda_n B_n Z_n | \mathcal{Y}_n]}{\mathbb{E}[\Lambda_n | \mathcal{Y}_n]},
\]
and by analogy to the \( \hat{A}(n) \) term, the numerator term is given by:
\[
\sigma_n(B_n Z_n) = \mathbb{E}[\Lambda_n B_n Z_n | \mathcal{Y}_n] = \mathbb{E}[\Lambda_n B_{n-1} Z_n + \Lambda_n \frac{D_{n-1} H_n}{K_{n-1}^2} \langle Z_{n-1}, e_j \rangle | \mathcal{Y}_n] = \mathbb{E}[\Lambda_n B_{n-1} Z_n | \mathcal{Y}_n] + \mathbb{E}[\Lambda_n \frac{D_{n-1} H_n}{K_{n-1}^2} \langle Z_{n-1}, e_j \rangle | \mathcal{Y}_n] = \sum_{i=1}^N \Gamma_i \langle \mathbb{E}[\Lambda_{n-1} B_{n-1} Z_{n-1} | \mathcal{Y}_{n-1}], e_i \rangle \Pi e_i + \frac{D_{n-1} H_n}{K_{n-1}^2} \mathbb{E}[\Lambda_{n-1} \langle Z_{n-1}, e_j \rangle Z_n | \mathcal{Y}_n] = \Pi diag(\Gamma^n) \sigma_{n-1}(\Lambda_{n-1} Z_{n-1}) + \frac{D_{n-1} H_n}{K_{n-1}^2} \Gamma_j^n \langle \sigma_{n-1}(Z_{n-1}), e_j \rangle \Pi e_j. \tag{5.5.23}
\]

Similarly for the \( \hat{C}(n) \) term: We write:
\[
C_n = C_{n-1} + \frac{D_{n-1}^2}{K_{n-1}^2} \langle Z_{n-1}, e_j \rangle.
\]
We have:
\[
\hat{C}(n) = \mathbb{E}[C_n | \mathcal{Y}_n] = \langle \mathbb{E}[C_n Z_n | \mathcal{Y}_n], 1 \rangle.
\]
CHAPTER 5. POWER PRICING USING FILTERING THEORY

Using the Bayes’ theorem:

\[ E[C_nZ_n|\gamma_n] = \frac{E[\Lambda_nC_nZ_n|\gamma_n]}{E[\Lambda_n|\gamma_n]} , \]

and by analogy to the \( \hat{A}(n) \) term, the numerator term is given by:

\[ \sigma_n(C_nZ_n) = E[\Lambda_nC_nZ_n|\gamma_n] = E[\Lambda_nC_{n-1}Z_n + \Lambda_n \frac{D_{n-1}^2}{K_{n-1}^2} (Z_{n-1}, e_j)]|\gamma_n] \]

\[ = E[\Lambda_nC_{n-1}Z_n|\gamma_n] + E[\Lambda_n \frac{D_{n-1}^2}{K_{n-1}^2} (Z_{n-1}, e_j)]|\gamma_n] \]

\[ = \sum_{i=1}^{N} \Gamma_n^i (E[\Lambda_{n-1}C_{n-1}Z_{n-1}|\gamma_{n-1}], e_i) \Pi e_i + \frac{D_{n-1}^2}{K_{n-1}^2} E[\Lambda_n (Z_{n-1}, e_j)Z_n|\gamma_n] \]

\[ = \Pi diag(\Gamma_n^0) \sigma_{n-1}(C_{n-1}Z_{n-1}) + \frac{D_{n-1}^2}{K_{n-1}^2} \Gamma_n^0 \langle \sigma_{n-1}(Z_{n-1}), e_j \rangle \Pi e_j . \]

\[ \square \]

5.5.4 Estimation of the volatility term \( \sigma \)

In our discretised model (5.5.2), the volatility occurs in the terms \( K(l) \) as:

\[ K_l = \sigma e^{\kappa h} \sqrt{h} . \]

Recall that the log –likelihood function is:

\[ \log \Lambda_n = -\sum_{l=1}^{n} \log(K_{l-1}) - \frac{1}{2} \sum_{l=1}^{n} \left( \frac{Y_l - H_l - D_{l-1}(\mu, Z_{l-1})}{K_{l-1}} \right)^2 + \frac{1}{2} \sum_{l=1}^{n} Y_l^2 . \]

Then,

\[ \frac{\partial \log \Lambda_n}{\partial \sigma} = \frac{\partial \log \Lambda_n}{\partial K} \frac{\partial K}{\partial \sigma} \]

\[ = \sum_{l=1}^{n} \left( -\frac{e^{\kappa h} \sqrt{h}}{K_{l-1}} + \left( \frac{Y_l - H_l - D_{l-1}(\mu, Z_{l-1})}{K_{l-1}} \right)^2 \left( K_{l-1}^{-3} e^{\kappa h} \sqrt{h} \right) \right) \]

\[ = \sum_{l=1}^{n} \left( -\frac{1}{\sigma} + \frac{\left( \frac{Y_l - H(l) - D(l)(\mu, Z_l)}{\sigma^2 e^{2\kappa h} h} \right)^2}{\sigma^3 e^{2\kappa h} h} \right) . \]

From the equation:

\[ \frac{\partial \log \Lambda_n}{\partial \sigma} = 0 \]
we obtain an estimate:

\[ \hat{\sigma}^2 = \frac{1}{n} \sum_{l=1}^{n} \frac{(Y_l - H_l - D_{l-1}(\mu, Z_{l-1}))^2}{e^{2\kappa l h}}. \]  

(5.5.24)

To calculate the new estimate of the volatility, we need to evaluate the above sum, which involves the unobserved Markov chain \( Z \). Again, we introduce a filter to calculate its expected value.

**Proposition 16.** The recursive filter estimator of the volatility \( \sigma \) given in (5.5.24) is:

\[ \mathbb{E}[\hat{\sigma}^2 | \mathcal{Y}_n] = \frac{1}{n} \sum_{l=1}^{n} \frac{1}{e^{2\kappa l h}} \left( Y_l^2 + H_l^2 - 2Y_l H_l \right) + \frac{1}{n} \left\langle \frac{\sigma_n(S_n Z_n)}{\sigma_n(1)}, 1 \right\rangle, \]

(5.5.25)

where

\[ \sigma_n(S_n Z_n) = \mathbb{E}[\Lambda_n S_n Z_n | \mathcal{Y}_n] \]

\[ = \Pi \text{diag}(\Gamma^n) (\sigma_{n-1}(S_{n-1} Z_{n-1}) + \text{diag}(F(\mu, n))\sigma_{n-1}(Z_{n-1})) , \]  

(5.5.26)

with \( F(\mu, l) = D_{l-1}^2 \mu^2 + 2(H_l D_{l-1} - Y_l D_{l-1})\mu \), and \( \sigma_n(1) \) is given in (5.5.13).

**Proof.** From (5.5.24) :

\[ \mathbb{E}[\hat{\sigma}^2 | \mathcal{Y}_n] = \mathbb{E} \left[ \frac{1}{n} \sum_{l=1}^{n} \frac{1}{e^{2\kappa l h}} \left( Y_l^2 + H_l^2 - 2Y_l H_l + \langle F(\mu, l), Z_{l-1} \rangle \right) | \mathcal{Y}_n \right] , \]

Here we denote by the product of 2 vectors \( u = (u_1, u_2, \ldots, u_N)' \) and \( v = (v_1, v_2, \ldots, v_N)' \) the point-wise product vector:

\[ uv = (u_1 v_1, u_2 v_2, \ldots, u_N v_N)' \]

Then:

\[ \mathbb{E}[\hat{\sigma}^2 | \mathcal{Y}_n] = \frac{1}{n} \sum_{l=1}^{n} \frac{1}{e^{2\kappa l h}} \left( Y_l^2 + H_l^2 - 2Y_l H_l \right) + \frac{1}{n} \mathbb{E} \left[ \sum_{l=1}^{n} \frac{1}{e^{2\kappa l h}} \langle F(\mu, l), Z_{l-1} \rangle | \mathcal{Y}_n \right] . \]
The second term in the right hand side of the above equality can be calculated similarly to the \( \hat{A}(n) \) term in the estimation of the drift \( \mu \). Write:

\[
S_n = \sum_{l=1}^{n} \frac{1}{e^{2\kappa nh}} \langle F(\mu, l), Z_{l-1} \rangle
= S_{n-1} + \frac{1}{e^{2\kappa nh}} \langle F(\mu, n), Z_{n-1} \rangle .
\]

Again, we consider \( S_n Z_n \) and calculate:

\[
E[S_n Z_n | \mathcal{Y}_n] .
\]

We then have:

\[
E[S_n | \mathcal{Y}_n] = E[S_n \langle Z_n, 1 \rangle | \mathcal{Y}_n] = \langle E[S_n Z_n | \mathcal{Y}_n], 1 \rangle .
\]

Using the probability measure \( \bar{P} \), and the conditional version of Bayes’ theorem we have:

\[
E[S_n Z_n | \mathcal{Y}_n] = \frac{E[\Lambda_n S_n Z_n | \mathcal{Y}_n]}{E[\Lambda_n | \mathcal{Y}_n]} . \tag{5.5.27}
\]

We have previously evaluated the denominator term. For the numerator term write:

\[
\sigma_n(S_n Z_n) = E[\Lambda_n S_n Z_n | \mathcal{Y}_n] .
\]

We shall find a recursive relation that allows us to calculate \( \sigma_n(S_n Z_n) \). We have:

\[
\sigma_n(S_n Z_n) = E[\Lambda_n S_n Z_n | \mathcal{Y}_n]
= E[\Lambda_n (S_{n-1} + \frac{1}{e^{2\kappa nh}} \langle F(\mu, n), Z_{n-1} \rangle) Z_n | \mathcal{Y}_n]
= E[\Lambda_n S_{n-1} Z_n | \mathcal{Y}_n] + \frac{1}{e^{2\kappa nh}} E[\Lambda_n \langle F(\mu, n), Z_{n-1} \rangle Z_n | \mathcal{Y}_n] .
\]

The first term in the right hand side can be derived by analogy from (5.5.16), and is given by:

\[
E[\Lambda_n S_{n-1} Z_n | \mathcal{Y}_n] = \sum_{i=1}^{N} \Gamma^n_i \langle \sigma_{n-1}(S_{n-1} Z_{n-1}), e_i \rangle \Pi e_i = \Pi diag(\Gamma^n) \sigma_{n-1}(S_{n-1} Z_{n-1}) .
\]
For the second term, we have:

\[
\mathbb{E}[\Lambda_n(F(\mu, n), Z_{n-1})Z_n|Y_n] = \mathbb{E}[\Lambda_{n-1}\Lambda_n(F(\mu, n), Z_{n-1})(\prod_{i=1}^{N}(Z_{n-1}, e_i))|Y_n]
\]

\[
= \sum_{i=1}^{N} \mathbb{E}[\Lambda_{n-1}\Lambda_n(F(\mu, n), e_i)\Pi E_i(Z_{n-1}, e_i)|Y_n]
\]

\[
= \sum_{i=1}^{N} \Gamma_i^n \mathbb{E}[\Lambda_{n-1}Z_{n-1}, e_i|Y_{n-1}](F(\mu, n), e_i)\Pi E_i
\]

\[
= \sum_{i=1}^{N} \Gamma_i^n \mathbb{E}[\Lambda_{n-1}Z_{n-1}|Y_{n-1}](e_i|F(\mu, n), e_i)\Pi E_i
\]

\[
= \prod_{i} \text{diag}(\Gamma_i^n) \text{diag}(F(\mu, n)) \sigma_{n-1}(Z_{n-1}).
\]

The denominator of (5.5.27) can be calculated from (5.5.19).

5.5.5 Estimation of the speed of return $\kappa$

For the estimation of the parameter $\kappa$, since it does not depend on the Markov process $Z$, one can use a straightforward naive estimation. One commonly used method for Ornstein-Uhlenbeck process is the least square minimisation. Recall from (5.3.2) that the discretised observed process can be written:

\[
y_t = \int_0^t a_u du + \int_0^t b_u \langle \mu, Z_u \rangle du + \sigma \int_0^t e^{\kappa u} dW_u,
\]

where

\[
a_u = \frac{d}{du}(e^{\kappa u} \ln(f_u))
\]

and

\[
b_u = \kappa e^{\kappa u}.
\]

Then

\[
y_{n+1} = y_n + e^{\kappa n}(\log f_n + \kappa \mu)h + \sigma e^{\kappa n} \sqrt{h}\epsilon.
\]

Where $\epsilon \sim \mathcal{N}(0, 1)$. Applying a linear regression, we obtain the equations with the unknown $\kappa$:

\[
\text{sd}(\epsilon) = \sqrt{\frac{nS_{xy} - S_y^2 - nS_{xy} - S_x S_y}{n(n-2)}} = \sigma e^{\kappa n} \sqrt{h},
\]

(5.5.30)
and
\[
\frac{S_y - S_x}{n} = e^{\kappa n} (\log f_n + \kappa \mu) h, \tag{5.5.31}
\]
where
\[
S_x = \sum_{i=1}^{n} y_{i-1}, \quad S_y = \sum_{i=1}^{n} y_i, \quad S_y^2 = \sum_{i=1}^{n} y_i^2
\]
are calculated from the observed process \( y \). The estimator \( \hat{\kappa} \) can then be obtained numerically. An average value of the estimators over different sequences of data is taken to compensate the regime switching effect.

### 5.5.6 Estimation of the transition rate matrix of the process

We consider the matrix \( \Pi = hA + I \). Using the results from [19], write \( N_{rs}^n \) for the number of jumps from state \( r \) to state \( s \), and \( J_r^n \) for the occupational time of the state \( r \) up to the time step \( n \). Then the estimates of the entries of \( \Pi = (\pi_{sr})_{1 \leq s,r \leq N} \), where \( \pi_{sr} \) is the probability of switching from state \( r \) to state \( s \), are given in the following result:

**Proposition 17.** The recursive filter estimate of the entries of the transition rate matrix, \( \hat{\pi}_{sr} \) is given by:
\[
\forall 1 \leq s, r \leq N : \hat{\pi}_{sr} = \frac{\langle \sigma_n(N_{rs}^n Z_n), 1 \rangle}{\langle \sigma_n(J_r^n Z_n), 1 \rangle}, \tag{5.5.32}
\]
where
\[
\sigma_n(N_{rs}^n Z_n) = \Pi \text{diag}(\Gamma^n)\sigma_{n-1}(N_{rs}^{n-1} Z_{n-1}) + \Gamma^n \mathbb{E}[\Lambda_{n-1}(Z_{n-1}, e_r)\pi_{sr} \Pi e_r | Y_n] \tag{5.5.33}
\]
and
\[
\sigma_n(J_r^n Z_n) = \Pi \text{diag}(\Gamma^n)\sigma_{n-1}(J_{r}^{n-1} Z_{n-1}) + \Gamma^n \mathbb{E}[\sigma_{n-1}(Z_{n-1}, e_r)\Pi e_r | Y_n]. \tag{5.5.34}
\]

**Proof.** The transition probability estimator is, from [19], given by:
\[
\hat{\pi}_{sr} = \frac{\mathbb{E}[N_{rs}^n | Y_n]}{\mathbb{E}[J_r^n | Y_n]}.
\]
Using the conditional version of Bayes’ theorem, write:

\[ \hat{\pi}_{sr} = \frac{\mathbb{E}[\Lambda_n N_{rs}^n | Y_n]}{\mathbb{E}[\Lambda_n J_{nr}^r | Y_n^r]}. \]  

(5.5.35)

We have:

\[ N_{rs}^n = \sum_{l=1}^{n} \langle Z_l, e_r \rangle \langle Z_{l-1}, e_r \rangle \]

\[ = N_{rs}^{n-1} + \langle Z_{n-1}, e_r \rangle \langle Z_n, e_s \rangle. \]

Define:

\[ \sigma_n(N_{rs}^n Z_n) = \mathbb{E}[\Lambda_n N_{rs}^n Z_n | Y_n]. \]

Now:

\[ \sigma_n(N_{rs}^n Z_n) = \mathbb{E}[\Lambda_n \lambda_n (N_{rs}^{n-1} + \langle Z_{n-1}, e_r \rangle \langle Z_n, e_s \rangle) Z_n | Y_n] \]

\[ = \mathbb{E}[\Lambda_n \lambda_n N_{rs}^{n-1} | Y_n] + \mathbb{E}[\Lambda_n \lambda_n \langle Z_{n-1}, e_r \rangle \langle Z_n, e_s \rangle Z_n | Y_n]. \]

For the first term on the right hand side,

\[ \mathbb{E}[\Lambda_n \lambda_n N_{rs}^{n-1} Z_n | Y_n] = \mathbb{E}[\Lambda_n \lambda_n N_{rs}^{n-1} \sum_{i=1}^{N} \langle Z_{n-1}, e_i \rangle (\Pi Z_{n-1} + M_n) | Y_n] \]

\[ = \sum_{i=1}^{N} \Gamma_i^n(n) \mathbb{E}[\Lambda_n \lambda_n N_{rs}^{n-1} Z_{n-1} e_i | Y_{n-1}] \]

\[ = \sum_{i=1}^{N} \Gamma_i^n(n) \mathbb{E}[\Lambda_n \lambda_n N_{rs}^{n-1} Z_{n-1} | Y_{n-1}, e_i] \Pi e_i \]

\[ = \Pi \text{diag}(\Gamma^n) \sigma_{n-1}(N_{rs}^{n-1} Z_{n-1}). \]

For the second term:

\[ \mathbb{E}[\Lambda_n \lambda_n \langle Z_{n-1}, e_r \rangle \langle Z_n, e_s \rangle Z_n | Y_n] = \Gamma_r^n \mathbb{E}[\Lambda_n \lambda_n \langle Z_{n-1}, e_r \rangle \langle Z_n, e_s \rangle Z_n | Y_n] \]

\[ = \Gamma_r^n \mathbb{E}[\Lambda_n \lambda_n \langle Z_{n-1}, e_r \rangle \langle (\Pi Z_{n-1} + M_n), e_s \rangle | (\Pi Z_{n-1} + M_n) | Y_n] \]

\[ = \Gamma_r^n \mathbb{E}[\Lambda_n \lambda_n \langle Z_{n-1}, e_r \rangle \pi_{sr} e_r | Y_n] \]

\[ = \Gamma_r^n \mathbb{E}[\Lambda_n \lambda_n \langle Z_{n-1}, e_r \rangle \pi_{sr} e_r | Y_n] \]

\[ = \Gamma_r^n \sigma_{n-1}(Z_{n-1}), e_r \pi_{sr} e_r \].
The estimate for the number of jumps in (5.5.35) is:

\[ \mathbb{E}[\Lambda_n N^{rs}_n | \mathcal{Y}_n] = \langle \mathbb{E}[\Lambda_n N^{rs}_n Z_n | \mathcal{Y}_n], 1 \rangle. \]

For the estimates of the occupation time, we have:

\[ J^n_r = \sum_{l=1}^{n} \langle Z_{l-1} e_l \rangle = J^{n-1}_r + \langle Z_{n-1}, e_r \rangle \]

so:

\[ \sigma_n(J^n_r Z_n) = \mathbb{E}[\Lambda_n J^n_r Z_n | \mathcal{Y}_n] \]

\[ = \mathbb{E}[\Lambda_{n-1} \lambda_n (J^{n-1}_r + \langle Z_{n-1}, e_r \rangle)(\Pi Z_{n-1} + M_n) | \mathcal{Y}_n] \]

\[ = \sum_{i=1}^{N} \Gamma_i^n \langle \mathbb{E}[\Lambda_{n-1} J^{n-1}_r Z_{n-1} | \mathcal{Y}_{n-1}], e_i \rangle \Pi e_i + \Gamma_r^n \mathbb{E}[\Lambda_{n-1} \langle Z_{n-1}, e_r \rangle \Pi Z_{n-1} | \mathcal{Y}_n] \]

\[ = \Pi diag(\Gamma^n) \sigma_{n-1}(J^{n-1}_r Z_{n-1}) + \Gamma_r^n \langle \sigma_{n-1}(Z_{n-1}), e_r \rangle \Pi e_r. \]

So the estimate of the smoother for the number of jumps in (5.5.35) is obtained by:

\[ \mathbb{E}[\Lambda_n J^n_r | \mathcal{Y}_n] = \langle \mathbb{E}[\Lambda_n J^n_r Z_n | \mathcal{Y}_n], 1 \rangle. \]

\[ \square \]

### 5.6 Empirical analysis and simulation

The electricity price database used here is the South Australian day-ahead half-hourly spot price for the period of three years from January 2010 to December 2012. Figure 5.6.1 is a normal scaled plot of the electricity spot price time series. We note here the magnitude of spikes that characterise this market. The extreme jumps in the spot price explain the incorporation of a jump process in our model as a Markov regime switching factor and a regime switching mean reverting Ornstein-Uhlenbeck process.

It is well known that power price time data plots accommodate common features summarised in three components: seasonality, mean reversion and jumps. The
South-Australian power prices in Figure 5.6.1, show very high spikes reaching price returns of 3000%, and in a normal scale such big jumps veil the seasonality effect as well the mean reversion. In figure 5.6.3 the logarithm of price is plotted and we can clearly notice the seasonality and the mean reversion of the price.

Figure 5.6.1: South Australian Electricity Spot Price

The definition of price spikes is not unique, but it is clear that extreme changes in the spot price in terms of frequency of occurrence and the size will abort the assumption of price return having a normal or log-normal distribution, Figure 5.6.2 is the normal probability test of the logarithm of the spot prices and shows a clear deviation of the data from the normal distribution, consequently the presence of heavy tail distribution and the requirement of introducing a jump process in the spot price dynamics in order to well simulate the South Australian power price. More details on price spikes filtering and analysis are given in subsection 5.6.2.

To capture other features that the spot price empirically exhibits, we can damp the scale size of the spikes by looking at the logarithm of prices. One common feature of the electricity markets is that prices turn negative few times over the year, more details and discussions on this peculiarity is given in subsection 5.6.3. For a negative argument a logarithm is not defined, so just for a graphing purpose,
we use a reflexion principle and define the log-price as:

$$\log(S_t) = \text{sign}(S_t) \log(|S_t|).$$

Other techniques used in the literature are given in subsection 5.6.3.
5.6.1 Seasonality

Figure 5.6.3 is a plot of the log-price for the period of Jan 2010 - Dec 2013. We can see a periodicity pattern, which is proven empirically due to the offer and demand seasonal fluctuations [57, 29, 46]. Spot price seasonality mostly arise due to changes in weather conditions and climate, as well as to daylight time length, in this case the demand is the main driver of electricity in reaching the market equilibrium price. The power supplies can also follow seasonal variations in output, and this occurs in some power generation technologies, mainly renewable one, where the later depend on climate and weather conditions.

It is essential in power market modelling to determine the seasonality function that will best fit the data seasonality. There are different approaches to model seasonal pattern depending on the nature of the data in question. In the case of noticeable difference in price from winter to summer, sinusoidal function are usually suitable, although there is no unique choice on the form of the function. Some authors [57, 44, 28] model the seasonality by a deterministic function, having a deterministic linear component and a sinusoidal function, this type of models prove efficiency in the case where data shows a clear yearly or monthly spot price seasonality. Other authors like in [15, 47] use constant piece-wise functions, changing values according to the hour, day and year as dummy variable.

In this work, from the data base in question, we choose to use the same seasonality model used in Bierbrauer [27], which is a hybrid approach taking into account both the yearly seasonality through a sinusoidal function as well the weekly and daily seasonality modelled by the piece-wise constant dummy variable:

\[
f(t) = \exp[\delta + \beta t + d\Delta + \epsilon \sin((t + \varphi)\frac{2\pi}{365})],
\]

(5.6.1)

where \(\delta, \beta, d, \epsilon\) and \(\varphi\) are constant parameters to be calibrated, and \(\Delta\) is a dummy variable taking values 1, 2, \ldots, 7 to include the weekly seasonality in spot price.
5.6.2 Identifying the spikes

Before simulating seasonality, one needs to identify and filter the spot price spikes. As our model suggests, the price spikes are unpredictable but observable process. The identification of spikes is however essential in the model calibration, and can influence the estimation of the deterministic and stochastic components of the model. There is no unique method to select the values determining the spikes, but the latter are defined when the price exceeds some threshold over a brief period of time. Different methods are suggested in the literature for identifying the spikes, Janczura [44] presents a literature review on different methods. Once the spikes are identified and filtered, one can proceed to estimate the seasonality components, the curve fitting with a least square is a favourite candidate technique for this type of data.

5.6.3 Negative prices

In this paper, we case-study the Australian power market, and we chose to take the South Australian market as a sample set because the latter is recognised as one of the most volatile power markets worldwide. This represents a challenge in terms of capturing different aspects of the market. The jumps in price seen as spikes is a usual property, but in our case these jumps are negative drops. From a technical perspective, a negative price means that there is more benefit in the destruction of the commodity rather than its creation. This usually happens during the night, when power demand is lower than supply. This imbalance is usually caused by the must-run character of the non-flexible generators. This makes shut-down hardly possible or involves high costs. These price drops are results of demand shortage, and are included in our model through the regime switching factor $\langle \alpha, J_t \rangle$, which can take negative values. Different techniques are used to solve the negative prices issue. Some authors solved the problem by an affine transformation of all prices into the positive range, so uplifting the power prices by a constant level will not affect
the dynamics of our model.

### 5.6.4 Sample path simulation

The main difficulty in modelling the electricity prices is how can the model mimic the price spikes. Some authors like in [50], [12] and [25] are modelling the spikes with a mean reverting pure jump diffusion dynamics. For these models to exhibit typical spikes the mean reversion rate must be extremely high, otherwise the jumps do not revert quickly enough. Our model simulates jumps using Markov processes to let the price change from a base load regime 'off-peak' to peak-load and similarly here, the return to base load is assured by means of the rate matrix entries. Figure 5.6.4 is a realisation of a simulated path of our model, we chose to 3 regimes for the price, or the jump factor $\alpha$, peak-load, base load and negative prices.

![Sample path](image)

Figure 5.6.4: Simulated log-prices realisation

### 5.7 Conclusion

In this chapter we present a new model for pricing power. We introduce a new way of modelling price spikes using a continuous time Markov jump process. In addition to Markov switching jump sizes, we introduce another independent, but unobservable, Markov jump process following which the long term mean switches from one state to another. We start from this spot price model to price a forward contract; this is
performed under a risk neutral measure determined using a risk premium principle. Some difficulty occurs when we wish to calculate the expectation of a quantity which depends on the unobservable Markov process. To overcome this issue a recursive filtering technique is derived and we then present an approximation of the forward price. The latter is used to price a call option on forward.

A major part of this chapter is dedicated to develop the estimators for the model parameters, again using filtering theory to overcome the issue of the non observable Markov jump process.

We conclude the chapter with some data analysis and a simulation of the spot price model as a background for the suggested model. Further investigations could focus on numerical work to calibrate the model parameters.
Chapter 6

Outlook and extensions

This thesis demonstrates the extent to which Markov regime switching models can be used to price derivatives with electricity as the underlying, taking into account different aspects characterising power prices: mean reversion, jumps and seasonality.

Using a regime switching stochastic discount factor is one way to treat the non-uniqueness of the risk neutral measure, which is usually the case when modelling power price randomness. Despite the complexity rising from the additional Markov jump process in the discount factor, we found it is mathematically tractable to price different financial instruments after giving a semi-closed formulas. We assume that the stock price is a derivative asset, and is the expected present value of future dividends. This is exploited under regime switching Black and Scholes dynamics for the underlying. This requires a natural condition on the model parameters. Further research could focus on analysing the limits of this condition. The results obtained from our model are analytically coherent; however, the model calibration and its relation with real world data merits further study.

The combination of the stochastic discount factor with a regime switching Ornstein-Uhlenbeck process enables the calculation of the price of forward contracts as a closed formula. However, the resulting call option price is only semi-analytical and can be seen as a modification of the Black and Scholes call price with Markov process de-
A second process used to price power, and showing positive improvements over previous literature, is the compensated pure jump process. We consider a general form of such a process and define a risk neutral measure consistent with the market spot price and interest rate. This enables us to price the forward and swap contract. This model can easily be adapted to different pure jump process depending on the availability of suitable data bases.

The final section of the thesis presents a different model for power pricing. An independent Markov chain factor models jump sizes, in addition to a non observable and independent Markov process acting on the long term mean. We then develop a recursive filter to price forwards and options on forwards. The hidden Markov process constrains us to use some approximations, and numerical solutions are required to evaluate derivative prices. Estimators of parameters are also developed which enable calibration from discrete time data. An interesting aspect to investigate would be to consider stochastic jump sizes, so introducing an infinite state Markov chain. Further research could focus on the numerical efficiency of the analytical results we have obtained and back test their consistency with market behaviour.
Appendix A

Probability and Stochastics

Appendix

A.1 The exponential of a matrix

Consider the systems of linear, constant coefficient ordinary differential equations:

$$\frac{dx(t)}{dt} = Ax(t).$$

Here $A$ is a fixed $n$-by-$n$ matrix. In principle, the solution is given by

$$x(t) = e^{tA}x_0$$

where $e^{tA}$ can be defined by the convergent power series:

$$e^{tA} = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$ 

A.2 The Fourier transform of a function or a measure

Consider the function $f : \mathbb{R} \to \mathbb{C}$ such that $f \in L^1$ and the measure $\mu$ on $\mathbb{R}$ with finite total mass $|\mu| < \infty$. Then the Fourier transform of the function $f$ and $\mu$ are
respectively given by:

\[ \hat{f}(u) = \int_{\mathbb{R}} e^{iux} f(x) \, dx \]

\[ \hat{\mu}(u) = \int_{\mathbb{R}} e^{iux} \mu(dx) . \]

Note that the characteristic function of a random variable is plainly is Fourier transform.

### A.3 Bayes’ rule for conditional expectation

**Theorem 5.** Let \( \mathbb{P} \) and \( \bar{\mathbb{P}} \) be two equivalent probability measures defined on a state space \( \Omega \), and let the martingale process \( \Lambda \) defined on the filtered probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) be the Radon-Nikodym derivative such that:

\[ \frac{d \bar{\mathbb{P}}}{d \mathbb{P}} |_{\mathcal{F}_t} = \Lambda_t . \]  

(A.3.1)

The process \( \{\Lambda_t\}_{t \geq 0} \) satisfies:

\[ \forall t \in \mathbb{R}^+ : \mathbb{E}[\Lambda_t | \mathcal{F}_t] = 1 . \]

Then for every \( \mathcal{F} \)-measurable process \( X \):

\[ \forall 0 \leq s \leq t : \mathbb{E}[X_t | \mathcal{F}_s] = \frac{\mathbb{E}[X_t \Lambda_t | \mathcal{F}_s]}{\mathbb{E}[\Lambda_t | \mathcal{F}_s]} = \frac{\mathbb{E}[X_t \Lambda_t | \mathcal{F}_s]}{Z_s} . \]  

(A.3.2)

### A.4 Lévy’s characterisation of Brownian motion

This result in probability measure theory enables to determine if a martingale process is Brownian motion in a filtered probability space.

**Theorem 6.** Let \( M \) be a local martingale with \( M_0 = 0 \). Then the following statements are equivalent:

1. \( M \) is standard Brownian motion on the underlying filtered probability space.
2. $M$ is continuous and $M_t^2 - t$ is a local martingale.

3. $M$ has quadratic variation $[M]_t = t$.

This result extends to a $d$-dimensional Brownian motion.

### A.5 Compensated jump process

If $N$ is a right continuous jump process with finite variation, then $N$ can be written, see [43, 18], as:

$$\forall t \geq 0, N_t = \sum_{s \leq t} \Delta N_s = \int_0^t \int_{\mathbb{R}} x \gamma(dx, du),$$

where $\Delta N_s = N_{s+} - N_{s-}$ and $\gamma$ is a random measure which selects the random times and random sizes of the jumps. Let $\mathcal{F}_t^N$ be the natural filtration of the process $N$ where $\mathcal{F}_t^N = \sigma(N_s, s \leq t)$. Then the compensator of the process $N$ or the predictable projection of the jump measure, is the unique $\mathcal{F}_t^N$-predictable process $\tilde{\gamma}$ such that:

$$M_t = \int_0^t \int_{\mathbb{R}} x \gamma(dx, du) - \int_0^t \int_{\mathbb{R}} x \tilde{\gamma}(dx, du)$$

(A.5.1)

is a $\mathcal{F}_t^N$-local martingale and $M$ is referred to as compensated jump process. We note that under the historical probability measure, all the statistical properties of the jump process are determined by its compensator.

**Remark 10.** For a standard Poisson counting process the intensity is a deterministic function of time $\lambda(s)$. The compensator in this case is $\lambda t$, and the process

$$(N_t - \int_0^t \lambda(s) ds)_{t \geq 0}$$

(A.5.2)

is an $\mathcal{F}_t^N$-martingale.

### A.6 Stochastic Fubini theorem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete filtered probability space, and $X$ be a semi-martingale. If $\mu$ is a bounded measure on $\mathbb{R}$ and $f$ is a bounded measure map on $\mathbb{R}^+ \times \Omega \times \mathbb{R}$,
then:

\[
\forall t \in \mathbb{R}^+ : \int_0^t dX_s \left( \int \mu(dx)f(s, \omega, x) \right) = \int \mu(dx) \left( \int_0^t dX_s f(s, \omega, x) \right) . \tag{A.6.1}
\]
Appendix B

List of Used Notations and Symbols

\( \mathbb{R}^N \)  \( N \)-dimensional Euclidian space

\( \text{sign}(x) \) \[
\begin{cases} 
1 & \text{if } x \geq 0 \\
-1 & \text{if } x < 0 
\end{cases}
\]

SDF Stochastic Discount Factor

\( \langle u, v \rangle \) Euclidienn vectorial inner product \( \sum_{i=1}^{N} u_i v_i \)

\( \text{diag}(u) \) The diagonal matrix where the diagonal entries are the entries of the vector \( u \)

\( \mathbb{E}^* \) The expectation under the risk neutral measure \( \mathbb{P}^* \)

\( \mathcal{F} \lor \mathcal{G} \) The smallest filtration including the filtrations \( \mathcal{F} \) and \( \mathcal{G} \)

\( \bar{u} \) The conjugate complex vector of the vector \( u \)

\( P' \) The transpose matrix of \( P \)

\( \ker(A) \) The kernel space of the linear application defined by the matrix \( A \)

\( \hat{f} \) The Fourier transform of the function \( f \)

\( PV \int_R f \) The principal value of the integral := \( \lim_{x \to +\infty} \int_M^M f \)

OU process Ornstein Uhlenbeck process

\( \mathcal{N} \) The normal distribution function

\( \langle X \rangle_t \) The quadratic variation of the process \( X \)
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{\gamma}$</td>
<td>The predictable process of the jump measure $\gamma$</td>
</tr>
<tr>
<td>$f(s,t)$</td>
<td>The price of a forward contract written at time $s$ for settlement at time $t$</td>
</tr>
<tr>
<td>$I(\chi \leq x)$</td>
<td>The indicator function of the set ${\omega : \chi(\omega) \leq x}$ for $\omega \in \Omega$</td>
</tr>
<tr>
<td>$\sigma_n(X)$</td>
<td>The recursive filter of the process $X$</td>
</tr>
<tr>
<td>$p_{ij}$</td>
<td>The probability of transition from state $j$ to state $i$</td>
</tr>
<tr>
<td>$sd(\epsilon)$</td>
<td>The standard deviation of the random variable $\epsilon$</td>
</tr>
</tbody>
</table>
Bibliography


