Emad Imreizeeq

Parameter Estimation of a New Energy Spot Model from Futures Prices

Parameter Estimation of a New Energy Spot Model from Futures Prices

Graduation Committee:

prof. dr. ir. A. J. Mouthaan (Chairman), University of Twente, EWI
prof. dr. A. Bagchi (First Promotor), University of Twente, EWI
prof. dr. S. Aihara (Second Promotor), Tokyo University of Science, Japan
dr. D. Dupont, (Referent), RWE, Germany
prof. dr. A. Stoorvogel, University of Twente, EWI
prof. dr. ir. M. Vellekoop, University of Amsterdam
dr. J.A.M. van der Weide, TU Delft
prof. dr. ir. M. Wouters, University of Twente, MB

Parameter Estimation of a New Energy Spot Model from Futures Prices

DISSERTATION

to obtain
the doctor's degree at the University of Twente,
on the authority of the rector magnificus,
prof. dr. H. Brinksma,
on account of the decision of the graduation committee,
to be publicly defended
on Friday 18 March 2011 at 14:45 hours

by

Emad Saleh Naser Imreizeeq
born on 19$^{\text{th}}$ July 1968
in Jerusalem

Dit proefschrift is goedgekeurd door de promotors
Prof. dr. A. Bagchi
Prof. dr. S. Aihara

This thesis is dedicated to
my parents

Contents

Chapter 1

Introduction

1.1 Energy Industry

The beginning of 1990s was the start of the liberalization of electricity and gas markets, resulting in the emergence of markets for corresponding spot and derivative products in numerous countries and regions all over the world. These assets are in many ways distinct in nature from the more "classical" commodity markets as oil, coal, metals and agriculture. One of the main characteristics of electricity is the existence of seasonality and spikes in prices and limited storage possibilities. As a result of the latter, the electricity markets tend to be regional. This means that different prices of electricity between two regions, do not necessarily imply an arbitrage opportunity. An arbitrageur cannot buy for storage and transportation, and therefore the spot asset cannot be used to set up dynamic hedging strategies exploiting price differentials.

Those features are also shared by temperature and natural gas markets. Temperature is obviously not possible to store. Natural gas on the other hand, can be stored, but most often it is quite costly. Benth *et al.* (2008) refer to these two markets as the *related markets* of electricity, because they share similarities with the electricity market from a modeling point of view.

1.1.1 Energy markets

The main power exchanges in Europe include Nord Pool (Norway, Sweden, Denmark and Finland), EEX (Germany), Powernext (France), APX (Netherlands, Belgium, and UK), Endex (Netherlands), Belpex (Belgium),

OMEL (Spain) and GME (Italy). With regard the gas markets, they are located around different hubs, which are connection and arrival points for gas transportation systems and where there are infrastructure capabilities like, for instance, storage and a concentration of buyers and sellers. Two important hubs are Henry Hub located in Louisiana (US) at the Gulf of Mexico and the National Balancing Point (NBP) in the UK. The latter is a notional hub without any physical location, where all UK gas flows through.

The market for short-term delivery of gas or electricity is usually referred to as the spot market, and the trading is mostly over-the-counter OTC. Futures contracts on both markets ensure the delivery of the underlying over longer time periods like weeks, months, quarters, or even years. Contracts on Nord Pool are not traded during the delivery period, and market participants typically close their position prior to the start of the delivery period.

1.1.2 Energy contracts

A portfolio of energy commodities typically consists of different trading products. As the spot is not tradable asset in the electricity, gas or temperature markets, derivatives such as futures and forwards on the spot are introduced for both trading and risk management. A forward contract is an agreement between two parties for delivery of the underlying commodity at a predetermined date at a predetermined price. A different delivery date leads to a different forward price. The relation between the delivery date and the forward price is called the forward curve. The spot price contract can be viewed as a forward contract with immediate delivery.

Closely related to the forward contracts are the futures contracts. There are some small differences in payment streams and other details in the contract, but for practical purposes, the two contracts can be regarded similar. In fact, the price of both is the same when the interest rate is assumed constant. Analysis of the behavior of these prices, which we will go through later within this chapter, shows that electricity and gas is a seasonal commodity, the change of seasons are a strong driver of the prices. Besides this structural price level, there is a difference between having an amount of gas at this moment and receiving it after a given time period. When a pipeline breaks or another event happens that makes the provision of gas impossible, there is an immediate shortage and prices are likely to increase on such an occasion. This typically only affects the spot market,

but not the long-term prices. One can profit from this situation with high prices only, if the alternative gas is readily available, for example from a storage facility.

Other trading instruments are for example fixed-for-floating swaps and options. Under a swap contract, one party pays a fixed price to a second party, who pays a market dependent price in return. A forward is an example of a simple swap, usually a swap consists of more than one payment stream. An options contract gives the buyer the right (not the obligation) to buy or sell a certain volume at a predetermined price and at a predetermined time. The biggest difference between a forward and an option is the pay-out at maturity. Under the forward contract, you are obliged to buy or sell at the negotiated price, even if it is far above or below the then prevalent market price. With regard to the option contract, as is captured in the name, the owner of an option can decide at maturity if he wishes to exercise the right to buy or sell or not. In the energy markets, the most developed instruments are forwards/futures and swaps. These contracts are often traded and the markets are liquid. In the remainder of this thesis, the main focus will be on forwards and futures.

1.1.3   Day ahead (DA) spot market

In this section we briefly describe the structure of the spot markets. In particular, the APX and EEX as examples of the electricity market and the UK-GAS-NBP as an example of the gas market. The spot market is 24-hour ahead market, which means that every day an auction takes place based on the bidding from buyers and sellers of electricity, and around 12-AM, prices for each hour of the next day are quoted, whence the term (DA) price is used. The spot price is an equilibrium price of demand and supply determined by market players, such as production and distribution companies, large consumers, brokers and traders. Thus the electricity spot market is not the same as in classical definition of spot market of some commodity, where delivery is carried out immediately. The hourly instruments are subject to physical delivery of electricity of a constant output on the electricity grid.

As a result of non-storability of electricity the immediate delivery of electricity is possible only in exceptional cases. On EEX and APX mostly hourly contracts are traded, but also the half-hourly contracts on APX are available. Since the contracts are settled against hourly (DA) prices, the

underlying amount of electrical energy is determined by

$$DP \times 24 \,\mathrm{MWh},$$

with DP being the delivery period measured in days. To be able to compare contracts with different delivery periods, prices are listed in Euros for 1 MWh of power delivered as a constant flow during the delivery period. In the Gas market, the energy content of gas is measured in units of "therms" or British thermal units Btu. By definition, there are 100,000 Btu in 1 therm, whereas 1 therm is the equivalent of 105.5 MJ. Since there are 3.6 GJ per MWh, we have the relation

$$1 \,\mathrm{therm} = 105.5 \,\mathrm{MJ} \times \frac{\mathrm{MWh}}{3.6 \times 1000 \,\mathrm{MJ}} \approx 0.029 \,\mathrm{MWh}.$$

Although the spot prices reflect only physical contracts, they are also underlying bases for many derivatives on the electricity market, which could be either with physical or financial delivery. See Figure 1.1, which shows a time series of average daily prices of APX and EEX electricity markets, respectively. Figure 1.2 shows a time series of (DA) prices of UK-Gas-NBP.

Figure 1.1: Up: APX daily spot prices in euros per MWh, Jan 2001 - Dec 2004. Down: EEX daily spot prices in euros per MWh, Jun 2000 - Dec 2005.

Figure 1.2: UK-Gas-NBP daily spot prices in euros per therm, Jan 2007 - April 2008

1.1.4 Characteristics of the DA spot prices

The spot prices dynamics are stochastic, for that we use stochastic factors to represent this dynamics. However, spot prices can contain some outliers in the form of large positive or negative prices. Hence, it is reasonable first, to remove the outliers from the data. The next step is to specify the deterministic components such as trend and seasonality. Typically, these deterministic components are absorbed into the stochastic factor model through a deterministic function.

Removing the Spikes

Looking at Figure 1.1 and Figure 1.2, we see from the plots that there may be some outliers present in the data. To detect the possible outliers, we analyze daily changes in the logarithm of the spot prices. See Figure 1.3, which shows the plot of the log returns of the APX data. Obviously, these price changes are not normally distributed as can be seen from the histogram in the same figure. To check and remove outliers in data that are not normally distributed, the following simple statistic can be used. Following Benth *et al.* (2008), given the lower and upper quartiles, Q_1 and

Q_3, respectively, and the interquartile range IQR, defined as the difference between the upper and the lower quartile, an observation is called an outlier if it is smaller than $Q_1 - 3 \times \text{IQR}$, or larger than $Q_3 + 3 \times \text{IQR}$. We then substituted the detected outliers in the time series with the average of the two closest observations. Figure 1.4 shows the data of the APX after removing these spikes, together with its histogram.

Figure 1.3: Up: APX daily log returns before removing the spikes. Down: Histogram of the above log returns

Figure 1.4: Up: APX daily log returns after removing the spikes. Down: Histogram of the APX log returns after removing the spikes

Similar analysis to the EEX spot data are shown in Figure 1.5 and Figure 1.6, and the data corresponding to the UK-Gas-NBP market before and after removing the spikes is also shown in Figure 1.7 and Figure 1.8, respectively.

Figure 1.5: Up: EEX daily log returns before removing the spikes. Down: Histogram of the above log returns

Figure 1.6: Up: EEX daily log returns after removing the spikes. Down: Histogram of the EEX log returns after removing the spikes

Figure 1.7: Up: UK-Gas-NBP daily log returns before removing the spikes. Down: Histogram of the above log returns

Figure 1.8: Up: UK-Gas-NBP daily log returns after removing the spikes. Down: Histogram of the gas log returns after removing the spikes

Identifying the Trend

To model the trend, we consider a linear function describing the increase in the logarithm of the price level, given by $T(t) = a_0 + a_1 t$. The two

8

parameters a_0 and a_1 are found by following the least square approach given by

$$\min_{a_0,a_1} \int_0^{t_f} |\log S(t) - (a_0 + a_1 t)|^2 dt, \qquad (1.1)$$

where t_f is the final time and $S(t)$ is the spot data at time t (after removing the outliers).

Characterizing the Seasonality

To get the seasonality component, we remove the trend component and consider the following functional form applied to the logarithmic of the cleaned spot prices:

$$h_p(t) = \sum_{k=1}^{N} b_k \sin(2\pi f_k t) + c_k \cos\left(2\pi f_k t\right), \qquad (1.2)$$

where the frequencies f_k are identified by using the Fast Fourier Transform (FFT). After determining these frequencies f_k, we determine the coefficients b_k, c_k by using the least square method again.

Trend and seasonality analysis for APX data

From the APX data and after removing the spikes as shown in Figure 1.4, we first identify the coefficients a_0, a_1 of the linear trend using equation (1.1). We get $a_0 = 3.2215$, $a_1 = 0.0643$. In Figure 1.9, the APX data (without spikes) and the fitted linear trend are shown.

Now following Moler (2004), we subtract the linear trend from the data, see Figure 1.10 and then, we apply the FFT to the new data. The periodogram which represents the plot of the power of the signal versus frequency is shown is Figure 1.11.

Figure 1.9: APX daily log returns after removing the spikes with a linear trend

Figure 1.10: APX daily log returns after removing the spikes and a linear trend

Figure 1.11: Periodogram of Apx data

To identify the seasonality part, we look at the frequencies that correspond to the largest four powers, as can be seen from Figure 1.12.

Figure 1.12: Zoomed periodogram of APX data

The identified frequencies are in terms of cycles per day. The corre-

sponding periods are the reciprocal of these frequencies. Table 1.1 summarizes the result.

Table 1.1: The largest frequencies and the corresponding periods of the APX spot data

Frequency f (cycles/day)	Period (in day)
0.0014	≈ 720.00 (2 years)
0.0027	≈ 365.00 (1 year)
0.1424	≈ 7.00 (1 week)
0.1431	≈ 7.00 (1 week)

Now, using these frequencies, we continue to identify the coefficients b_k, c_k of equation (1.2) by using the least square method again to fit the data. We get

$$b_1 = 0.0828 \; b_2 = -0.1026 \; b_3 = -1.6841 \; b_4 = 1.7449$$
$$c_1 = 0.0163 \; c_2 = 0.0652 \; c_3 = -1.0601 \; c_4 = 0.8434$$

The original spot price together with the identified seasonality and trend functions are shown in Figure 1.13.

Figure 1.13: log of APX data together with the imposed seasonality plus trend

Trend and seasonality analysis for EEX data

Repeating the same procedure to the EEX data. The coefficients a_0, a_1 of the linear trend function are given by $a_0 = 2.8243$, $a_1 = 0.1566$. Figure 1.14, shows the linear trend function imposed on the data.

Figure 1.14: log of EEX data and a linear trend

Also the periodogram is shown in Figure 1.15.

Figure 1.15: Periodogram of the EEX data

Form Figure 1.15, we can identify the following four prominent frequencies. Again, we then zoom in to the windows that corresponds to the high powers. Figure 1.16 shows the plot and Table 1.2 shows the necessary frequencies and its periods.

Figure 1.16: Zoomed periodogram of the EEX data

Table 1.2: The largest frequencies and the corresponding periods of the EEX spot data

Frequency f(cycles/day)	Period (in day)
0.0005	$\approx$ 2000.00 (5.4 years)
0.1427	$\approx$ 7.00 (1 week)
0.2891	$\approx$ 2.60
0.4286	$\approx$ 2.30

Hence the seasonality of this data is concentrated around a week and two or three days. Figure 1.17 presents the original spot price data together with the imposed trend and seasonality functions.

Figure 1.17: log of Eex data and seasonality plus trend

Finally, the coefficients b_k, c_k characterizing the whole seasonality function, using least square method is given by

$$b_1 = 0.0484 \; b_2 = -0.1107 \; b_3 = -0.0147 \; b_4 = 0.0544$$
$$c_1 = 0.0918 \; c_2 = -0.1692 \; c_3 = 0.0039 \; c_4 = -0.0610.$$

Trend and seasoning analysis for UK-Gas-NBP data

From the UK-Gas-NBP data, we can identify the following parameters for the linear trend function as, $a_0 = 2.8122, a_1 = 1.0295$.

Figure 1.18, shows the trend function imposed on the data after removing the spikes.

Figure 1.18: log of UK-Gas-NBP data together with the trend function

The periodogram is also shown in Figure 1.19

Figure 1.19: Periodogram of UK-Gas-NBP data

Notice here that most of the power is related to the short frequency

region, i.e. The subplot in the upper left side of Figure 1.19. Notice that we have presented the x axis in terms of cycles/year. Zooming into this subplot, we get the enlarged figure appearing as Figure 1.20.

Figure 1.20: Zoomed periodogram of UK-Gas-NBP data

Table 1.1.4 shows the details of the four main frequencies and its corresponding periods

Frequency f (cycles/year)	Period (in year)
0.7668	≈ 1.30 (15.6 months)
1.5336	≈ 0.60 (7 months)
2.3004	≈ 0.40 (5 months)
6.1345	≈ 0.16 (2 months)

Moreover, The identified coefficients b_k, c_k are given by

$$b_1 = -0.1131 \; b_2 = 0.0679 \; b_3 = 0.0936 \; b_4 = 0.0779$$
$$c_1 = 0.0586 \; c_2 = 0.0600 \; c_3 = 0.0871 \; c_4 = -0.0454.$$

Finally we present the original UK NBP spot prices together with the imposed trend and seasonality functions, shown in Figure 1.21.

Figure 1.21: log of Gas data with the imposed linear trend and seasonality functions

1.1.5 Energy futures markets

Although contracts for future delivery of power are called futures or forwards, this denomination may be misleading because the exchange traded contracts at these markets are written on the (weighted) average of the (hourly) system price (spot price) over a specified delivery period, see Benth *et al.* (2008). During the delivery period the contract is settled in cash against the system price. Hence, financial electricity contracts are in fact swap contracts, exchanging a floating spot price against a fixed price. The specified reference price is typically the day ahead (DA) spot price, which has been discussed in Section 1.1.3.

More specifically, a futures contract is a contract that obliges the seller of the contract to deliver and the buyer to receive a given quantity of electricity or gas over a fixed period $[T_0, T]$ at a price K specified in advance. The payoff of these futures are based on the arithmetic average of the spot price $\frac{1}{n} \sum_{t=T_0}^{T} S(t)$ and not one fixed spot price $S(T)$ as in most financial and commodity futures markets. Here n is the number of days during the delivery period $[T_0, T]$. See Figure 1.22 for a typical energy futures

contract.

Figure 1.22: Payoff of energy futures contract

1.1.6   Energy modeling approaches

In the energy market, there are two types of models used to model the dynamics of the forward curve of energy commodities. One type of models is the spot price model which explicitly defines the spot dynamics, from which the forward dynamics can be constructed. A typical example is Schwartz and Smith (2000) (SS) model, which is popular in the commodity markets because it gives analytical solutions for the futures prices. See also Schwartz (1997). Because of this feature, particularly (SS) model has been adapted in the energy market with some modifications and approximations to the payoff structure without affecting its mathematical tractability. The other type of models uses the arbitrage-free framework of Heath *et al.* (1992) (HJM), which describes the forward curve dynamics directly via the use of volatility function(s). For examples, see Jamshidian (1991) and Clewlow and Strickland (1999*a,b*).

1.2   Thesis Motivation

The two approaches of modeling, either by using factors or volatilities, both have appealing features that they lead to tractable solutions for the

derivatives' prices. Both of these approaches have been used in the energy market following the trend in multi-factor modeling of the term structure of interest rates. However, these two markets are fundamentally different. The main difference is related to the payoff structure of the same type of contract, such as futures. This makes the methodology used in the interest rate market inapplicable in the energy market.

Another important issue is related to the parameter estimation of the models representing the dynamics of both the spot and the futures. As a result of dealing with unobservable factors, a popular estimation method that has been proposed in the literature is the maximum likelihood estimation (MLE) method under the assumption that observations are corrupted with additive Gaussian noise. In this framework, the state space representation is used together with the Kalman filtering techniques, and the parameter estimates are obtained through maximization of a likelihood functional.

To make this approach mathematically feasible, some *ad hoc* observation noise has to be added to the model in order to derive the Kalman filter as has been made by numerous authors, see Elliott and Hyndman (2007). The additional noise in the observation has been interpreted to take into account bid-ask spreads, price limits, non-simultaneity of the observations, or errors in the data. The argument is clearly forced, unconvincing and hard to verify. Even if one ignores this aspect, there is an additional complication with this approach. Since there is no feedback of the observation noise to the spot price, this leads to a model that is not anymore an arbitrage free model.

1.3 Goal

The goal of this thesis is two folds: Starting from the two factor model of Schwartz and Smith (2000), we formulate and implement a new arbitrage free model for the futures prices of energy which can be used in a mathematically sound way when estimating the parameter of the model, using the method of maximum likelihood (MLE). In this respect, we extend the idea proposed in Aihara and Bagchi (2010*a*) to the energy market. Following this approach, we assume that a slightly different model will lead to a slightly perturbed futures price from that derived using Schwartz and Smith (2000). We consider this perturbation to be generated by an error

term represented by a stochastic integral involving an infinite dimensional noise term, reflecting the fact that it should depend on all times of or to maturity.

In this setup, on the one hand, the added measurement noise is built in within the model. On the other hand, the modeling of the correlation structure between the futures (observation) is a natural component of our formulation. Hence the model parameters can be calibrated through the derived likelihood functional without adding any *ad hoc* observation noise. The second goal is to estimate the parameters of the model without any approximation to the nonlinear payoff of the futures (observation). For that, we use the particle filtering methodology. Moreover, and on the empirical side, we identify the parameters of the model using real data from the European energy market.

1.4 Structure of the Thesis

The thesis is organized in six chapters. The contents of the remaining chapters are briefly summarized as follows.

Chapter 2 This chapter contains the mathematical preliminaries needed throughout the thesis. It starts with a description of the filtering problem, then it goes through a review of the finite dimensional filter algorithm. Here, we present the discrete Kalman filter algorithm, together with its corresponding likelihood function. Also, we present the continuous version of the filtering equations, with its likelihood functional. These algorithms will be used in Chapter 3 and Chapter 4, respectively. Then we discuss the particle filtering algorithm and discuss its main properties. For that, we present a very basic review of Monte Carlo methods and importance sampling. Next, we explain the Sequential Importance Sampling method. We also discuss the problem of degeneracy associated with this method which is followed by the resampling step to mitigate this problem. At the end of this section, we present a generic particle filter algorithm. This section will be used in both Chapter 3 and Chapter 5. The end of this chapter contains a discussion of the infinite dimensional filter. For that we discuss the infinite dimensional Brownian motion and explain why it is a natural component in any term structure model. Then, we present the algorithm of the infinite dimensional Kalman filter. The use of this algorithm will be

presented in both Chapter 4 and Chapter 5.

Chapter 3 This chapter represents an illustration of the standard use of both the Kalman filter and the particle filter in the finance literature. By that we mean the assumption used in the literature that the measurements (futures) are corrupted by noise. In the first part of this chapter we use the Schwartz and Smith (2000) model as our basis. Using this model, we derive the futures price that depends on a delivery period by employing the geometric approximation instead of the arithmetic average. The main aim of this approximation is to keep the linearity of the resulting model so that the Kalman filter can be used. Moreover, we implement a sensitivity analysis for the likelihood function and show that the optimal parameters are hard to find when using the MLE method. In the second part of this chapter, we avoid this approximation and we use the particle filter algorithm for the estimation of the parameters. This chapter is based on Imreizeeq (2010).

Chapter 4 The main focus of this chapter is the development of a new model for the energy market. Using reverse engineering concept, we start by assuming that the dynamics of the futures is perturbed by extra term that takes into account the uncertainties in both the time, and time of maturity, of the term structure of the futures. After deriving the explicit relation between the observed energy futures prices and the factor processes, we employ the geometric approximation for the payoff of the futures. However, in contrast to Chapter 3, we employ the infinite dimensional Kalman filter together with the MLE method to estimate the parameters. As the observation noise covariance in this case is unknown, we derive a quasi likelihood functional. Moreover, we discuss the invertibility of the covariance of the real observation data. At the end of this chapter, we use simulation and real data of the spot and futures on the UK-Gas-NBP, and establish the feasibility of the proposed filter. This chapter is based on Aihara *et al.* (2009*b*).

Chapter 5 In this chapter, we extend the results of the previous chapter by employing the particle filtering methodology as our method for the parameter estimation of the new model of the futures prices. In other words, we avoid the use of the geometric approximation to the

payoff of the futures. Here, we propose to use a variant of the particle filter, which is based on the convolution kernel approximation techniques termed convolution particle filter. In order for this chapter to be self-contained, we briefly repeat our discussion of Chapter 4 and present the forward model together with the mechanism of the observation equation. Also, we discuss the discrete approximation for the system and observation. Moreover, we show how to generate the two dimensional noise term $\frac{\partial w(t,x)}{\partial t}$ which appears in the dynamics of our state equation. In this chapter, we employ the Bayesian framework where the augmented state variable that contains the state and the unknown parameters is processed by a filtering procedure. Finally, we run a simulation study to test the feasibility of the proposed filter. This chapter is based on Aihara *et al.* (2009*a*) and Aihara *et al.* (2011).

Chapter 6 This chapter presents conclusions and recommendations on the possible directions for future research.

Chapter 2

Preliminaries

2.1 Introduction

We begin with a brief review of the filtering problem. We describe both the discrete and continuous finite dimensional Kalman filters. Then we describe the maximum likelihood functions which can be used for the estimation of parameters. Then we go through the particle filtering approach for state estimation and its variant algorithms. We conclude this chapter with stating the infinite dimensional Kalman filter algorithm.

2.2 The Filtering Problem

Following Bagchi (1993), consider a stochastic dynamical system, which can be represented by either continuous/discrete stochastic differential/ difference equations, respectively. These equations are termed the system equation. Also we have a sequence of observations of some functions of the states. Typically, the state equation represents unobserved states, and the measurement of some functions of these states are typically corrupted with noise (measurement error). The measurement and state equations represent together what is called the state-space form of the model. To explain the filtering problem, suppose that the n-dimensional state vector X_k of a stochastic dynamical system is not directly observed and is only available through an observation mechanism which generates measurements

$Y_1, Y_2, \cdots, Y_l$. Suppose that we want to obtain an estimator of X_k on the basis of the available observations $Y_1, Y_2, \cdots, Y_l$ in some optimal way. One intuitively appealing approach is to consider the minimum variance estimator of X_k based on the observations $Y_1, Y_2, \cdots, Y_l$. To make this idea clear let us stack all the m-dimensional observations into one random vector $\mathbf{Y}(l) = [Y_0 \ Y_1 \cdots Y_l]'$. The minimum variance estimator of X_k, denoted $\hat{X}_{k|l}$, is a function of $\mathbf{Y}_l$ such that for any other function F of $\mathbf{Y}_l$,

$$\mathbf{E}\left[\left\|X_k - \hat{X}_{k|l}\right\|^2\right] \leq \mathbf{E}\left[\|X_k - F(\mathbf{Y}_l)\|^2\right]$$

The solution of this problem is well-known and is given by

$$\hat{X}_{k|l} = \mathbf{E}[X_k|\mathbf{Y}_l]$$

The estimation problem is called filtering if $k = l$, prediction if $k > l$ and smoothing if $k < l$.

2.3 Finite Dimensional Filter

2.3.1 The discrete Kalman filter

Following Bagchi (1993), Consider the following discrete-time linear stochastic dynamical system

$$
\begin{array}{rcl}
X_{k+1} & = & A_k X_k + B_k U_k + F_k W_k \hspace{2em} (2.1) \\
Y_k & = & C_k X_k + D_k + V_k, \ k \geq 0 \hspace{1.5em} (2.2)
\end{array}
$$

where X_k is an n-dimensional random vector denoting the state at the time instant k, the r-dimensional random vector W_k is the system disturbance, U_k is a p-dimensional input (control) vector which is either a deterministic sequence, or is such that U_k, for each fixed k, is a (Borel measurable) function of $\mathbf{Y}_k$. B_k is a matrix of order $n \times p$. The m-dimensional random vector Y_k denotes the observation and V_k is the observation error. We assume that X_0, $\{W_k\}$ and $\{V_k\}$ are jointly Gaussian, X_0 has mean $\bar{x}_0$, variance $\bar{P}_0$ and is independent of $\{W_k\}$ and $\{V_k\}$, $k \geq 0$. Assume that $EW_k = EV_k = 0$, $k \geq 0$ and

$$E\left[\begin{pmatrix} W_k \\ V_k \end{pmatrix}\begin{pmatrix} W_l' \ V_l' \end{pmatrix}\right] = \begin{pmatrix} \Sigma_k^s & \mathbf{0} \\ \mathbf{0} & \Sigma_k^o \end{pmatrix}\delta_{kl} \hspace{2em} (2.3)$$

with $\Sigma_k^o > 0$ for all k.

The Kalman filtering problem is that of calculating $E\left[X_k|\mathbf{Y}_k\right]$ or, equivalently, $\hat{X}_k \triangleq \hat{X}_{k|k}$, where $\mathbf{Y}_k = \{Y_0, Y_1, \cdots, Y_k\}$. Kalman filter recursive equations are:

1. Time update $\hat{X}_{k+1|k}$ and $P_{k+1|k}$

$$\hat{X}_{k+1|k} = A_k\hat{X}_k + B_kU_k \tag{2.4}$$

$$P_{k+1|k} = A_kP_kA_k' + F_k\Sigma_k^sF_k' \tag{2.5}$$

where $P_{k+1|k} \triangleq \mathbf{E}\left[\left(X_{k+1} - \hat{X}_{k+1|k}\right)\left(X_{k+1} - \hat{X}_{k+1|k}\right)'\right]$

2. Measurement update $\hat{X}_{k+1}$ and P_{k+1}

$$\hat{X}_{k+1} = \hat{X}_{k+1|k} + K_{k+1}\nu_{k+1} \tag{2.6}$$

$$K_{k+1} \triangleq P_{k+1|k}C_{k+1}'\left[C_{k+1}P_{k+1|k}C_{k+1}' + \Sigma_{k+1}^o\right] \tag{2.7}$$

$$\nu_{k+1} = Y_{k+1} - C_{k+1}\hat{X}_{k+1|k} - D_k \tag{2.8}$$

$$\begin{aligned} P_{k+1} = \ & P_{k+1|k} - P_{k+1|k}C_{k+1}' \\ & \times \left[C_{k+1}P_{k+1|k}C_{k+1}' + \Sigma_{k+1}^o\right]^{-1} \\ & \times C_{k+1}P_{k+1|k} \end{aligned} \tag{2.9}$$

where $P_{k+1} \triangleq \mathbf{E}\left[(X_{k+1} - \hat{X}_{k+1})(X_{k+1} - \hat{X}_{k+1})'\right]$

As for the initial conditions, we can take: $\hat{X}_{0|-1} = \bar{x}_0$, and $P(0|-1) = \bar{P}_0$.

Notice that the calculation of equations (2.5, 2.7, 2.9) do not depend on the measurements Y_k, but depend only on A_k, C_k, Σ_k^s and Σ_k^o. That means that the Kalman gain K_k can be calculated offline before the system operates and saved in memory. Only equation (2.6) need to be implemented in real time. This has a great advantage in practice.

2.3.2 Likelihood function

Let us now assume that the system is time invariant; that is, A_k, B_k, F_k, C_k, D_k are constants (independent of k) and some or all components of these matrices are unknown. Let θ denote the vector of all unknown parameters of the system (2.1, 2.2). One possible way of estimating θ is to use the method of maximum likelihood. For that we have to calculate the likelihood function which can be obtained from the joint probability density of $\mathbf{Y}_N = \{Y_0, \cdots, Y_N\}$. This density using Bayes' rule is given by

$$p_{\mathbf{Y}_N}(y_N; \theta) = p_{Y_1}(y_1) \prod_{k=2}^{N} p_{Y_k|\mathbf{Y}_{k-1}, \theta}(y_k|\mathbf{y}_{k-1}, \theta)$$

where $p_{Y_k|\mathbf{Y}_{k-1}, \theta}(y_k|\mathbf{y}_{k-1}, \theta)$ denotes the conditional probability density of Y_k given $\mathbf{Y}_{k-1}$ and θ. From results of Kalman filtering, we know that the innovation process ν_k given by (2.8) is a realization of ν_k for a true parameter θ, is a Gaussian white noise with zero mean and variance given by

$$\mathbf{Q}_k = CP_{k|k-1}C' + \Sigma_o \qquad (2.10)$$

Therefore,

$$p_{\mathbf{Y}_N}(y_N; \theta) = p_{\mathbf{Y}_1}(y_1) \prod_{k=2}^{N} [(2\pi)^n |\det \mathbf{Q}_k|]^{\frac{1}{2}}$$
$$\exp\left(-\frac{1}{2}\nu_k' \mathbf{Q}_k^{-1} \nu_k\right) \qquad (2.11)$$

The likelihood function is obtained from (2.11) by substituting $\mathbf{Y}_N$ in place of the actual observation $\mathbf{y}_N$. Hence, the likelihood function is given by

$$L(\mathbf{Y}_N; \theta) = p_{\mathbf{Y}_N}(\mathbf{Y}_N; \theta) \qquad (2.12)$$
$$= p_{Y_1}(Y_1) \prod_{k=2}^{N} [(2\pi)^n |\det \mathbf{Q}_k|]^{\frac{1}{2}} \exp\left(-\frac{1}{2}\nu_k' \mathbf{Q}_k^{-1} \nu_k\right)$$

where

$$\nu_k = Y_k - C\hat{X}_{k|k-1} - D.$$

A maximum likelihood estimator of θ, denoted $\hat{\theta}_N$, is that value of θ for which $L(\mathbf{Y}_N; \theta)$ is a maximum, or equivalently, $\log L(\mathbf{Y}_N; \theta)$ is a maximum.

2.3.3 The continuous Kalman filter

Consider the following continuous -time linear stochastic dynamical system

$$X(t) = X(t_0) + \int_{t_0}^{t} A(\tau) X(\tau) d\tau + \int_{t_0}^{t} B(\tau) d\tau + \int_{t_0}^{t} F(\tau) dW^s(\tau)$$

(2.13)

$$Y(t) = \int_{t_0}^{t} C(\tau) X(\tau) ds + \int_{t_0}^{t} D(\tau) d\tau + W^o(t), \ 0 \le t \le T \quad (2.14)$$

where the state $X(t)$, as in the discrete case, takes values in $\mathbb{R}^n$, the observation $Y(t)$ takes values in $\mathbb{R}^m$, $\{W^s(t), t \ge 0\}$ and $\{W^o(t), t \ge 0\}$ are r and m dimensional Brownian motions, and $A(t), F(t), C(t)$, are matrices of order $n \times n$, $n \times r$, $m \times n$, respectively. $B(t)$ and $D(t)$ are deterministic functions of t of appropriate dimensions. $X(t_0)$ is a Gaussian random vector that follows $N(\bar{x}(t_0), \bar{P}(t_0))$, with mean $\bar{x}(t_0)$ and covariance matrix $\bar{P}(0)$.

Assume that

$$\mathbf{E}\left[\begin{pmatrix} W^s(t) \\ W^o(t) \end{pmatrix} \left(W^{s'}(\tau) \ W^{o'}(\tau)\right)\right]$$

$$= \begin{pmatrix} \int_0^{\min(t,\tau)} \Sigma_s(\sigma) d\sigma & \mathbf{0} \\ \mathbf{0} & \int_0^{\min(t,\tau)} \Sigma_o(\sigma) d\sigma \end{pmatrix} \quad (2.15)$$

and $\Sigma_o(t) > 0$ for all t.

In this case, we may also try to determine the minimum variance estimator. Let $\mathcal{F}_t^Y$ be the smallest σ- algebra generated by $\{Y(\tau);$

$t_0 \leq \tau \leq t$}. The minimum variance estimator of $X(t)$, based on the observation $\mathcal{F}_t^Y$, denoted by $\hat{X}(t)$, is an $\mathcal{F}_t^Y$- measurable function, such that for any $\mathcal{F}_t^Y$- measurable function F_t,

$$\mathbf{E}\left[\left\|X(t) - \hat{X}(t)\right\|^2\right] \leq \mathbf{E}\left[\|X(t) - F_t\|^2\right]$$

The solution of this problem is well-known and is given by

$$\hat{X}(t) = \mathbf{E}\left[X_t | \mathcal{F}_t^Y\right]$$

The recursive equation for the filter $\hat{X}(t) = \mathbf{E}\left[X_t | \mathcal{F}_t^Y; t_0 \leq \tau \leq t\right]$ is given by

$$
\begin{aligned}
d\hat{X}(t) &= A(t)\hat{X}(t)dt + B(t)dt + K(t)\Sigma_o(t)^{-1}d\nu(t) \\
d\nu(t) &= dY(t) - C(t)\hat{X}(t)dt - D(t)dt \\
K(t) &= P(t)C(t)' \\
\hat{X}(t_0) &= \bar{x}(t_0)
\end{aligned}
$$

$$
\begin{aligned}
\dot{P}(t) &= A(t)P(t) + P(t)A(t)' + F(t)\Sigma_s(t)F(t)' \\
&\quad - K(t)\Sigma_o(t)^{-1}K(t)' \\
P(t_0) &= \bar{P}(t_0)
\end{aligned}
$$

2.3.4 The likelihood functional

We now consider the parameter estimation problem, where the system is time-invariant; that is, $A(t), B(t), F(t), C(t), D(t)$ are all constants. Suppose the set of observations $\{Y(s), 0 \leq s \leq T\}$ are available for the purpose of parameter estimation. In addition to the assumption that $Y(t)$ satisfies the measurement equation (2.14) and the non observed factors $X(t)$ satisfies the state equation (2.13), we also assume that $\Sigma_o > 0$ and is known completely. Without loss of generality we can assume that $\Sigma_o = I_m$. (Simply redefine Y by $\left(\sqrt{\Sigma_o}\right)^{-1}Y$). In this case, a likelihood function can be defined if we find a fixed measure on $\mathbf{C} = C([0,T]; R^m)$, (the space of continuous functions from $[0,T]$ into R^m) such that the measure induced on $\mathbf{C}$ by the observation process $Y(t)$, $0 \leq t \leq T$, denoted by p_Y is absolutely continuous with

respect to that fixed measure. If we use p_{W^o}, the measure induced by the process $W^o(t)$, $0 \leq t \leq T$, on **C** as the fixed measure, we can define the likelihood functional as the corresponding Radon-Nikodym derivative evaluated at the sample trajectory of the observation. In this case, the likelihood functional is given by, see Balakrishnan (1973) and Bagchi (1975) for more details,

$$
L\left(Y\left(\cdot\right);\theta\right) = \exp\left\{ -\frac{1}{2} \int_0^T \left\| C\hat{X}\left(t\right) + D \right\|^2 dt \right.
$$
$$
\left. - \int_0^T \left\langle C\hat{X}\left(t\right) + D, dY\left(t\right) \right\rangle \right\} \qquad (2.16)
$$

where θ denotes the vector containing all unknown parameters that describe the dynamics of the system. The bracket term denotes an inner product $\langle a, b \rangle = a'b$ for $a, b \in R^m$ and $\|a\|^2 = \langle a, a \rangle$. where

$$
\hat{X}\left(t\right) = \mathbf{E}\left[X\left(t\right) | \mathcal{F}_t^Y\right]
$$

The estimate of the unknown vector θ can be found by maximizing the likelihood functional (2.16); or equivalently, its logarithm, that is,

$$
\hat{\theta} = \arg\max_{\theta} L\left(\theta\right)
$$

2.4 Particle Filtering

Kalman filter is based on the assumption of linear model and Gaussian disturbance, so that at every time step the states and observations are Gaussian. In many real world applications, these assumptions cannot be expected to hold. Other, sub-optimal filters have therefore been developed to deal with non-linear functions and non-Gaussian disturbances. Particle filter is one such sub-optimal filter which is widely used these days.

In the Bayesian framework, all relevant information about $\mathbf{X}_k \equiv \{X_0, \ldots X_k\}$ given observation $\mathbf{Y}_k \equiv \{Y_1, \ldots Y_k\}$ up to and including time k can be obtained from the posterior probability density function

$p_{\mathbf{X}_k|\mathbf{Y}_k}(\mathbf{x}_k|\mathbf{y}_k) \equiv p(\mathbf{x}_k|\mathbf{y}_k)$. In real applications we are mainly interested in estimating recursively in time the filtering density given by $p_{X_k|\mathbf{Y}_k}(x_k|\mathbf{y}_k) \equiv p(x_k|\mathbf{y}_k)$. From this density one can get any filtered point estimates such as the posterior mode or mean of the state. The recursive filter consists of two steps of prediction and updating. Following Ristic *et al.* (2004) and Bølviken and Storvik (2001), consider the following general discrete state-space model given by

$$X_k = f(X_{k-1}, W_k) \tag{2.17}$$

$$Y_k = h(X_k, V_k) \tag{2.18}$$

where X_k is the unobservable system equation, taking values in $\mathbb{R}^n$ with initial (prior) density $p(x_0) \equiv p(x_0|\mathbf{y}_0)$. Y_k is the measurements process, taking value in $\mathbb{R}^m$. The process noises $W_k, k = 1, 2, \cdots$ are assumed to be independent, so are the measurement noises $V_k, k = 1, 2, \cdots$ Furthermore, W_k is assumed to be independent of V_k. $f(x, w)$ and $h(x, v)$ are functions of (x, w) and (x, v), respectively, where both can be nonlinear. In this model, we assume that the probability density functions for W_k and V_k are known.

The above model can also be characterized in terms of its probabilistic description via the state transition density $p(x_k|x_{k-1})$ and the observation density $p(y_k|x_k)$. This follows from the fact that X_k is a Markov process, i.e. the conditional density of X_k given the past state $\mathbf{X}_{k-1}$, depends only on X_{k-1}, and, the conditional density of Y_k given the state $\mathbf{X}_k$ and the past observations $\mathbf{Y}_{k-1}$, depends only on X_k. Then, in principle, the filtered density $p(x_k|\mathbf{y}_k)$ may be obtained recursively in two stages: prediction and update.

Suppose that the filtered density $p(x_{k-1}|\mathbf{y}_{k-1})$ at time $k - 1$, termed the prior, is available. The prediction stage involves using the system equation (2.17) to obtain the prior pdf of the state at time k via the Chapman-Kolmogorov equation given by

$$p(x_k|\mathbf{y}_{k-1}) = \int p(x_k|x_{k-1}) p(x_{k-1}|\mathbf{y}_{k-1}) dx_{k-1} \tag{2.19}$$

At time step k, a measurement y_k becomes available, and it can be used to update the prior via Bayes' rule, that is

$$p(x_k|\mathbf{y}_k) = \frac{p(y_k|x_k) p(x_k|\mathbf{y}_{k-1})}{p(y_k|\mathbf{y}_{k-1})} \tag{2.20}$$

where the normalized constant

$$p(y_k|\mathbf{y}_{k-1}) = \int p(y_k|x_k)p(x_k|\mathbf{y}_{k-1})dx_k \qquad (2.21)$$

depends on the likelihood function $p(y_k|x_k)$. In the update stage (2.20), the measurement y_k is used to modify the prior density to get the required posterior density $p(x_k|\mathbf{y}_k)$ of the current state at time k. Thus, starting from the initial density $p(x_0)$ one can, at least in principle, recursively arrive at the desired density $p(x_k|\mathbf{y}_k)$.

However, it is not possible, in general, to derive optimal closed-form estimations of the state and we must adopt numerical strategies. Particle filter is precisely used for that purpose, where a sequential Monte Carlo method is used to represent the required posterior density function. This occurs through discrete approximations to the exact posterior distributions. Let $\hat{p}(x_t|\mathbf{y}_t)$ be some discrete analogue to the exact density $p(x_t|\mathbf{y}_t)$. The points $x_t^{(i)}$ on which $\hat{p}(x_t)$ assigns positive probabilities are known as the particles. Their numbers N_k may vary. Suppose a reasonable approximation $\hat{p}(x_{k-1}|\mathbf{y}_{k-1})$ is available at time $k-1$. When inserted for the exact density $p(x_{k-1}|\mathbf{y}_{k-1})$ on the right in (2.19), we obtain

$$\hat{p}(x_k|\mathbf{y}_{k-1}) = \sum_{i=1}^{N_{k-1}} p(x_k|x_{k-1}^{(i)})\hat{p}(x_{k-1}^{(i)}|\mathbf{y}_{k-1}) \qquad (2.22)$$

The main point of the design is to ensure a good approximation to the exact predictive density $p(x_k|\mathbf{y}_{k-1})$. When (2.22) replaces its exact counterpart in (2.20), and we get as an approximate update density

$$\widetilde{p}(x_k|\mathbf{y}_k) = \frac{p(y_k|x_k)\hat{p}(x_k|\mathbf{y}_{k-1})}{\hat{p}(y_k|\mathbf{y}_{k-1})} \qquad (2.23)$$

where the normalized constant $\hat{p}(y_k|\mathbf{y}_{k-1})$ is a discrete approximation of (2.21). To complete the recursion, (2.23) must be replaced by a particle approximation $\hat{p}(x_t|\mathbf{y}_k)$. The form of this approximation can be represented as:

$$p(x_k|\mathbf{y}_k) \approx \sum_{i=1}^{N_s} w_k^i \delta(x - x_k^i) \qquad (2.24)$$

where it can be shown that as $N_s \to \infty$, the approximation (2.24), approaches the true posterior density $p(x_k|\mathbf{y}_k)$.

Particle filtering has different forms and methods, one important method which represents the basis for most sequential Monte Carlo filters developed over the past decades, is the Sequential Importance Sampling (SIS) algorithm. Using this algorithm, it can be shown that the weight corresponding to each particle i denoted by w_k^i satisfies the following recursive relation, see Arulampalam *et al.* (2002) for the details,

$$w_k^i = w_{k-1}^i \frac{p(\mathbf{y}_k|x_k^i)p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{k-1}^i, \mathbf{y}_k)} \tag{2.25}$$

where $q(\cdot)$ represents a proposal density (importance function). Ideally, the proposal density should be the posterior density itself $p(\mathbf{x}_k|\mathbf{y}_k)$, but this quantity is unknown (it is what we are looking for). A pseudocode description of the SIS algorithm is given by algorithm 1.

Algorithm 1. *SIS PARTICLE FILTER*

$$\left[\{x_k^i, w_k^i\}_{i=1}^{N_s}\right] = SIS\left[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^{N_s}, y_k\right]$$

- *FOR $i = 1 : N_s$*

    - *Draw $x_k^i \sim q\left(x_k|x_{k-1}^i, y_k\right)$*
    - *Assign the particle a weight, w_k^i, according to (2.25)*

- *END FOR*

Although the SIS particle filter is easy to implement, there is a common problem known as the degeneracy phenomenon. It means that after a few iterations, most weights will be carried by few particles and the algorithm fails to represent the posterior density. This problem can be partially tackled using two methods: The first is a good choice of importance density, and the second is the use of a new step within the SIS algorithm called resampling. For the first method, it has been shown by Doucet *et al.* (2000), that the optimal importance density function which minimizes the variance of the true

weights, w_k^{*i}, conditioned upon x_{k-1}^i and y_k is given by

$$
\begin{aligned}
q\left(x_k|x_{k-1}^i, y_k\right)_{opt} &= p\left(x_k|x_{k-1}^i, y_k\right) \\
&= \frac{p\left(y_k|x_k, x_{k-1}^i\right) p\left(x_k|x_{k-1}^i\right)}{p\left(y_k|x_{k-1}^i\right)}
\end{aligned}
\tag{2.26}
$$

Substituting this optimal importance density in (2.25), we get

$$
\begin{aligned}
w_k^i &\propto w_k^i p\left(y_k|x_{k-1}^i\right) \\
&= w_{k-1}^i \int p\left(y_k|x_k'\right) p\left(x_k'|x_{k-1}^i\right) dx_k'
\end{aligned}
\tag{2.27}
$$

However, this optimal importance density suffers from two major drawbacks. The first is to be able to sample from $p\left(x_k|x_{k-1}^i, y_k\right)$ and the second is the evaluation of the integral in (2.27). We now consider the second method of using the resampling step by which the degeneracy problem can be reduced. The basic idea of the resampling method is to eliminate particles which have small weights and to concentrate on particles with large weights. It involves generating a new set $\{x_k^i\}_{i=1}^{N_s}$ by resampling with replacement N_s times from an approximate discrete representation of $p\left(x_k|\mathbf{y}_k\right)$ given by

$$
p\left(x_k|\mathbf{y}_k\right) \approx \sum_{i=1}^{N_s} w_k^i \delta\left(x_k - x_k^i\right)
\tag{2.28}
$$

so that $\Pr\left(x_k^{i*} = x_k^j\right) = w_k^j$. The resulting sample is an i.i.d sample from the discrete density (2.28), and so the weights are reset to $w_k^i = \frac{1}{N_s}$. However, although the resampling step reduces the effects of the degeneracy problem, it introduces other practical problem known as sample impoverishment, in which the particles which have high weights are statistically selected many times, leading to a loss of diversity among the particles. Methods to counter these problems have led to many variants of particle filter algorithms, such as sample importance resampling (SIR), auxiliary sampling importance resampling (ASIR), regularized particle filter (RPF). However, it can be introduced within a generic framework of the sequential importance sampling (SIS). Following Arulampalam *et al.* (2002), a generic particle filter can be described by algorithm 2:

Algorithm 2. *GENERIC PARTICLE FILTER*

$$\left[\{x_k^i, w_k^i\}_{i=1}^{N_s}\right] = PF\left[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^{N_s}, y_k\right]$$

- *FOR $i = 1 : N_s$*

  - *Draw $x_k^i \sim q\left(x_k | x_{k-1}^i, y_k\right)$*
  - *Assign the particle a weight, w_k^i, according to (2.25)*

- *END FOR*

- *Calculate total weights: $t = SUM\left[\{w_k^i\}_{i=1}^{N_s}\right]$*

- *For $i = 1 : N_s$:*

  - *Normalize: $w_k^i = \frac{w_k^i}{t}$*

- *END FOR*

- *Calculate what is called the effective number of particles as $N_{eff} = \frac{1}{\sum\limits_{i=1}^{N}\left(\hat{w}_k^{(i)}\right)^2}$*

- *IF $N_{eff} < N_{thr}$, where N_{thr} is a given threshold, then do resampling*

  - *resample from $\left(x_k^{(i)}\right)_{i=1}^{N}$ with probabilities $\left(\hat{w}_k^{(i)}\right)_{i=1}^{N}$ to get a new set of particles*
  - *put $\left(w_k^{(i)}\right)_{i=1}^{N} = \frac{1}{N}$*

- *END IF*

2.5   Infinite Dimensional Filter

2.5.1   Infinite dimensional Brownian motion

In term structure modeling, as in modeling the futures prices of electricity, the state is a function of two variables; t (the time) and x (the

time to maturity). For stochastic modeling, it is then natural to introduce two-parameter Brownian motion $w(t, x)$. One way of defining this is to consider $w(t, x)$ to be a stochastic process in t with values in the space of functions of x. If these functions are in a (separable) Hilbert space, we may think of $w(t, x)$ as a Hilbert space valued stochastic process in t.

For simplicity we set the Hilbert space $H = L^2(0, \hat{T})$. Hence we can choose the orthogonal sequence $\{e_k\}$ in $L^2(0, \hat{T})$ as

$$e_k(x) = \sin(\pi k \frac{x}{\hat{T}}),$$

i.e., for any function $g \in L^2(0, \hat{T})$ we have

$$g(x) = \sum_{k=1}^{\infty} g_k e_k(x),$$

where

$$g_k = \int_0^{\hat{T}} g(x) e_k(x) dx = (g, e_k).$$

Furthermore $g \in L^2(0, \hat{T})$ implies

$$\sum_{k=1}^{\infty} g_k^2 < \infty. \tag{2.29}$$

The two parameter Brownian motion $w(t, x)$ is formulated to follow the above procedure. We assume that for each t, $w(t, x)$ is in $L^2(0, \hat{T})$, i.e.,

$$w(t, x) = \sum_{k=1}^{\infty} \beta_k(t) e_k(x),$$

where

$$\beta_k(t) = (w(t, x), e_k(x)) = \int_0^{\hat{T}} w(t, x) e_k(x) dx.$$

Now we set $\{\beta_k(t)\}$ are set as the mutually independent Brownian motion processes in R^1:

$$\begin{aligned} E\{\beta_k(t)\} &= 0, \quad E\{\beta_k(t)\beta_\ell(t)\} = 0 \text{ for } k \neq \ell \\ E\{\beta_k^2(t)\} &= \lambda_k t. \end{aligned} \tag{2.30}$$

To support the square integrability for $w(t, x)$ with respect to x we need to set

$$\sum_{k=1}^{\infty} \lambda_k < \infty.$$

This implies that

$$
\begin{aligned}
E\{\int_0^{\hat{T}} w^2(t, x)dx\} &= \int_0^{\hat{T}} E\{(\sum_{k=1}^{\infty} \beta_k(t)e_k(x))^2\}dx \\
&= \int_0^{\hat{T}} \sum_{k=1}^{\infty} E\{\beta_k^2(t)\}e_k^2(x)dx \\
&= \sum_{k=1}^{\infty} \lambda_k \int_0^{\hat{T}} e_k^2(x)dxt \\
&= t\sum_{k=1}^{\infty} \lambda_k < \infty.
\end{aligned}
$$

It is also easy to see that $\forall \phi_1, \phi_2 \in L^2(0, \hat{T})$

$$
\begin{aligned}
&E\{\int_0^{\hat{T}} w(t, x)\phi_1(x) \int_0^{\hat{T}} w(t, y)\phi_2(y)dxdy\} \\
&= \int_0^{\hat{T}} \int_0^{\hat{T}} \phi_1(x)E\{w(t, x)w(t, y)\}\phi_2(y)dxdy \\
&= \int_0^{\hat{T}} \int_0^{\hat{T}} \phi_1(x) \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} E\{\beta_k(t)\beta_j(t)\}e_k(x)e_j(y)\phi_2(y)dxdy \\
&\quad \text{from (2.30)} \\
&= \int_0^{\hat{T}} \int_0^{\hat{T}} \phi_1(x) \sum_{k=1}^{\infty} \lambda_k e_k(x)e_k(y)\phi_2(y)dxdyt \\
&= (\phi_1, Q\phi_2)t,
\end{aligned}
$$

where

$$Q = \int_0^{\hat{T}} \sum_{k=1}^{\infty} \lambda_k e_k(x)e_k(y)(\cdot)dy.$$

Hence the operator Q is called the covariance operator of the H-valued Brownian motion $w(t, x)$ and it turns out to be a trace-class operator:

$$Tr\{Q\} = \int_0^{\hat{T}} \sum_{k=1}^{\infty} \lambda_k e_k(x) e_k(x) dx = \sum_{k=1}^{\infty} \lambda_k < \infty.$$

Intuitively $\sum_{k=1}^{\infty} \lambda_k e_k(x) e_k(y)$ denotes the covariance kernel $q(x, y)$ which satisfies

$$E\{w(t, x) w(t, y)\} = q(x, y) t.$$

If we set

$$\lambda_k = 1, \text{ for all } k,$$

the kernel $q(x, y) = \delta(x - y)$. In this case, $w(t, x)$ has no correlation for the spatial variable, i.e. $w(t, x)$ is a white noise in x. This situation, of course, is ruled out if we assume that the $\sum_{k=1}^{\infty} \lambda_k < \infty$.

2.5.2 Infinite dimensional Kalman filter

Following Bagchi and Borkar (1984), we consider an integral abstract signal process

$$X(t) = \int_0^t S_{t-s} D dw(s), t \geq 0 \tag{2.31}$$

and the observation equation

$$Y(t) = \int_0^t C X(s) ds + F W^o(t) \tag{2.32}$$

where S_t, $t \geq 0$, is a strongly continuous semigroup with generator A on a separable Hilbert space H; $w(t)$ is a Brownian motion on a separable Hilbert space K and has incremental covariance W, see Curtain and Pritchard (1978) for details. $D \in \mathcal{L}(K, H)$, $W^o(t)$ is a vector valued Brownian motion on $\mathbb{R}^q$ and has incremental covariance matrix V; V, V^{-1}, F, $F^{-1} \in \mathcal{L}(\mathbb{R}^q)$; $C \in \mathcal{L}(H, \mathbb{R}^q)$ and w, W^o are mutually independent, $\mathcal{L}(A, B)$ stands for the class of all bounded linear operators from A into B).

We denote by $\hat{X}(t)$ the filtered estimate of $X(t)$ based on $Y(s)$, $0 \leq s \leq t$. Then

$$\hat{X}(t) = \int_0^t S_{t-s} P(t) C^* d\nu(s) \tag{2.33}$$

where $\nu(t)$, the so-called *innovations process* is defined by

$$\nu(t) = Y(t) - \int_0^t C\hat{x}(s)ds \qquad (2.34)$$

and $P(t)$ is the unique solution of the functional Riccati differential equation

$$\frac{d}{dt}\langle P(t)h, k\rangle = \langle P(t)h, A^*k\rangle + \langle A^*h, P(t)k\rangle$$
$$+\langle DWD^*h, k\rangle - \langle P(t)C^*CP(t)h, k\rangle; \qquad (2.35)$$
$$h, k \in D(A^*)$$

where $\langle \cdot, \cdot \rangle$ denotes inner product in H and $D(A^*)$ denotes the domain of the unbounded operator A^*.

Chapter 3

Parameter Estimation of Two Factor Model of Schwartz-Smith Using Particle Filter

3.1   Introduction

In the finance literature, two main approaches for modeling commodity price dynamics stand out. The first approach starts with the explicit modeling of the spot price dynamics, from which forward price dynamics can be constructed; see for example Schwartz (1997) and Schwartz and Smith (2000). The other approach uses the arbitrage-free framework of Heath *et al.* (1992) (HJM), which describes the forward price dynamics directly using explicit volatility functions. Examples of application of such framework to commodity price modeling are Jamshidian (1991) and Clewlow and Strickland (1999*a,b*).

However, in the empirical implementation of either of the above approaches, one of the main difficulties is the estimation of the parameters of the model. The estimation problem becomes even more difficult because some factors of these models are not directly observable. For instance, the spot price is proxied with the forward/futures price with the closest time to maturity, which can go as far as one

month in the case of coal[1]. In the case of convenience yield models like in Gibson and Schwartz (1990), the convenience yield cannot be observed. Another difficulty stems from the fact that the prices themselves contain some observation noise attributed to the lack of liquidity or high bid-ask spread. For these issues, many recent papers like Schwartz and Smith (2000), Elliott and Hyndman (2007), Manoliu and Tompaidis (2002) and Geman and Roncoroni (2006) use a filtering approach as a better alternative to estimate parameters.

In this chapter, we illustrate the use of filtering in the energy market, both by Kalman filtering and particle filtering techniques. We consider the problem of estimating the parameter of the two factor model of Schwartz and Smith (2000), which is a popular model in the commodity market. However, to extend the model to be suited to the energy market, we need to deal with the delivery period of the futures contract.

In this respect, our first contribution to the literature is to get an expression for the futures price that takes into account the delivery period of the contract. For that we use the the geometric approximation in continuous time to the payoff structure of the contract. As a result, we can be able to express the model in state space form, and once it is cast in a state space form, the Kalman filter can be applied to estimate the unobservable state variables and the parameters of the model. However, since we are dealing with many parameters, a sensitivity analysis of the use of MLE method shows that it is hard to find reliable estimates of the parameters. In this regard, our approach resembles the contribution of Kholopova (2006) in her thesis. The only difference is that we adopt the continuous version of the geometric average of the payoff instead of the discrete one.

In the second part of the chapter is our second contribution, which is to avoid the use of any approximation to the futures prices (observation equation). In this case we are dealing with a nonlinear system, so we use the particle filtering methodology to estimate the states and the parameters of the system. Finally, we present a simulation study that shows the result of the filter.

---

[1]Through the text, we use the term futures or forward to represent the same contract, since we assume that the interest rate is constant, both products will have the same price.

3.2 Review of Schwartz and Smith (SS) Model

We consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ where $\mathcal{F}_t$ represents the information available at time t. We assume that this space satisfies the "usual conditions", see Oksendal (2003). For any process $X(t)$, we use the notation $E_t[X(T)]$ to denote $E(X(T)|\mathcal{F}_t)$ for the expectation conditional on filtration $\mathcal{F}_t$.

Let us denote by $S(t)$ the spot price of electricity at time t. As in (Schwartz and Smith, 2000) we decompose spot prices into two stochastic factors as

$$\ln S(t) = \chi(t) + \xi(t) + h(t) \tag{3.1}$$

where $\chi(t)$ will be referred to as the short-term deviation in price, $\xi(t)$ is the equilibrium price level and $h(t)$ is a deterministic function. The short-term deviations $\chi(t)$ are assumed to revert toward zero following an Ornstein-Uhlenbeck process

$$d\chi(t) = -\kappa \chi(t) dt + \sigma_1 dW_\chi(t) \tag{3.2}$$

and the equilibrium level $\xi(t)$ is assumed to follow a Brownian motion process given by

$$d\xi(t) = \mu_\xi dt + \sigma_2 dW_\xi(t) \tag{3.3}$$

where under the real probability measure denoted by $\mathbb{P}$, the two Brownian motion are correlated with $\rho dt = dW_\chi(t) dW_\xi(t)$. Spot price process is adapted to the filtration $\mathbb{F}$. In integral form, (3.2) and (3.3) are given by

$$\chi(t) = e^{-\kappa(t-t_0)} \chi(t_0) + \sigma_1 \int_{t_0}^{t} e^{-\kappa(t-u)} dW_\chi(u) \tag{3.4}$$

and

$$\xi(t) = \xi(t_0) + \mu_\xi(t - t_0) + \sigma_2 \int_{t_0}^{t} dW_\xi(u) \tag{3.5}$$

For the valuation of futures, we need to represent the model under the risk neutral measure denoted by $\mathbb{Q}$. Assuming a constant market

prices of risk for both processes $\chi(t)$ and $\xi(t)$ denoted by λ_χ and λ_ξ respectively, we get the following dynamics of both under the risk neutral measure

$$d\chi(t) = (-\kappa\chi(t) - \lambda_\chi)\, dt + \sigma_1 dW_\chi^*(t) \qquad (3.6)$$

and

$$d\xi(t) = (\mu_\xi - \lambda_\xi)\, dt + \sigma_2 dW_\xi^*(t) \qquad (3.7)$$

where W_χ^* and W_ξ^* are correlated standard Brownian motions under $\mathbb{Q}$, and $dW_\chi^* dW_\xi^* = \rho dt$. Denote the current time by t, the time of maturity of the futures by T, the time to maturity τ where $\tau = T - t$, and by T^* a fixed time horizon where $t_0 \leq t \leq T < T^*$. The spot price is still given by

$$\ln S(t) = \chi(t) + \xi(t) + h(t) \qquad (3.8)$$

similar to (3.1), but its dynamics under $\mathbb{Q}$ take into account the dynamics of the underlying factors under $\mathbb{Q}$ given by (3.6) and (3.7).

We know that the futures price denoted by $F(t, T)$, is given by $\mathbf{E}_t[S(T)]$, see Musiela and Rutkowski (1997). Hence, to get this expectation, we express (3.6) and (3.7) in integral form and get

$$\chi(t) = -\frac{\lambda_\chi}{\kappa} + e^{-\kappa(t-t_0)}\left(\chi(t_0) + \frac{\lambda_\chi}{\kappa}\right) + \sigma_1 \int_{t_0}^{t} e^{-\kappa(t-u)} dW_\chi^*(u) \quad (3.9)$$

$$\xi(t) = \xi(t_0) + (\mu_\xi - \lambda_\xi)(t - t_0) + \sigma_2 \int_{t_0}^{t} dW_\xi^*(u) \qquad (3.10)$$

Substituting (3.9) and (3.10) in (3.8), we get

$$\ln(S(t)) = -\frac{\lambda_\chi}{\kappa} + e^{-\kappa(t-t_0)}\left(\chi\left(t_0\right) + \frac{\lambda_\chi}{\kappa}\right) + \sigma_1 \int_{t_0}^{t} e^{-\kappa(t-u)}dW_\chi^*\left(u\right)$$

$$(3.11)$$

$$+ \xi\left(t_0\right) + \left(\mu_\xi - \lambda_\xi\right)\left(t - t_0\right) + \sigma_2 \int_{t_0}^{t} dW_\xi^*\left(u\right) + h(t).$$

Hence, $\ln\left(S\left(T\right)\right)$ is a normal random variable with mean and variance given by

$$\mathbf{E}_t\left[\ln\left(S\left(T\right)\right)\right] = -\frac{\lambda_\chi}{\kappa} + e^{-\kappa(T-t)}\left(\chi\left(t\right) + \frac{\lambda_\chi}{\kappa}\right) \qquad (3.12)$$
$$+ \xi\left(t\right) + \left(\mu_\xi - \lambda_\xi\right)\left(T - t\right) + h(T)$$

and

$$\mathbf{Var}_t\left[\ln\left(S\left(T\right)\right)\right] = \frac{\sigma_1^2}{2\kappa}\left(1 - e^{-2\kappa(T-t)}\right) + \sigma_2^2\left(T - t\right)$$
$$+ \frac{2\rho\sigma_1\sigma_2}{\kappa}\left(1 - e^{-\kappa(T-t)}\right) \qquad (3.13)$$

where $\mathbf{E}_t$ and $\mathbf{Var}_t$ represent the expectation and variance at time T under the risk-neutral measure $\mathbb{Q}$, conditional on the information available at earlier time t. Using (3.12), and (3.13), we get,

$$F\left(t, T\right) = \mathbf{E}_t\left(S\left(T\right)\right) = \exp\left(\mathbf{E}_t\left[\ln S\left(T\right)\right] + \frac{1}{2}\mathbf{Var}_t\left[\ln S\left(T\right)\right]\right)$$
$$(3.14)$$

which can be written as

$$F\left(t, T\right) = \exp\left(e^{-\kappa(T-t)}\chi\left(t\right) + \xi\left(t\right) + A(t, T)\right) \qquad (3.15)$$

where

$$A(t, T) = \left[\frac{\lambda_\chi}{\kappa}\left(e^{-\kappa(T-t)} - 1\right)\right] + \left(T - t\right)\left(\mu_\xi - \lambda_\xi\right) \qquad (3.16)$$
$$+ \frac{1}{2}\mathbf{Var}_t\left[\ln\left(S\left(T\right)\right)\right] + h(T)$$

The logarithm of the futures price, using (3.15) is given as

$$\ln F\left(t, T\right) = e^{-\kappa(T-t)}\chi\left(t\right) + \xi\left(t\right) + A(t, T) \qquad (3.17)$$

3.3 Extension of The Model to Delivery Period

In the electricity market, the market prices of electricity futures are different than the standard futures traded in other markets, such as futures on interest rates or futures on bonds. In the electricity market, the futures prices are based on the arithmetic averages of the spot prices over the delivery period $[T_0, T]$, given by

$$\frac{1}{T - T_0} \int_{T_0}^{T} S(\eta) \, d\eta \tag{3.18}$$

Now, for $t < T$, the futures price is given by

$$F(t, T_0, T) = \mathbf{E} \left\{ \frac{1}{T - T_0} \int_{T_0}^{T} S(\eta) \, d\eta \Big| \mathcal{F}_t \right\} \tag{3.19}$$

where $\mathcal{F}_t = \sigma \{ S(\eta) ; 0 \le \eta \le t \}$. Assume that $S(t) \in L^2(T_0, T) \, \forall t \in [T_0, T]$ and using the linearity of the expectation operator, see (Benth *et al.*, 2008). Then (3.19) can be represented as

$$F(t, T_0, T) = \frac{1}{T - T_0} \int_{T_0}^{T} \mathbf{E}_t [S(\eta)] \, d\eta \tag{3.20}$$

and using the definition of the futures price, (3.20) can be written as

$$F(t, T_0, T) = \frac{1}{T - T_0} \int_{T_0}^{T} F(t, \eta) \, d\eta \tag{3.21}$$

Hence, (3.20) using (3.15) is given by

$$F(t, T_0, T) = \frac{1}{T - T_0} \int_{T_0}^{T} \exp[e^{-\kappa(\eta - t)} \chi(t) + \xi(t) + A(t, \eta)] d\eta \tag{3.22}$$

The integrand represents a lognormal random variable because it represents the exponential of the sum of two normal random variables.

However, the integral itself which represents the futures price is clearly not lognormal, as the sum of lognormal random variables is not lognormal. To treat this system, one way is to approximate the above lognormality by a Gaussian system using the geometric approximation. This will be stated in the next section. Another way to treat this nonlinear system is to use the nonlinear filtering theory with the particle filter. The application of the particle filter is shown in the last section of this chapter.

3.4   Observation Mechanism

As the given data of the observations is available in daily basis and already transformed such that the time-to-delivery $\tau_i = T_0^i - t$ is fixed as a constant through time for each future i. Hence, we need to make adjustments for the futures price using the time to delivery $\tau_i = T_0^i - t$ instead of the time of delivery T_0^i. This means that, for each t, $T_0^i - t$ is set as a constant time period τ_i for $i = 1, \cdots, m$ where m is the number of futures available. Moreover, the delivery period $T - T_0 = \theta$ (1-month) is set as a constant for all the futures.
So, before deriving our observation mechanism, we rewrite the futures price given by (3.22) by using the time-to-delivery variable $x = T - t$:

$$F(t, T_0, T) = \frac{1}{T - T_0} \int_{T_0 - t}^{T - t} g(t, x) dx, \qquad (3.23)$$

where

$$g(t, x) = \exp\{f(t, x)\},$$

$$f(t, x) = A(t, x) + e^{-\kappa x} \chi(t) + \xi(t)$$

and

$$A(t, x) = \frac{\lambda_\chi}{\kappa}(e^{-\kappa x} - 1) + (\lambda_\xi - \mu_\chi)x + h(x + t)$$

$$+ \frac{1}{2}\{\frac{\sigma_\chi^2}{2\kappa}(1 - e^{-2\kappa x}) + \sigma_\xi^2 x + 2\frac{\rho \sigma_\xi \sigma_\chi}{\kappa}(1 - e^{-\kappa x})\} \qquad (3.24)$$

where $h(x + t)$ is a seasonality function identified in Chapter 1. In practice we observe the fixed time-to-delivery futures prices;

$$\bar{F}(t, \tau_i + t, \tau_i + t + \theta) = \frac{1}{\theta} \int_{\tau_i}^{\tau_i + \theta} g(t, x) dx \qquad (3.25)$$

for $\theta = T - T_0$ and $\tau_1 < \tau_2 < \cdots < \tau_n$. Hence we observe the market data $\bar{F}(t, \tau_i + t, \tau_i + t + \theta)$ for $i = 1, 2, \cdots, m$. This equation exhibits the heavy nonlinearity for χ and ξ.

3.5 Geometric Average Approximation

To find an analytical expression for the futures price, we consider the continuous time geometric average approximation to equation (3.25). The geometric average in the continuous form becomes

$$
\begin{aligned}
\frac{1}{\theta} \int_{\tau i}^{\tau_i + \theta} g(t, x) dx &\approx \exp\{\frac{1}{\theta} \int_{\tau i}^{\tau_i + \theta} \log\{g(t, x)\} dx \\
&= \exp\{\frac{1}{\theta} \int_{\tau_i}^{\tau_i + \theta} f(t, x) dx\} \\
&= \exp\left[\frac{e^{-\kappa \tau_i} - e^{-\kappa(\tau_i + \theta)}}{\kappa \theta} \chi(t) + \xi(t) + \tilde{A}(t, \tau_i; \theta)\right], \quad (3.26)
\end{aligned}
$$

where

$$\tilde{A}(t, \tau_i; \theta) = \frac{1}{\theta} \int_{\tau_i}^{\tau_i + \theta} A(t, x) dx. \qquad (3.27)$$

Hence the geometric approximation of the logarithm of the observation becomes

$$y(t, \tau_i; \theta) = \frac{e^{-\kappa \tau_i} - e^{-\kappa(\tau_i + \theta)}}{\kappa \theta} \chi(t) + \xi(t) + \tilde{A}(t, \tau_i; \theta)$$
$$\text{for } i = 1, 2, \cdots, m. \qquad (3.28)$$

3.5.1 Filtering equations

The transition system is represented using the two unobservable factors $\chi(t)$ and $\xi(t)$ given by equations (3.4) and (3.5). Having characterized the state and observation equations, we formulate these equations in a state-space so that we can calculate the likelihood function. For this we need to introduce a noise term in the observation equation. This is often justified by the transaction costs and bid ask spread in the prices. For that, we proceed by a discretization to the transition system. Thus, let the time points be given by $t_k = k\frac{T}{N}$, $k = 1, \cdots, N$ and let $\Delta t = t_k - t_{k-1}$ is the time step, and where N is the number of time periods in the data set. Then we use the following substitutions

$$
\begin{aligned}
dX(t_k) &\cong X(t_k) - X(t_{k-1}) \\
dt &\cong \Delta t
\end{aligned}
\tag{3.29}
$$

Hence, The transition system, can be written as

$$
X_{k+1} = AX_k + U_k + W_k, \quad k = 1, 2, \cdots, N
\tag{3.30}
$$

$$
\text{where} \quad X_k := \begin{bmatrix} \chi(t_k) \\ \xi(t_k) \end{bmatrix}_{2\times 1}
$$

$$
A := \begin{bmatrix} e^{-\kappa\Delta t} & 0 \\ 0 & 1 \end{bmatrix}_{2\times 2}
$$

$$
\tag{3.31}
$$

$$
U := \begin{bmatrix} 0 \\ \mu_\xi \Delta t \end{bmatrix}_{2\times 1}
$$

$$
W_k := \begin{bmatrix} \Delta W_\chi(t_k) \\ \Delta W_\xi(t_k) \end{bmatrix}_{2\times 1}
$$

and

$$
\begin{aligned}
\Delta W_k &\sim N(m^s, \Sigma_s) \\
m^s &= \mathbf{E}(\Delta W_k)' = \begin{bmatrix} 0 & 0 \end{bmatrix}_{1\times 2} \\
\Sigma_s &= \mathbf{Cov}(\Delta W_k) = \mathbf{E}(\Delta W_k \Delta W_k') \\
&= \begin{bmatrix} \left(1 - e^{-2\kappa\Delta t}\right)\frac{\sigma_1^2}{2\kappa} & \left(1 - e^{-\kappa\Delta t}\right)\frac{\rho\sigma_1\sigma_2}{\kappa} \\ \left(1 - e^{-\kappa\Delta t}\right)\frac{\rho\sigma_1\sigma_2}{\kappa} & \sigma_2^2\Delta t \end{bmatrix}_{2\times 2}
\end{aligned}
\tag{3.32}
$$

For the measurement equation, we use a sequence of m futures $f_1, f_2, \cdots, f_m$ with time to delivery $\tau_1, \tau_2, \cdots, \tau_m$, respectively. From equation (3.28), adding the noise term mentioned above, we get

$$Y_k = d_k + CX_k + \eta_k, \quad k = 1, 2, \cdots, N \tag{3.33}$$

where

$$Y_k := \begin{bmatrix} \ln F(t_k, \tau_1 + t_k; \theta) \\ . \\ . \\ \ln F(t_k, \tau_m + t_k; \theta) \end{bmatrix}_{m \times 1} , \quad X_k := \begin{bmatrix} x_1(t_k) \\ x_2(t_k) \end{bmatrix}_{2 \times 1}$$

$$d_k := \begin{bmatrix} \tilde{A}(t_k, \tau_1) \\ . \\ . \\ \tilde{A}(t_k, \tau_m) \end{bmatrix}_{m \times 1} \tag{3.34}$$

where $\tilde{A}(t_k, \tau_i)$ is given in (3.27) and

$$C := \begin{bmatrix} \frac{e^{-\kappa \tau_1} - e^{-\kappa(\tau_1 + \theta)}}{\kappa \theta} & 1 \\ . & . \\ . & . \\ \frac{e^{-\kappa \tau_m} - e^{-\kappa(\tau_m + \theta)}}{\kappa \theta} & 1 \end{bmatrix}_{m \times 2} \tag{3.35}$$

and where the measurement noise $\eta_k \sim N(m_o, \Sigma_o)$,

$$m_o = \mathbf{E}(\eta_k)' = \begin{bmatrix} 0 & \cdot & \cdot & 0 \end{bmatrix}_{1 \times m}$$

$$\Sigma_o = \mathbf{Cov}[(\eta_k)] = \mathbf{E}(\eta_k \eta_k') = R$$

where $R = \sigma_\epsilon I_m$ is a constant diagonal matrix of size m

The specifications of the state space system is completed by introducing further assumptions:

1. The initial state space vector $X_{t_0} = (\chi(t_0), \xi(t_0))'$ has a mean of $\hat{\mathbf{x}}_0$ and a covariance matrix $\mathbf{P}_0$

$$
\begin{aligned}
\mathbb{E}(X_{t_0}) &= \hat{\mathbf{x}}_0 \\
\mathbb{V}ar(X_{t_0}) &= \mathbf{P}_0
\end{aligned}
$$

2. The distributions of ΔW_k and η_k are uncorrelated with the initial state, i.e.

$$
\begin{aligned}
\mathbb{E}\left(\Delta W_k X'_{t_0}\right) &= 0 \quad \text{for } k = 1, \cdots, N \\
\mathbb{E}\left(\eta_k X'_{t_0}\right) &= 0 \quad \text{for } k = 1, \cdots, N
\end{aligned}
$$

3.5.2 Likelihood function

Now we are in a position to calculate the likelihood function as explained in Section 2.3.2 of Chapter 2. The log-likelihood function is given as

$$
\ln L(y_N, \psi) = \sum_{k=1}^{N} \ln(p(y_k, \psi | \mathcal{Y}_{k-1})) \tag{3.36}
$$

$$
= -\frac{nN}{2} \ln 2\pi - \frac{1}{2} \sum_{k=1}^{N} \ln |\det \mathbf{Q}_k| - \frac{1}{2} \sum_{k=1}^{N} \omega'_k \mathbf{Q}_k^{-1} \omega_k
$$

where

$$
\begin{aligned}
\omega_k &= Y_k - \bar{\mathbf{y}}_k, \quad k = 1, \cdots, N \\
\bar{\mathbf{y}}_k &= C_k \hat{X}_{k|k-1} + d_k \\
\mathbf{Q}_k &= C_k \mathbf{P}_{k|k-1} C'_k + \Sigma_0
\end{aligned}
$$

3.6 Sensitivity Analysis to The MLE

Our aim in this section is to examine the sensitivity of changing the covariance of the observation noise and changing the parameters, on the corresponding changing value of the maximum Likelihood function. For that we need to simulate the data first.

3.6.1 The generation of data for simulation

We set the parameters for simulation as follows which are denoted by the true parameters given by,

$$\kappa=1.321 \quad \lambda_\chi= 0.623 \quad \sigma_1 = 0.7 \quad \mu_\xi = .1$$
$$\lambda_\xi = 0.1 \quad \sigma_2 = 0.3 \quad \rho=0.6$$

Also, we consider two values for the standard deviation "square root of the covariance" of the artificial observation noise,

$$\sigma_\epsilon = 0.01 \text{ and } \sigma_\epsilon = 0.05.$$

To simplify our simulations, we only generate the observation data from a model where seasonality and linear trend are removed completely. In other words, we assume that the deterministic component of the spot price model given by the function $h(t)$ has been found as discussed in Chapter 1. Figure 3.1 show the result of the generated term structure of futures when $\sigma_\epsilon = 0.01$, while Figure 3.2 show the term structure when the $\sigma_\epsilon = 0.05$

Figure 3.1: Simulated $y(t, \tau_i)$ with $\sigma_\epsilon = 0.01$

Figure 3.2: Simulated $y(t, \tau_i)$ with $\sigma_\epsilon = 0.05$

3.6.2   The effect of the parameters on the MLE

To check the effect of the change of the parameters values on the ML function, we assume that the parameters are given in the following range, summarized in Table 3.1.

Parameter	Lower bound	Upper bound
κ	0.01	10
λ_χ	1.2	3
σ_1	0.3	1
μ_ξ	0.1	2
λ_ξ	0.5	2
σ_2	0.1	.1
ρ	-0.9	0.9

Table 3.1: Lower and upper bounds for all parameters

Now we consider five different parameters and the resulting corresponding values of its ML function. These five cases for the parameters are chosen as follows:

Case 1	=	Lower bound
Case 2	=	Lower bound + 0.01
Case 3	=	True value
Case 4	=	Upper bound - 0.01
Case 5	=	Upper bound

Considering two different cases for σ_ϵ, the values of the corresponding log-likelihood function are shown in Figure 3.3,

Figure 3.3: Change of log-likelihood around $\sigma_\epsilon = 0.01$ (left) and 0.05 (right)

From these figures, we find that whenever the value of the artificial noise becomes bigger, the value of log-likelihood becomes small. Furthermore the MLE seems to have a global maximum at the true value of the parameters.

Now we slightly perturbed each of the parameters, keeping all the other ones fixed, and get the corresponding ML values. The result of this analysis is shown in the following figures.

Figure 3.4: Log likelihood for changing κ

Figure 3.5: Log likelihood for changing λ_χ

Figure 3.6: Log likelihood for changing σ_1

Figure 3.7: Log likelihood for changing μ_ξ

Figure 3.8: Log likelihood for changing λ_ξ

Figure 3.9: Log likelihood for changing σ_2

Figure 3.10: Log likelihood for changing ρ

Figure 3.11: Log likelihood for changing σ_ϵ, keeping the other parameters fixed to their true values, (left: $\sigma_\epsilon = 0.01$, right: $\sigma_\epsilon = 0.05$).

Now from above figures, we find that the precise estimates for the unknown parameters using the MLE method will be difficult when using the usual optimization algorithm with the gradient method. Notice that the values of likelihood with respect to the unknown parameters do not change, and the plot is almost flat. One possible option to get the parameters is to apply the Genetic Algorithm Tool-

box GA in MATLAB. This will be performed in Chapter 4 when dealing with the infinite dimensional system. At this stage, we need to recognize that the last figure above implies that it will be hard to identify the artificially added noise covariance, because the likelihood function is also monotone with respect to σ_ϵ. In the next Chapter, a new modeling approach is proposed where the artificial noise is build in within the model and it is not exogenously imposed to the observation equation.

3.7 The Arithmetic Average Case

In this section, our aim is to avoid the use of any approximation to the payoff of the futures. For that, we propose the use of particle filtering algorithm as a method for parameter estimation. Hence, our futures price will be given by its non-linear form,

$$y_i(t_j) \equiv y(t_j, \tau_i; \theta) = \bar{F}(t_j, \tau_i + t_j, \tau_i + t_j + \theta), \qquad (3.37)$$

where

$$\bar{F}(t_j, \tau_i + t_j, \tau_i + t_j + \theta) = \frac{1}{\theta} \int_{\tau_i}^{\tau_i+\theta} g(t_j, x) dx.$$

Now, we assume that the observation is changed as a result of, for example, bid-ask spread to $\bar{F}$. So, we impose extra noise term affecting the original observation. As a result, the observation will be given by

$$y_i(t_j) = \bar{F}(t_j, \tau_i + t_j, \tau_i + t_j + \theta) \exp(\epsilon_j) \qquad (3.38)$$

where ϵ_j is a zero mean Gaussian white noise with

$$E\{\epsilon_j \epsilon_k\} = \sigma_\epsilon^2 \delta_{jk}.$$

Now the log price $\tilde{y}_i(t_j) = \log y_i(t_j)$ becomes

$$\tilde{y}_i(t_j) = \log \bar{F}(t_j, \tau_i + t_j, \tau_i + t_j + \theta) + \epsilon_j.$$

3.8   The Discrete Version of Models

Now we present the discrete version of our system model and we also reset $\lambda_\xi = \lambda_\xi - \mu_\xi$. We set

$$t_0 < t_1 < \cdots < t_n, \text{for } \Delta t = t_{j+1} - t_j.$$

Hence

$$
\begin{aligned}
\chi_{j+1} &= \chi_j + (-\kappa\chi_j - \lambda_\chi)\Delta t + \sigma_\chi \Delta W_{\chi j} & (3.39)\\
\xi_{j+1} &= \xi_j + \lambda_\xi \Delta t \sigma_\xi + \Delta W_{\xi j}. & (3.40)
\end{aligned}
$$

The observation mechanism is given by

$$\tilde{y}_\ell(t_j) = \log\left[\frac{1}{\theta}\int_{\tau_\ell}^{\tau_\ell+\theta} \exp\{A(t_j,x) + e^{-\kappa x}\chi_j + \xi_j\}dx\right] + \epsilon_j, \quad (3.41)$$

for $\ell = 1, 2, \cdots, m$. We set

$$\tilde{Y}_j = [\tilde{y}_\ell(t_j)]_{m\times 1} \tag{3.42}$$

3.9   Particle Filter Algorithm

It is again possible to identify the system parameters by using the maximum likelihood method. Here noting that the particle filter works for the nonlinear systems, we estimate our parameters by augmenting the unknown parameters with the state of the system. Suppose that the parameters

$$\Theta = [\kappa, \lambda_\chi, \sigma_\chi, \lambda_\xi, \sigma_\xi, \rho];$$

are random variables with uniform distributions with known bounds, and consider the augmented system state variable (χ, ξ, Θ) with

$$\chi(0) \in N(m_1, \sigma_1), \ \xi(0) \in N(m_2, \sigma_2),$$

where $m_1, \sigma_1, m_2, \sigma_2$ are known. Now our generating algorithm for the particle filter is as follows:

- Generate NS particles: for $i = 1, 2, \cdots, NS$

$$\chi^{(i)}(0) \in N(m_1, \sigma_1) \text{ and } \xi^{(i)}(0) \in N(m_2, \sigma_2)$$

and

$$\Theta^{(i)} \in \text{Some uniform distribution.}$$

- Generate the system states $\chi_j^{(i)}, \xi_j^{(i)}$ from (3.39) and (3.40).

- Calculate

$$h(j, i) = \left[\log\left[\frac{1}{\theta} \int_{\tau_k}^{\tau_k + \theta} \exp\{A(t_j, x) + e^{-\kappa x}\chi_j^{(i)} + \xi_j^{(i)}\}dx\right] \right]_{m \times 1}$$

for $i = 1, 2, \cdots, NS$

- Get the likelihood;

$$p(\tilde{Y}_j | \chi_j^{(i)}, \xi_j^{(i)}, \Theta^{(i)}) = \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} \exp\{-\frac{1}{2\sigma_\epsilon^2}||\tilde{Y}_j - h(j, i)||^2\}$$

for $i = 1, 2, \cdots, NS$.

- Updated weight $w_j^{(i)}$ is obtained for $W_0 = [1/NS]_{NS \times 1}$,

$$w_j^{(i)} = w_{j-1}^{(i)} p(\tilde{Y}_j | \chi_j^{(i)}, \xi_j^{(i)}, \Theta^{(i)}).$$

- Normalize

$$\hat{w}_j^{(i)} = \frac{w_j}{\sum_{i=1}^{NS} w_j^{(i)}}.$$

- Get the conditional probability densities;

$$p(\chi_j | \tilde{Y}_j) = \sum_{i=1}^{NS} \hat{w}_j^{(i)} \delta(\chi_j - \chi^{(i)}), \tag{3.43}$$

$$p(\xi_j | \tilde{Y}_j) = \sum_{i=1}^{NS} \hat{w}_j^{(i)} \delta(\xi_j - \xi_j^{(i)}), \tag{3.44}$$

$$p(\Theta | \tilde{Y}_j) = \sum_{i=1}^{NS} \hat{w}_j^{(i)} \delta(\Theta - \Theta_j^{(i)}). \tag{3.45}$$

- We use the systematic resampling procedure, if we need.

3.10   Simulation Studies

For the simulation study, we set the initial conditions and system parameters as

$$[\chi(0)\ \xi(0)] = [0.02\ 0.5]$$

$$[\kappa\ \lambda_\chi\ \sigma_\chi\ \lambda_\xi\ \sigma_\xi\ \rho\ \sigma_\epsilon] = [1.321\ 0.623\ 0.2\ 0.01\ 0.2\ 0.6\ 0.1]$$

and the the seasonality function $h(t)$ is assumed to be known and given by

$$\begin{aligned} h(t) = {}& 2(27.9\sin(2\pi t) + 40\cos(2\pi t) \\ & + \ 13.9\sin(4\pi t)) + 11.5\cos(4\pi t))/251 + 0.27t + 3.04. \end{aligned}$$

We also set the time difference $dt = 0.004 = \frac{1}{250}$ and $\tau_i = 0.0873$ for all $i = 1, 2, \cdots, 44$. The simulated observation data is given in Figure 3.12, and we also present the log data with the observation noises in Figure 3.13.

Figure 3.12: Simulated observation

Figure 3.13: Simulated observation with the extra noise term

For performing the proposed particle filter algorithm, we set the initial probability densities for the initial states as

$$\chi(0) \ \in \mathcal{N}(0.02, 0.01) \text{ and } \ \xi(0) \in \mathcal{N}(0.5, 0.1)$$

and the bounds for the parameters are set as

$$1 < \kappa < 2,$$
$$0.1 < \lambda_\chi < 1,$$
$$0.1 < \sigma_\xi < 1,$$
$$0.001 < \lambda_\xi < 0.1,$$
$$0.1 < \sigma_\xi < 1,$$
$$0.1 < \rho < 1.$$

The number of particle is given by $NS = 350$ and we perform the algorithm proposed in the previous section for $\sigma_\epsilon = 0.3, 0.6$. In the original data, the true value for σ_ϵ is set to be 0.1.

We demonstrate the results for estimating the states χ, ξ and the parameter Θ for the two cases ($\sigma_\epsilon = 0.3, 0.6$). It should be noted that

the results for $\sigma_\epsilon = 0.6$ are better than the results for σ_ϵ, even though the value of $\sigma_\epsilon = 0.3$ is near the true value of σ_ϵ.

Figure 3.14: Estimation of χ process

Figure 3.15: Estimation of ξ process

Figure 3.16: True and estimated κ

Figure 3.17: True and estimated λ_χ

Figure 3.18: True and estimated σ_χ

Figure 3.19: True and estimated λ_ξ

Figure 3.20: True and estimated σ_ξ

Figure 3.21: True and estimated ρ

3.11 Concluding Remarks

In this chapter, we have described the two factor model of Schwartz and Smith (2000) used to model the behavior of futures contracts for energy commodities. In this model, the futures prices are defined in terms of a spot price where the log of this price is represented by two unobservable stochastic factors. We use the geometric approximation to take care of the delivery period. Hence, we are able to find an analytical solution for the futures price. Then, we implement a sensitivity analysis of the ML function to the parameters and the extra noise imposed on the measurements. It seems from the sensitivity analysis that the parameters are difficult to find using standard optimization methods. The second part of the chapter focuses on the use of particle filtering method to estimate the state and the parameters of the model. We implement a simulation study to check the feasibility of this approach.

Chapter 4

Infinite Dimensional Kalman Filtering Applied to New Energy Spot Model

4.1 Introduction

As we have discussed in the previous chapter, an important issue is the estimation of the model parameters, where we used the maximum likelihood method. In order to derive the likelihood functional, the Kalman filter is constructed in linear models. However, in energy futures modeling, in spite of the mathematically elegant derivation of the futures prices which are the observed data, one needs to add some *ad hoc* observation noise in order to derive the Kalman filter. This assumption has been made by numerous authors, either in the commodity or interest rate markets, see Elliott and Hyndman (2007) and its reference. The additional noise in the observation has been interpreted to take into account bid-ask spreads, price limits, or errors in the data. The argument is clearly forced and unconvincing. Here, we approach the modeling differently. In our setup, on one hand, the added measurement noise is built in within the model. On the other hand, the modeling of the correlation structure between the futures (observation) is a natural component of our formulation.

In this Chapter, we follow the approach of Aihara and Bagchi (2010*a*), Aihara and Bagchi (2010*b*) applied to the interest rate market. The

main idea is to assume that the term structure of futures prices on electricity given by Schwartz and Smith (2000) model is affected by an error term represented by a stochastic integral that generates infinite dimensional noise as it should depend on all time of, or to maturity. This extended model does not need addition of artificial noises to the observation equation, in order to use the filtering methodology. Using the martingale property of the modified price under the risk neutral measure, we derive the arbitrage free model for the spot and futures prices. We first derive the futures price formula taking into account the arithmetic average of the payoff function. We then approximate this by using the geometric average and use infinite-dimensional Kalman filter to estimate the parameters.

In Section 4.2, we present the new model for the future price for one maturity date T where we introduce a perturbation term to the futures price. In Section 4.3, we focus our attention on the electricity futures situation which are based on the arithmetic averages of the spot prices over a delivery period. In Section 4.4, we derive the explicit relation between the observed futures prices and the factor processes. In Section 4.6, we discuss the parameter estimation problem in relation to the covariance of the noise term of the observation and derive a quasi likelihood functional. In Section 4.7, we state the optimal filtering and kernel equations. The last two Sections contain the simulation work and conclusion, respectively.

As discussed in Chapter 3, the spot price $S(t)$ of a commodity (electricity) at time t is given by

$$\ln(S(t)) = \chi(t) + \xi(t) + h(t) \tag{4.1}$$

where $\chi(t)$ represents the short-term deviation in the price, $\xi(t)$ is the equilibrium price level and $h(t)$ is a deterministic seasonality function. Assume that the risk-neutral stochastic process for the two factors are of the form

$$\begin{cases} d\chi(t) = (-\kappa\chi(t) - \lambda_\chi)dt + \sigma_\chi dW_\chi^*(t) \\ d\xi(t) = \lambda_\xi dt + \sigma_\xi dW_\xi^*(t) \end{cases} \tag{4.2}$$

where W_χ^* and W_ξ^* are correlated standard Brownian motions, where $dW_\chi^* dW_\xi^* = \rho dt$. We denote the current time by t, the maturity of

the futures by T, and by T^* a fixed time horizon where $t \leq T < T^*$. The futures price $F(t, T - t)$ is given by

$$F(t, T - t) = \exp(B(T - t)\chi(t) + C(T - t)\xi(t) \\ + A(t, T - t)) \qquad (4.3)$$

where

$$B(T - t) = e^{-\kappa(T-t)}, \ C(T - t) = 1 \qquad (4.4)$$

$$A(t, T - t) = \frac{\lambda_\chi}{\kappa}\left(e^{-\kappa(T-t)} - 1\right) + \lambda_\xi(T - t) \\ + \frac{1}{2}\sigma_A^2(T - t) + h(T) \qquad (4.5)$$

and

$$\sigma_A^2(T - t) = \frac{\sigma_\chi^2}{2\kappa}\left(1 - e^{-2\kappa(T-t)}\right) + \sigma_\xi^2(T - t) \\ + 2\frac{\rho\sigma_\chi\sigma_\xi}{\kappa}\left(1 - e^{-\kappa(T-t)}\right)$$

4.2 A New Model for the Electricity Prices

We start with the Schwartz-Smith model discussed in detail in the previous section. We assume that the correct model for the spot price is not exactly the same as in (4.1), but is close to it. Given this, the futures price will be somewhat perturbed from the formula given in (4.3). Hence, suppose that the correct futures price at time t where $t \leq T$ is given by

$$F^{corr}(t, T - t) = \exp[\bar{B}(T - t)\chi(t) + \bar{C}(T - t)\xi(t) \qquad (4.6)$$

$$+ \bar{A}(t, T - t) + \int_0^t \sigma dw^*(s, T - s)]$$

where

$$\int_0^t \sigma dw^*(s, T - s) = \sum_{k=1}^\infty \int_0^t \sigma\frac{1}{\lambda_k}e_k(T - s)\,d\beta_k(s) \qquad (4.7)$$

and where e_k is a sequence of differentiable functions forming an orthonormal basis in $L^2(0, T^*)$ and $\{\beta_k(t)\}$ are mutually independent Brownian motions processes. See (section 2.5.1 for details). Let $q(x, y)$ represent the correlation of $w^*(t, x)$ and $w^*(t, y)$. The extra stochastic integral term (4.7) which appears in (4.6), represents the modeling error between the futures price given by (4.3) and the correct futures price. The term $dw^*(s, T - s)$ represents a two parameter Brownian motion, it has both time and spatial dimensions. The first dimension is a representation of the uncertainty through time while the latter is for the uncertainty for futures with different maturities. The integral term shows that these uncertainties increase.

When $T - t = 0$, the correct spot price process is given by

$$S^{corr}(t) \equiv F^{corr}(t, 0) \tag{4.8}$$

To get the corresponding (correct) dynamics for the spot, we need the dynamics of the futures taking into account that this dynamics under the risk-neutral measure is a martingale. Applying Ito's formula to (4.6) , we get

$$\frac{dF^{corr}(t, T - t)}{F^{corr}(t, T - t)} = \left[\frac{d\bar{A}(t, T - t)}{dt} + \frac{d\bar{B}(T - t)}{dt}\chi(t) \right. \tag{4.9}$$

$$+ \frac{d\bar{C}(T - t)}{dt}\xi(t) + \bar{B}(T - t)(-\kappa\chi(t) - \lambda_\chi) + \bar{C}(T - t)\lambda_\xi$$

$$+ \frac{1}{2}\sigma^2 q(T - t, T - t) + \frac{1}{2}\sigma_\chi^2 \bar{B}^2(T - t) + \frac{1}{2}\sigma_\xi^2 \bar{C}^2(T - t)$$

$$+ \rho\sigma_\chi\sigma_\xi \bar{B}(T - t)\bar{C}(T - t) \bigg] dt + \sigma_\chi \bar{B}(T - t)dW_\chi^*(t)$$

$$+ \sigma_\xi \bar{C}(T - t)dW_\xi^*(t) + \sigma dw^*(t, T - t).$$

For the futures price to be a martingale, the dt-part of (4.9) has to be zero. For that, we get $\bar{B}(t, T - t) = B(t, T - t)$, $\bar{C}(t, T - t) = C(t, T - t)$ given by (4.4) and $\bar{A}$ satisfies

$$\frac{d\bar{A}(t, T - t)}{dt} - \lambda_\chi e^{-\kappa(T-t)} + \lambda_\xi$$

$$+ \frac{1}{2}\sigma_\chi^2 e^{-2\kappa(T-t)} + \frac{1}{2}\sigma_\xi^2 + \rho\sigma_\chi\sigma_\xi e^{-\kappa(T-t)}$$

$$+ \frac{1}{2}\sigma^2 q(T - t, T - t) = 0, \quad \bar{A}(T, 0) = h(T).$$

The solution of this is given by

$$\bar{A}(t, T-t) = A(t, T-t) + \frac{1}{2}\sigma^2 \int_0^{T-t} q(x, x)\, dx$$

where $A(t, T-t)$ is given by (4.5). Substituting $\xi(t)$ in (4.6), we obtain

$$F^{corr}(t, T-t) = \exp(B(T-t)\chi(t) + \tilde{A}(t, T-t) \qquad (4.10)$$
$$+ \int_0^t [\sigma dw^*(s, T-s) + \sigma_\xi dW_\xi^*(s)])$$

where

$$\tilde{A}(t, T-t) = \bar{A}(t, T-t) + \lambda_\xi t + \xi(0). \qquad (4.11)$$

Using (4.8), the correct spot price process is given by

$$S^{corr}(t) = F^{corr}(t, 0)$$
$$= \exp(\chi(t) + h(t) + \lambda_\xi t$$
$$+ \int_0^t \{\sigma dw^*(s, t-s) + \sigma_\xi dW_\xi^*(s)\}).$$

After here, we omit writing the expression "corr" for $S(t)$ and $F(t, T)$ processes.

4.3 Practical Model for the Electricity Prices (Geometric Average Approximation)

As we have done in Section 3.3 in the previous chapter, we are interested in pricing the electricity futures contracts which are based on the arithmetic average of the spot prices over a delivery period $[T_0, T]$. Hence, we follow the same procedure and we get the following formula for the futures price

$$F(t, T_0, T) = \frac{1}{T - T_0} \int_{T_0}^T F(t, \eta)\, d\eta,$$

This price using (4.10) satisfies

$$F(t, T_0, T) = \frac{1}{T - T_0} \int_{T_0}^{T} \exp \Big[B(\eta - t) \chi(t) + \tilde{A}(t, \eta - t)$$

$$+ \int_{0}^{t} \{\sigma dw^*(s, \eta - s) + \sigma_\xi dW_\xi^*(s)\} \Big] d\eta, \qquad (4.12)$$

where B and $\tilde{A}$ satisfy (4.4) and (4.11), respectively. To simplify the pricing formula, we adapt again the same method of Section 3.3 in using the geometric average approximation. In this case the futures price for this average satisfies

$$\tilde{F}(t, T_0, T) = E\{\exp\{\frac{1}{T - T_0} \int_{T_0}^{T} \log S(\eta) d\eta\} | \mathcal{F}_t\}$$

and for $t < T_0$

$$\tilde{F}(t, T_0, T) = \exp\{\frac{1}{T - T_0} \int_{T_0}^{T} [B(\eta - t) \chi(t) + \tilde{A}(t, \eta - t) \qquad (4.13)$$

$$+ \int_{0}^{t} \{\sigma dw^*(s, \eta - s) + \sigma_\xi dW_\xi^*(s)\}] d\eta\}$$

where it is obvious that B and $\tilde{A}$ satisfy the same equations (4.4) and (4.11), respectively.
Using $x = \eta - t$ in (4.13) and setting

$$f(t, x) = B(x) \chi(t) + \tilde{A}(t, x) + \int_{0}^{t} \{\sigma dw^*(s, x + t - s),$$

we also have

$$\tilde{F}(t, T_0, T) = \exp\{\frac{1}{T - T_0} \int_{T_0 - t}^{T - t} f(t, x) dx\}.$$

4.4 Observation Mechanism (Geometric Average Approximation Case)

In practice, the observed real data for the futures are on daily basis and transformed such that the time-to-delivery $\tau = T_0 - t$ is fixed as a constant. See section 3.5.1 for similar adjustment to the futures price formula. For each t, the time to delivery $T_0^i - t$ is set as a constant time period τ_i for $i = 1, 2, \cdots, m$ and $T - T_0$ is set as a constant θ (1 month) for all i. Hence our observation data becomes

$$y(t, \tau_i) = \log \tilde{F}(t, \tau_i + t, \tau_i + t + \theta). \qquad (4.14)$$

Setting $z = \eta - t$ in (4.13), $y(t, \tau_i)$ satisfies

$$y(t, \tau_i) = \frac{1}{\theta}\{ \int_{\tau_i}^{\theta+\tau_i} B(z)\,dz\chi(t) + \int_{\tau_i}^{\theta+\tau_i} \tilde{A}(t, z)\,dz$$

$$+ \int_0^t \int_{\tau_i}^{\theta+\tau_i} [\sigma dw^*(s, z+t-s) + \sigma_\xi dW_\xi^*(s)]dz\} \qquad (4.15)$$

In differential form, this observation mechanism becomes

$$dy(t, \tau_i) = \left\{ -\kappa H(\tau_i)\chi(t) + \frac{1}{\theta}(f_w(t, \theta + \tau_i) - f_w(t, \tau_i)) \right.$$

$$+ \frac{1}{\theta}\int_{\tau_i}^{\theta+\tau_i} \frac{\partial\tilde{A}(t, z)}{\partial t}dz - \lambda_\chi H(\tau_i)\}dt$$

$$+ \sigma_\chi H(\tau_i)dW_\chi^*(t) + \sigma_\xi dW_\xi^*(t) + \frac{\sigma}{\theta}\int_{\tau_i}^{\theta+\tau_i} dw^*(t, z)dz, \qquad (4.16)$$

where

$$df_w(t, x) = \frac{\partial f_w(t, x)}{\partial x}dt + \sigma dw^*(t, x) + \sigma_\xi dW_\xi^*(t) \qquad (4.17)$$

$$f_w(t, 0) = 0 \qquad (4.18)$$

and

$$H(\tau) = \frac{1}{\theta}\int_{\tau_i}^{\theta+\tau_i} B(z)dz = \frac{1}{\kappa\theta}[e^{-\kappa\tau_i} - e^{-\kappa(\tau_i+\theta)}]$$

We set the observation state as

$$Y(t) = [y(t, \tau_1), y(t, \tau_2), \cdots, y(t, \tau_m)].$$

Remark 3. *Notice that the identification usually is performed under the real world measure denoted by $\mathbb{P}$, because the historical data of the futures are available to us under this measure. In our derivation, the extra infinite dimensional noise term, see equation (4.6) is added to the futures price under the risk neutral measure $\mathbb{Q}$. This implies that the market price of risk terms are included in (4.17). In this chapter and also in the next chapter, we assume that this perturbation is small and we can safely ignore this market price of risk. Theoretically it is possible to include this risk which is denoted by $\lambda_w(x)$. The Brownian motion process $w(t, x)$ in the physical measure $\mathbb{P}$ is transformed to the new Brownian motion process $w^*(t, x)$ in the risk neutral measure $\mathbb{Q}$ such that*

$$w(t, x) = w^*(t, x) - \lambda_w(x)t,$$

if the following Novikov condition is satisfied:

$$\sum_{k=1}^{\infty} \frac{1}{\lambda_k^2} (\lambda_w, e_k)^2 < Constant.$$

Where λ_k is defined by (2.30) in Chapter 2, see Da Prato and Zabczyk (1992) for details.

Hence, For including this market price of risk, we only need to reset the term $\frac{1}{2}\sigma^2 q(x, x)$ as $\frac{1}{2}(\sigma^2 q(x, x) + \sigma \lambda_w(x))$. For more details, we refer to Aihara and Bagchi (2010a).

4.5 The Kalman Filter Algorithm

The main purpose of introducing Section 2.5 is to show that the finite-dimensional systems (2.13) and (2.14) could be extended to the infinite-dimensional ones. Here we propose a heuristic review of the infinite-dimensional stochastic systems theory to be applied to the modeling of the energy futures market without using any functional analysis tools.

The system state $X(t) \in \mathbb{R}^n$ in (2.13) is extended to $X(t, x)$ where x denotes the spatial variable in $]0, \hat{T}[$. (In the mathematical finance

field, x normally represents the time-to-maturity or maturity variables.) i.e.,

$$X(t) \in \mathbb{R}^n \text{ is now replaced by } X(t, x) \text{ for } x \in]0, \hat{T}[\qquad (4.19)$$

The system matrix $A(t)$ in (2.13) is now replaced by a partial differential operator with respect to x. Here we only treat the first order partial differential operator, i.e.,

$$A(t) \in \mathbb{R}^n \times \mathbb{R}^n \text{ is replaced by } A = \frac{\partial}{\partial x} \qquad (4.20)$$

where we use the same symbol A but it should be noted that the partial differential operator A is not a bounded operator, as in the finite dimensional case where A is given by a finite dimensional matrix[1] Now our infinite-dimensional system becomes

$$dX(t, x) = A\, X(t, x)dt + dw^*(t, x), \qquad (4.21)$$

where the Brownian motion $w^*(t, x)$ which depends on the new variable x has been described in Subsection 2.5.1. For a fixed x, this new Brownian motion $w(t, x)$ is just a standard Brownian motion, where,

$$E\{w^*(t, x)w^*(t, y)\} = q(x, y)t. \qquad (4.22)$$

4.5.1 Observation mechanism with additive noise

Although equation (4.21) represents an infinite dimensional system, the corresponding observation mechanism in practice is a finite dimensional one. For each t, we can not observe the state $X(t, x)$ for all x. One possible observation mechanism for the state $X(t, x)$, adjusting the observation matrix C in (2.14) is given by

$$CX(t, x) = \frac{1}{\theta} \left[\int_{T_1}^{T_1+\theta} X(t, x)dx, \cdots, \int_{T_m}^{T_m+\theta} X(t, x)dx \right]', \quad (4.23)$$

---

[1]For a mathematical treatment of a non-bounded operator, we need tools from semigroup or variational form in functional analysis.

where $0 < T_1 < \cdots < T_m < \hat{T}$ and $\theta > 0$, and this form often appears when modeling the energy market. This occurs when we connect the system factor model of the spot price with the observations represented by the futures prices with different times to delivery.

The infinite dimensional Kalman filter can be summarized as follows:

$$\begin{cases} dX(t,\cdot) = AX(t,\cdot)dt + dw^*(t,\cdot) \\ \quad dY(t) = CX(t,\cdot)dt + dW^o(t) \end{cases} \qquad (4.24)$$

where $W^o(t)$ is still a finite-dimensional Brownian motion as used in (2.14).

With some minor technical modifications of subsection 2.5.2, Kalman filter algorithm takes the form

$$\begin{cases} d\hat{X}(t,\cdot) = A\hat{X}(t,\cdot)dt + K(t,\cdot)\Sigma_o^{-1}d\nu(t) \\ d\nu(t) = dY(t) - C\hat{X}(t,\cdot)dt \\ K(t,\cdot) = P(t)C^* \end{cases} , \qquad (4.25)$$

It has the same form of the finite dimensional case, except that we need to define the gain K and its related operators $P(t)$ and C^*. Before showing the exact forms of P and C^*, we shall first show the estimation error, i.e., the error covariance for this system. In this case, $\hat{X}(t,x)$ is a function of the spatial variable x and we need to define the estimation error for one point x in relation to another point in the spatial domain y, i.e.,

$$p(t,x,y) = E\{(X(t,x) - \hat{X}(t,x))(X(t,y) - \hat{X}(t,y))|\mathcal{Y}_t\}. \quad (4.26)$$

to characterize the kernel of the covariance operator $P(t)$. The exact form of $P(t)C^*$ is then given by

$$\begin{aligned} P(t)C^* &= \int_0^{\hat{T}} p(t,\cdot,y)C^*dy \\ &= \frac{1}{\theta}\left[\int_{T_1}^{T_1+\theta} p(t,\cdot,y)dy, \cdots, \int_{T_m}^{T_m+\theta} p(t,\cdot,y)dy\right]. \end{aligned}$$

Hence to solve the above Kalman filter we need the exact equation for $p(t,x,y)$. To derive this equation, we set $A = \frac{\partial}{\partial x}$ and $A^* = \frac{\partial}{\partial y}$ in

the finite dimensional Riccati equation and we get

$$
\frac{\partial p(t,x,y)}{\partial t} = \frac{\partial p(t,x,y)}{\partial x} + \frac{\partial p(t,x,y)}{\partial y} + q(x,y)
$$

$$
-\Sigma_{i=1}^{m}\Sigma_{j=1}^{m}\frac{1}{\theta}\int_{T_i}^{T_i+\theta} p(t,x,z_1)dz_1[\Sigma_o^{-1}]_{i,j}\frac{1}{\theta}\int_{T_j}^{T_j+\theta} p(t,z_2,y)dz_2.
$$

4.5.2   Observation mechanism without additive noise

In the financial market, we often encounter the case that some parts of system state can be observed without observation noise,i.e.

$$
Y(t) = CX(t,\cdot) \tag{4.27}
$$

The differential form of the above equation becomes

$$
dY(t) = C(AX(t,\cdot))dt + Cdw^*(t,\cdot), \tag{4.28}
$$

where Cdw^* is a part of the system noise. Now we shall calculate this noise covariance. It is easy to show that

$$
\frac{E\{Cdw^*(t,\cdot)\ Cdw^*(t,\cdot)'\}}{dt} = \frac{1}{\theta^2}\left[\int_{T_i}^{T_i+\theta}\int_{T_j}^{T_j+\theta} q(x,y)dxdy\right]_{m\times m} \tag{4.29}
$$

The infinite-dimensional noise $w(t,x)$ always guarantees the positivity of the above matrix (4.29) for finite m. This property was also checked in the US-bond market by Aihara and Bagchi (2010b). From this fact, we will develop our modeling procedure by using the infinite-dimensional systems, which will be discussed further in this chapter. Now, we present the Kalman filter in its original form:

$$
\begin{cases} d\hat{X}(t,x) = \frac{\partial \hat{X}(t,x)}{\partial x}dt + K(t,x)[C(CQ)^*]^{-1}(dY(t) \\ \quad -\frac{1}{\theta}[X(t,T_i+\theta)-X(t,T_i)]_{m\times 1}dt) \\ K(t,x) = \frac{1}{\theta}\left[\int_{T_i}^{T_i+\theta} p(t,x,y)dy\right]_{1\times m} + \frac{1}{\theta}\left[\int_{T_i}^{T_i+\theta} q(x,y)dy\right]_{1\times m} \end{cases} \tag{4.30}
$$

where

$$
[C(CQ)^*] = \frac{1}{\theta^2}\left[\int_{T_i}^{T_i+\theta}\int_{T_j}^{T_j+\theta} q(x,y)dxdy\right]_{m\times m}. \tag{4.31}
$$

The error covariance $p(t, x, y)$ is a solution of

$$\frac{\partial p(t, x, y)}{\partial t} = \frac{\partial p(t, x, y)}{\partial x} + \frac{\partial p(t, x, y)}{\partial y} + q(x, y)$$

$$-\Sigma_{i=1}^{m}\Sigma_{j=1}^{m}\frac{1}{\theta}\int_{T_i}^{T_i+\theta}(p(t, x, z_1) + q(x, z_1))dz_1$$

$$\times[(C(CQ)^*)^{-1}]_{i,j}\frac{1}{\theta}\int_{T_j}^{T_j+\theta}(p(t, z_2, y) + q(z_2, y))dz_2.$$

4.6 Parameter Estimation Problem

Our objective now is to estimate the unknown system parameters. Our first difficulty is the covariance kernel $q(x, y)$. If we can parameterize it with one or more parameter(s), say c, then the parameters we need to estimate are $\kappa, \sigma_\chi, \sigma_\xi, \sigma, \lambda_\chi, \lambda_\xi, c$ and the seasonality function $h(t)$. The standard approach is to use the method of maximum likelihood, for which we need to calculate the likelihood functional from the observation data $\{Y(t); 0 \le t \le t_f\}$, where t_f denotes a final time. However, since the observation noise covariance

$$\Phi = [\sigma_\chi^2 H(\tau_i)H(\tau_j) + \rho\sigma_\chi\sigma_\xi(H(\tau_i) + H(\tau_j)) + \sigma_\xi^2$$

$$+ \frac{\sigma^2}{\theta^2}\int_{\tau_i}^{\theta+\tau_i}\int_{\tau_j}^{\theta+\tau_j}q(x, z)dxdz]_{ij} \qquad (4.32)$$

is unknown, we do not have an obvious likelihood functional. To understand this problem, note that in our discussion in Section 2.3.4, we write down the likelihood on the assumption that the covariance of the observation noise is known. However, in our case here, the observation noise covariance, see equation (4.32) contains unknown parameters which need to be estimated. In this case, there is no apparent fixed measure on $\mathbf{C} = C([0, T]; R^m)$ such that the measure induced by $\{Y(t); t \le t \le t_f\}$ on $\mathbf{C}$ is absolutely continuous with respect to that fixed measure. However, since our model is linear and Gaussian, we may circumvent this problem by working with a quasi likelihood functional as proposed in Bagchi (1975). This is done by defining an appropriate functional and showing that minimizing this functional with respect to all the unknown parameters, including those in the

observation noise covariance, yields consistent estimates of those parameters. We can, in fact, replace the observation noise covariance appearing as weight in the usual likelihood functional expression by any positive definite matrix which we take here to be the identity matrix I. The noise covariance matrix does appear in the quasi likelihood functional through the filtered states and may be estimated from there. The quasi likelihood functional for our problem is

$$QL(t_f, Y, I) = \int_0^{t_f} (\mathcal{H} \begin{bmatrix} \hat{\chi}(s) \\ \hat{f}_w(s) \end{bmatrix} + \hat{G})^* dY(s)$$
$$- \frac{1}{2} \int_0^{t_f} ||(\mathcal{H} \begin{bmatrix} \hat{\chi}(s) \\ \hat{f}_w(s) \end{bmatrix} + \hat{G})||^2 ds, \qquad (4.33)$$

where $\hat{x}(s)$ and $\hat{f}_w(s)$ are the "best" estimates of the states $x(s)$ and $f_w(s)$ given the observation data $\sigma\{Y(\tau); 0 \leq \tau \leq s\}$,

$$\mathcal{H} = [-\kappa H(\tau_i), \frac{1}{\theta} \int_0^{T^*} \{\delta(x - \theta - \tau_i) - \delta(x - \tau_i)\}(\cdot)dx]_i \qquad (4.34)$$

and

$$\hat{G} = [\frac{\partial \hat{A}(t, \tau_i)}{\partial t} - \lambda_\chi H(\tau_i)]_i$$
$$= [(h(\tau_i + \theta + t) - h(\tau_i + t))/\theta + \lambda_\xi - \lambda_\chi H(\tau_i)]_i \qquad (4.35)$$

The MLE of the unknown parameters are then given by

$$\begin{bmatrix} \hat{\kappa} & \hat{\sigma}_\chi & \hat{\sigma}_\xi & \hat{\sigma} & \hat{\lambda}_\chi & \hat{\lambda}_\xi & \hat{\omega}_S \end{bmatrix} = \arg\max \ QL(t_f, Y, I)$$

where we set the function form of $Q = \int_0^{\hat{T}} \sigma^2 q(x, y)(\cdot)dy$ and the seasonality function $h(t) = h(\omega_S; t)$ for an unknown periodic factor ω_S.

4.7 State Estimation Problem

Now we summarize the system and observation mechanism in the usual vector notation;

$$d \begin{bmatrix} \chi(t) \\ f_w(t, x) \end{bmatrix} = \begin{bmatrix} -\kappa \chi(t) - \lambda_\chi \\ \frac{\partial f_w(t, x)}{\partial x} \end{bmatrix} dt + \begin{bmatrix} \sigma_\chi dW_\chi^*(t) \\ d\tilde{w}(t, x) \end{bmatrix},$$

where $d\tilde{w}(t, x) = \sigma dw^*(t, x) + \sigma_\xi dW_\xi^*(t)$ and

$$dY(t) = \mathcal{H} \left[\begin{array}{c} \chi(t) \\ f_w(t, \cdot) \end{array} \right] dt + \hat{G} dt + \sigma_\chi \bar{H} dW_\chi^*(t) + K d\tilde{w}(t, \cdot),$$

where $\mathcal{H}$ is defined by (4.34),

$$\bar{H} = [H(\tau_i)]_{m \times 1}, \text{ and } K(\cdot) = [\frac{1}{\theta} \int_{\tau_i}^{\theta + \tau_i} (\cdot) dz]_{m \times 1}.$$

As is mentioned in Chapter 2, the observation covariance Φ is positive definite and can be obtained numerically. (The details will be stated in the next section). Hence from Chapter 2, we can derive the optimal filter equation, see also Kallianpur (1980).The optimal estimates for $x(t)$ and $f_w(t, x)$ are given by

$$d \left[\begin{array}{c} \hat{\chi}(t) \\ \hat{f}_w(t, x) \end{array} \right] = \left[\begin{array}{c} -\kappa \hat{\chi}(t) - \lambda_\chi \\ \frac{\partial \hat{f}_w(t,x)}{\partial x} \end{array} \right] dt$$

$$+ \left(\mathcal{P}(t) \mathcal{H}^* + \left[\begin{array}{c} \sigma_\chi^2 \bar{H}^* + \rho \sigma_\chi \sigma_\xi \mathbf{1}_m^* \\ \rho \sigma_\chi \sigma_\xi \bar{H}^* + \sigma_\xi^2 \mathbf{1}_m^* + \sigma^2 Q K^* \end{array} \right] \right) \Phi^{-1}$$

$$\times (dY(t) - \mathcal{H} \left[\begin{array}{c} \hat{\chi}(t) \\ \hat{f}_w(t, \cdot) \end{array} \right] dt - \hat{G} dt), \qquad (4.36)$$

where $\mathbf{1}_m^* = [1, 1, \cdots, 1]$,

$$QK^* = [\frac{1}{\theta} \int_{\tau_1}^{\theta + \tau_1} \sigma^2 q(x, y) dy, \cdots, \frac{1}{\theta} \int_{\tau_m}^{\theta + \tau_m} q(x, y) dy],$$

$$\mathcal{P}(t) = \left(\begin{array}{cc} \mathbf{P}_x(t) & \mathbf{P}_{xw} \\ \mathbf{P}_{wx}(t) & \mathbf{P}_w \end{array} \right),$$

$\mathcal{P} = \mathcal{P}^*$ and

$$\mathbf{P}_x(t) = p_x(t), \quad \mathbf{P}_{xw}(t) = p_{xw}(t, x),$$

$$\mathbf{P}_w(t) = \int_0^{\hat{T}} p_w(t, x, y)(\cdot) dy,$$

and

$$\mathcal{P} \mathcal{H}^* = \left(\begin{array}{c} p_x(t)(-\kappa \bar{H}) + [\frac{1}{\theta}(p_{xw}(t, \tau_i + \theta) - p_{wx}(t, \tau_i)]_{1 \times m} \\ p_{xw}(t, x)(-\kappa \bar{H}) + [\frac{1}{\theta}(p_w(t, x, \tau_i + \theta) - p_w(t, x, \tau_i)]_{1 \times m} \end{array} \right).$$

The kernel equations are given by

$$
\frac{dp_x(t)}{dt} = -2\kappa p_x(t) + \sigma_\chi^2 - \left[-\kappa p_x(t)\bar{H}^* \right]
$$

$$
+ \frac{1}{\theta}\left[p_{xw}(t, \theta + \tau_i) - p_{xw}(t, \tau_i) \right]^* + \sigma_\chi^2 \bar{H}^* + \rho\sigma_\chi\sigma_\xi \mathbf{1}_m^*
$$

$$
\times \Phi^{-1}\left[-\kappa p_x(t)\bar{H} + \frac{1}{\theta}[p_{xw}(t, \theta + \tau_i) - p_{xw}(t, \tau_i)] \right]
$$

$$
+ \sigma_\chi^2 \bar{H} + \rho\sigma_\chi\sigma_\xi \mathbf{1}_m, \quad p_x(0) = Cov\{x(0)\}, \tag{4.37}
$$

$$
\frac{\partial p_{xw}(t, x)}{\partial t} = -\kappa p_{xw}(t, x) + \frac{\partial p_{xw}(t, x)}{\partial x} + \rho\sigma_\chi\sigma_\xi
$$

$$
- \left[-\kappa p_x(t)\bar{H}^* + \frac{1}{\theta}[p_{xw}(t, \theta + \tau_i) - p_{xw}(t, \tau_i)]^* + \sigma_\chi^2 \bar{H}^* \right.
$$

$$
+ \rho\sigma_\chi\sigma_\xi \mathbf{1}_m^* \Big] \Phi^{-1}[-\kappa p_{xw}(t, x)\bar{H} + \frac{1}{\theta}[p_w(t, x, \theta + \tau_i)
$$

$$
- p_w(t, x, \tau_i)] + [\rho\sigma_\chi\sigma_\xi \bar{H} + \sigma_\xi^2 \mathbf{1}_m + \frac{\sigma^2}{\theta}\int_{\tau_i}^{\theta + \tau_i} q(x, y)dy]],
$$

$$
\frac{\partial p_w(t, x, y)}{\partial t} = \frac{\partial p_w(t, x, y)}{\partial x} + \frac{\partial p_w(t, x, y)}{\partial y} + \sigma^2 q(x, y)
$$

$$
+ \sigma_\xi^2 - \left[-\kappa p_{xw}(t, x)\bar{H}^* + \frac{1}{\theta}[p_w(t, x, \theta + \tau_i) \right.
$$

$$
-p_w(t, x, \tau_i)]^* + \rho\sigma_\chi\sigma_\xi \bar{H}^* + \sigma_\xi^2 \mathbf{1}_m^* + \frac{\sigma^2}{\theta}\int_{\tau_i}^{\theta + \tau_i} q(x, y)dy \Big]
$$

$$
\times \Phi^{-1}\left[-\kappa p_{xw}(t, y)\bar{H} + \frac{1}{\theta}[p_w(t, \theta + \tau_i, y) - p_w(t, \tau_i, y)]^* \right.
$$

$$
+\rho\sigma_\chi\sigma_\xi \bar{H} + \sigma_\xi^2 \mathbf{1}_m + \frac{\sigma^2}{\theta}\int_{\tau_i}^{\theta + \tau_i} q(x, y)dx \Big], \tag{4.38}
$$

with $p_{xw}(0, x) = p_w(0, x, y) = 0$.

Hence we obtain

$$
\hat{f}(t, x) = E\{f(t, x)|\mathcal{Y}_t\}
$$

$$
= \hat{\chi}(t)B(x) + \tilde{A}(t, x) + \hat{f}_w(t, x). \tag{4.39}
$$

4.8 Simulation Studies

4.8.1 Numerical analysis of observation covariance from real data

Here we first check the invertibility of the covariance of the real observation data. We used a historical time-series of UK-Gas-NBP spot prices quoted daily from 2-Jan-2007 to 28-Dec-2008. From Chapter 2, we calculate the covariance of $Y(t)$ numerically such that

$$\frac{1}{n}\sum_{i=1}^{n}(Y(t_{i+1}) - Y(t_i))'(Y(t_{i+1}) - Y(t_i)) \sim \Phi. \qquad (4.40)$$

In Figure 4.1, the used data is presented.

Figure 4.1: Historical real data of UK-gas-NBP

From (4.40), we get the covariance Φ shown in Figure 4.2.

Figure 4.2: Numerically obtained Φ

Now we will calculate the inverse of Φ and at the same time we check this invertibility to calculate $\Phi\Phi^{-1}$ numerically. These results are demonstrated in Figure 4.3 and Figure 4.4, respectively.

Figure 4.3: Numerically obtained Φ^{-1}

Figure 4.4: Numerically obtained $\Phi\Phi^{-1}$

From these results, we find that it is possible to realize the maximization of quasi likelihood shown in Section 4.6 through the Kalman filter given in Section 4.7. As a value for the noise covariance Φ, we use the numerically obtained value of (4.40). Furthermore, we can construct the proxy function for Φ as used in the work of Aihara and Bagchi (2010b). We set the function form of $q(x, y)$ such that

$$q(x,y) = \sum_{i=1}^{100} \exp(-cx/T^*) \sin(i\pi x/T^*) \exp(-cy/T^*) \sin(i\pi y/T^*)/i^2. \quad (4.41)$$

Hence from (4.32), we can choose $\sigma_\chi, \sigma_\xi, \rho, \sigma$ and c to fit the diagonal shape of Φ. The usual mean square method dose not work properly. For details, please consult Aihara and Bagchi (2010b). Here by using trial and error procedure, we consider the maximum of the delivery horizon to be $T^* = 5$

$$\sigma = 0.5, \kappa = 1.421, \sigma_\chi = 0.07, \sigma_\xi = 0.01, \rho = 0.3, c = 3.$$

The obtained shape is shown in Figure 4.5.

Figure 4.5: Numerically obtained $\Phi(i,i)$ and its proxy shape

4.8.2 Simulation studies

We set the system parameters as follows:

$$\kappa = 1.321, \ \lambda_\chi = 0.623, \ \sigma_\chi = 1.2,$$
$$\lambda_\xi = 0.04, \ \sigma_\xi = 1.2, \ \rho = 0.6, \ \sigma = 1.$$

The seasonality function is set as

$$h(t) = 14.2521 + 4.0052t + h_p(t)$$

where we choose the seasonality function that represents UK-NBP gas market, derived in Subsection 1.1.4. The initial conditions for χ, ξ are set as

$$\chi(0) = 0.8, \, \xi(0) = 20.$$

We assume that the covariance kernel of $\sigma w(t,x)$ is given by

$$q(x,y) = \sum_{k=1}^{100} \sin(k\pi x/5)\sin(k\pi y/5),$$

where $q(x, y) \sim \delta(x-y)$, where we check the feasibility of the proposed algorithm in the severe noise kernel case for simulation studies.

The simulated observation data for a fixed delivery period of $T - T_0 = 1$ month, is shown in Figure 4.6. The factor process $f(t, x)$ is also demonstrated in Figure 4.7.

Figure 4.6: Observation data

Figure 4.7: $f(t, x)$-process

4.8.3 MLE results

We assume that the unknown parameters are $\kappa, \lambda_\chi, \sigma_\chi, \lambda_\xi, \sigma_\xi, \rho$ and σ. For finding MLE, we used the GA-algorithm in MATLAB. The initial values are set as

$$\kappa = 1.5, \ \lambda_\chi = 0.5, \ \sigma_\chi = 0.15, \ \lambda_\xi = 0.05,$$
$$\sigma_\xi = 0.1, \ \rho = 0.5, \ \sigma = 0.5$$

with the upper and lower bounds given by

$$1 \leq \kappa \leq 2, \ 0.1 \leq \lambda_\chi \leq 1.0, \ 0.1 \leq \sigma_\chi \leq 0.2, \ 0.01 \leq \lambda_\xi \leq 0.1,$$
$$0.05 \leq \sigma_\xi \leq 0.3, \ 0.1 \leq \rho \leq 1.0, \ 0.1 \leq \sigma \leq 2$$

We demonstrate the running time procedure of GA in Figure 4.8.

Figure 4.8: Evolutions of quasi-likelihood

$\hat{\kappa}$	$\hat{\lambda}_\chi$	$\hat{\sigma}_\chi$	$\hat{\lambda}_\xi$	$\hat{\sigma}_\xi$	$\hat{\rho}$	$\hat{\sigma}$
1.3757	0.9131	0.1292	0.0481	0.0772	0.3528	1.8840

Table 4.1: Estimated parameters

The estimates of the parameters are listed in Table 4.1. The estimated $\log \tilde{F}(t, x)$ with MLE parameters is shown in Figure 4.9, and the true and estimated spot and $\log \tilde{F}(t, 1.2698)$ processes are also shown in Figure 4.10 and Figure 4.11.

Figure 4.9: Estimated $f(t, x)$ process

Figure 4.10: True and estimated $S(t)$-processes

Figure 4.11: True and estimated $f(t, 1.2698)$-processes

4.8.4 Real data analysis

The data we consider, is the data of the UK-NBP Gas prices. For that, we use the estimated parameters for $h_p(t)$ calculated in Chapter 1. The real observation data is presented in Figure 4.12.

Figure 4.12: Real observed data

For the covariance of the observation noise Φ, we use the numerically calculated value shown in Figure 4.2 of subsection 4.8.1. The kernel form of $q(x, y)$ is selected as in (4.41) with $c = 3$. The unknown parameters are $\kappa, \lambda_\chi, \sigma_\chi, \lambda_\xi, \sigma_\xi, \rho, \sigma, \chi(0), \xi(0)$, and λ_q, where λ_q is the market price of risk included in the initial condition of $f(0, x)$ given by

$$f(0, x) = \frac{1}{2}\sigma^2 \lambda_q \int_0^x q(z, z; c)dz + \xi(0) + \frac{\lambda_\kappa}{\kappa}(e^{-\kappa x} - 1) + \lambda_\xi x$$

$$+ h(x) + \frac{\sigma_\chi}{\kappa}(1 - e^{-2\kappa x}) + \frac{1}{2}\sigma_\xi^2 x + e^{-\kappa x}\chi(0) + \rho\frac{\sigma_\chi \sigma_\xi}{\kappa}(1 - e^{-\kappa x}).$$

See Remark 3 of Section 4.4. The upper and lower bounds of these parameters are set as

$$\begin{array}{ccc}
0.1 \le \kappa \le 2, & 0.1 \le \lambda_\chi \le 21.0, & 0.01 \le \sigma_\chi \le 0.2, \\
0.01 \le \lambda_\xi \le 4, & 0.01 \le \sigma_\xi \le 3.0, & 0.01 \le \rho \le 1.0, \\
0.001 \le \sigma \le 5, & 0.1 \le \chi(0) \le 2.0, & 10 \le \xi(0) \le 50.0, \\
& -2.0 \le \lambda_q \le 2 &
\end{array}$$

The running time procedure of GA in Figure 4.13 and the estimates of the parameters are listed in Table 4.2.

Figure 4.13: Evolutions of quasi-likelihood for real data

$\hat{\kappa}$	0.9365
$\hat{\lambda}_\chi$	0.2006
$\hat{\sigma}_\chi$	0.0298
$\hat{\lambda}_\xi$	3.9216
$\hat{\sigma}_\xi$	1.1593
$\hat{\rho}$	0.0767
$\hat{\sigma}$	0.0228
$\hat{\chi}(0)$	0.2939
$\hat{\xi}(0)$	12.006
$\hat{\lambda}_q$	1.5582

Table 4.2: Estimated parameters

The estimated $f(t, x)$ process is shown in Figure 4.14. Also, Figure 4.15 and Figure 4.16 demonstrate the estimate of $f(t, x)$ when the time to maturity x is $\frac{50}{126}$ and 2 years, respectively. The case of $x = 0$ which corresponds to the estimate of the spot price, is also shown in Figure 4.17.

Figure 4.14: Estimated $f(t, x)$ process

4

Figure 4.15: Estimated $f(t, 50/126)$-processes

Figure 4.16: Estimated $f(t, 2)$-processes

Figure 4.17: Estimated $f(t, 0)$-process

4.9 Concluding Remarks

In this chapter, we propose a new arbitrage free model for the futures prices of energy. The new model can be used in a mathematically sound way when estimating the parameters of the model using the method of maximum likelihood. Using reverse engineering type modeling, we start by assuming that the term structure of futures prices on electricity given by Schwartz and Smith (2000) model is affected by an error term represented by a stochastic integral that generates infinite dimensional noise as it should depend on all time of, or to maturity. Hence, we do not need to add artificial noises to the observation equation in order to use the filtering methodology. We extend the model taking into account the delivery period in the futures prices by employing the geometric average approximation in the payoff of the futures. We then show the observation mechanism and formulate the state space representation of the problem. The factors then are estimated as solutions of the resulting filtering problem. We discuss the difficulty in relation to the noise covariance of the measurement equation and how to remediate it to get a quasi likelihood functional

that can be used to estimate the unknown parameters using simulation and real data of the spot and futures on the UK-Gas-NBP, the feasibility of the proposed filter is established.

Chapter 5

Convolution Particle Filter Applied to New Energy Spot Model

5.1 Introduction

As we have discussed in Chapter 4, the formula of the futures price on electricity is highly nonlinear and contains an infinite dimensional noise term, see equation (4.12). To overcome this problem we use the geometric average approximation, so that the logarithm of the futures price becomes linear with respect to the unobservable factors, see equation (4.14). This enabled us to use the infinite dimensional Kalman filter methodology together with the maximum likelihood estimation method to estimate the parameters of the state space model. The aim of this chapter is to keep the futures price as in equation (4.12), and look for the parameter estimation problem without using the geometric approximation. As the observation given by the futures price is highly non-linear, we propose to use a variant of the particle filter for the identification of the factors and the parameters of the system. This filter is based on the convolution kernel approximation techniques and termed convolution particle filter, see Rossi and Vila (2006).

In Section 5.2, we briefly repeat our discussion of chapter 4 and present the forward model together with the mechanism of the ob-

servation equation. Furthermore, we present our model in a form ready to be used in this framework. This will be followed by simulation work and studies using real data, which will be shown in Section 5.5. The final section will present the conclusion.

5.2 Practical Model for the Electricity Prices

In this section we briefly repeat our discussion of chapter 4. The market prices of electricity futures are different from the standard futures traded in other markets. The electricity futures prices are based on the arithmetic averages of the spot prices over a delivery period $[T_0, T]$, given by

$$\frac{1}{T - T_0} \int_{T_0}^{T} S(\eta) d\eta. \tag{5.1}$$

Now, for $t < T$, we can calculate the futures price by

$$F(t, T_0, T) = \mathbf{E}\{\frac{1}{T - T_0} \int_{T_0}^{T} S(\eta) d\eta | \mathcal{F}_t\}, \tag{5.2}$$

where $\mathcal{F}_t = \sigma\{S(\eta); 0 \leq \eta \leq t\}$. Assuming that $S(t) \in L^2(T_0, T)$, and using the linearity of the expectation operator, (5.2) can be written as

$$F(t, T_0, T) = \frac{1}{T - T_0} \int_{T_0}^{T} \mathbf{E}\{S(\eta) | \mathcal{F}_t\} d\eta, \tag{5.3}$$

Using the definition of futures price, it can be simplified as

$$F(t, T_0, T) = \frac{1}{T - T_0} \int_{T_0}^{T} F(t, \eta) d\eta, \tag{5.4}$$

This price using (4.10) satisfies

$$
\begin{aligned}
F\left(t, T_0, T\right) \;=\; & \tfrac{1}{T-T_0} \int\limits_{T_0}^{T} \exp\Big[B\left(\eta - t\right) \chi\left(t\right) + \tilde{A}\left(t, \eta - t\right) \\
& + \int\limits_{0}^{t} \{\sigma dw^*(s, \eta - s) + \sigma_\xi dW_\xi^*(s)\} \Big]\, d\eta, \qquad (5.5)
\end{aligned}
$$

where B and $\tilde{A}$ satisfy the following equations, respectively:

$$
\begin{aligned}
B\left(T - t\right) &= e^{-\kappa(T-t)} & (5.6)\\
\tilde{A}(t, T - t) &= \bar{A}(t, T - t) + \lambda_\xi t + \xi(0). & (5.7)
\end{aligned}
$$

where

$$
\bar{A}\left(t, T - t\right) = A\left(t, T - t\right) + \frac{1}{2}\sigma^2 \int\limits_{0}^{T-t} q\left(x, x\right) dx
$$

and where $A\left(t, T - t\right)$ is given by (4.5).

5.2.1 A forward model

In (5.5), setting $x = \eta - t$ to consider the time to maturity notation instead of the time of maturity. Let $f(t, x)$ represents the exponential part of the integrand given by

$$
\begin{aligned}
f(t, x) = B\left(x\right) \chi\left(t\right) + \tilde{A}\left(t, x\right) \\
+ \int\limits_{0}^{t} \{\sigma dw^*(s, x + t - s) + \sigma_\xi dW_\xi^*(s)\} \qquad (5.8)
\end{aligned}
$$

we get

$$
df(t, x) = \frac{\partial f(t, x)}{\partial x} dt - \frac{1}{2}\tilde{q}(x, x)dt + d\tilde{w}(t, x), \qquad (5.9)
$$

where

$$
\tilde{w}(t, x) = \sigma w^*(t, x) + e^{-\kappa x}\sigma_\chi W_\chi^*(t) + \sigma_\xi W_\xi^*(t), \qquad (5.10)
$$

and

$$\tilde{q}(x_1, x_2) = \sigma^2 q(x_1, x_2) + \frac{\rho \sigma_\chi \sigma_\xi}{2} (e^{-\kappa x_1} + e^{-\kappa x_2})$$
$$+ \sigma_\chi^2 e^{-\kappa(x_1 + x_2)} + \sigma_\xi^2. \qquad (5.11)$$

Hence the futures price in (5.5) becomes

$$F(t, T_0, T) = \frac{1}{T - T_0} \int_{T_0}^{T} \exp\left[f(t, x)\right] dx.$$

Notice that the identification usually is performed under the real world measure. This implies that the market price of risk terms are included in (5.9). Here, we neglect these terms because our identification procedure is easily applied to the model under the real world measure. See remark 3 in Chapter 4 for further discussion about this point.

5.2.2 Observation mechanism (arithmetic average case)

In chapter 4, we used the geometric average approximation to derive the filtering algorithm. In this chapter, we reformulate this observation mechanism without using any artificial approximation method. In practice, the futures are available on a daily basis and transformed as the time-to delivery $\tau_i = T_0^i - t$ is fixed as a constant time period τ_i for $i = 1, 2, \cdots, m$, where m is the number of futures (observations). Moreover, the delivery period $\theta = T - T_0$ is also fixed to be 1-month for all the futures. Hence the usual observation data becomes

$$
\begin{aligned}
y_i(t) &= \log F(t, \tau_i + t, \tau_i + t + \theta) \\
&= \log \frac{1}{\theta} \int_{\tau_i}^{\tau_i + \theta} \exp\{f(t, x)\} dx, \\
&\quad \text{for } \tau_1 < \tau_2 < \cdots < \tau_m. \qquad (5.12)
\end{aligned}
$$

Also, we set the observation vector as

$$\vec{Y}(t) = [y_1(t), y_2(t), \cdots, y_m(t)]'$$

where it should be noted that the observation contains the nonlinear state for $f(t, x)$.

5.3   Discrete Approximation for
System and Observation

In this section, we are interesting in estimating the parameters of the
new nonlinear state space model given by equation (5.9) for the state
and equation (5.12) for the observation. In this case, the parameter
estimation procedure is often based on an approximation of the op-
timal nonlinear filter using the extended Kalman filter, coupled with
maximum likelihood estimation techniques. Another approach is to
employ the Bayesian framework where the augmented state variable
$(f(t, x), \Theta)$ which represents the state and the unknown parameters,
is processed by a filtering procedure. The main issue regarding the
use of the extended Kalman filter and its various alternatives, is that
they do not always give good results. On the other hand, particle fil-
ters propose a good alternative in which Θ is considered as a random
variable with a prescribed a priori density function. Then the usual
augmented state approach can be considered. Such an identification
problem can be formulated as the problem of estimating the initial
state of a known, but albeit more complex system, see Balakrishnan
(1969). To show this point, consider we have the following dynamical
system:

$$
\begin{aligned}
X(t) &= F(X(t); \Theta) & (5.13) \\
\nu(t) &= C\,X(t) + n(t) & (5.14)
\end{aligned}
$$

where $X(t)$ is the state and $\nu(t)$ is the observation. If we form a new
vector $Z(t) = [X, \theta(t)]$ where

$$
\begin{aligned}
\dot{\theta}(t) &= 0 & (5.15) \\
\theta(0) &= \Theta & (5.16)
\end{aligned}
$$

Then we can write (5.13) in the form:

$$
\begin{aligned}
\dot{Z} &= \widetilde{F}[Z] & (5.17) \\
\nu(t) &= H(Z) + n(t) & (5.18)
\end{aligned}
$$

and the problem of estimating $X(t)$ and Θ simultaneously is the same
as that of determining $Z(t)$ from $\nu(s)$ where $0 < s < t$. In the follow-
ing section, we show how this system can be represented when using

the convolution particle filter. For that, we need to transform our state equation from stochastic partial differential equation to stochastic partial difference equation.

In the following, we present the approximation method to convert the infinite-dimensional system (5.9) to the finite-dimensional one.

The spatial region $]0, T^*[$ is discretized as

$$0 = x_1 < x_2 < \cdots < x_i < \cdots < x_k = T^*,$$

where we set for all i

$$\Delta x = x_{i+1} - x_i.$$

Now we approximate the partial differential operator $\frac{\partial(\cdot)}{\partial x}$ as the finite-dimensional matrix A;

$$A = \frac{1}{\Delta x} \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}_{k \times k}. \tag{5.19}$$

We also discretize the time region $[0, t_f]$ as

$$0 = t_0 < t_1 < t_2 < \cdots < t_i < \cdots < t_\ell = t_f,$$

where $\Delta t = t_{i+1} - t_i$ for all i. Hence our system state $f(t, x)$ is projected into the finite points $\{t_i, x_j\}$ for $i = 1, 2, \cdots, \ell, j = 1, 2, \cdots, k$, i.e., denoting $f_{ij} = f(t_i, x_j)$, we construct the finite dimensional vector;

$$\vec{f_i} = [f_{i1}, f_{i2}, \cdots, f_{ik}]'. \tag{5.20}$$

So the partial time derivative of f is approximated as

$$\frac{\partial f(t, x)}{\partial t} \sim \frac{\vec{f}_{i+1} - \vec{f_i}}{\Delta t}.$$

The deterministic input $\frac{1}{2}\tilde{q}(x, x)$ is also transformed to

$$\frac{1}{2}\vec{q} = \frac{1}{2}[q(x_1, x_1), q(x_2, x_2), \cdots, q(x_k, x_k)]'.$$

Hence the finite-dimensional version of the system (5.9) becomes

$$
\frac{\vec{f}_{i+1} - \vec{f}_i}{\Delta t} = A\vec{f}_i - \frac{1}{2}\vec{q} + \underbrace{(\frac{\partial \tilde{w}(t,x)}{\partial t})}_{\text{noise term}}. \tag{5.21}
$$

Next we shall show how to generate the noise term $\frac{\partial \tilde{w}(t,x)}{\partial t}$ in a discrete time and spatial space. We list up the detail steps

- $w(t,x)$ is approximated as

$$
w^*(t,x) \sim \sum_{i=1}^{100} e^{-cx/T^*} \sin(i\pi x/T^*)\frac{w_i(t)}{i}
$$

  where $w_i(t), i = 1, 2, \cdots$ are mutually independent standard Brownian motion processes.

- Consider the finite points $\{x_j\}$ for $j = 1, 2, \cdots, k$,i.e.,

$$
\tilde{w}(t,x) \sim [\tilde{w}(t,x_1), \cdots, \tilde{w}(t,x_\ell), \cdots, \tilde{w}(t,x_k)]',
$$

  where

$$
\tilde{w}(t,x_\ell) = \sum_{i=1}^{100} e^{-cx_\ell/T^*} \sin(i\pi x_\ell/T^*)\frac{w_i(t)}{i} + e^{-\kappa x_\ell}\sigma_\chi W_\chi^*(t) + \sigma_\xi W_\xi^*(t).
$$

- At each time t_p, we generate mutually independent normal random numbers $N_{pi}, i = 1, 2, \cdots, 102$. Then our Brownian motion $w_i(t)$ becomes

$$
w_i(t_p) \sim N_{pi}\sqrt{\Delta t}, \text{for } i = 1, 2, \cdots, 100
$$

  and

$$
\begin{cases}
W_\chi^*(t_p) \sim (N_{p101}\rho + N_{p102}\sqrt{1-\rho^2})\sqrt{\Delta t} \\
W_\xi^*(t_p) \sim N_{p101}\sqrt{\Delta t}
\end{cases}
$$

- Finally we get

$$\frac{\partial \tilde{w}(t_p, x_\ell)}{\partial t} \sim \left(\sum_{i=1}^{100} e^{-cx_\ell/T^*} \sin(i\pi x_\ell/T^*) \frac{N_{pi}\sqrt{\Delta t}}{i} \right.$$
$$\left. + e^{-\kappa x_\ell} \sigma_\chi (N_{p101}\rho + N_{p102}\sqrt{1-\rho^2})\sqrt{\Delta t} + \sigma_\xi N_{p101}\sqrt{\Delta t} \right) \frac{1}{\Delta t}.$$

We denote the vector version of above quantity by

$$\Delta \vec{W}_p \frac{1}{\Delta t}.$$

Now we obtain the discrete version of (5.9):

$$\vec{f}_{i+1} = (1 + A\Delta t)\vec{f}_i - \frac{1}{2}\vec{q}\Delta t + \Delta \vec{W}_i, \qquad (5.22)$$

where it should be noted that in order to support the discretely approximated system converges to the original one as $\Delta t, \Delta x \to 0$, we need the following condition:

$$|\frac{\Delta t}{\Delta x}| < 1.$$

Here noting that the spatial variable x is a time-to-maturity variable, we set

$$\Delta x = 2\Delta t$$

to fit the discrete time points to the discrete time-to-maturity points. We also mentioned about the special property for the 1st order hyperbolic equation. As the time goes by, the spatial region is shrinking without a boundary condition on $x = T^*$. So we set the artificial boundary condition $f_{i+1,k} = f_{i,k}$, because in our situation T^* is set to be longer than t_f. Summarizing above procedure, we denote the discretized system as

$$\vec{f}_{i+1} = F(\vec{f}_i, \Delta \vec{W}_i, \Theta) \qquad (5.23)$$

where Θ denotes the parameters which we need to identify ,i .e.,

$$\Theta = [\sigma, c, \xi(0), \lambda_\chi, \kappa, \lambda_\xi, \sigma_\chi, \sigma_\xi, \rho].$$

Now the observation $\vec{Y}(t)$ given by (5.12) is also generated from $\vec{f}_i$, i.e.

$$\vec{Y}(t_i) = [y_\ell(t_i)]_{m \times 1},$$

where

$$
\begin{aligned}
y_\ell(t_i) &= \log \frac{1}{\theta} \int_{\tau_i}^{\tau_i+\theta} \exp(f(t_i, x)) dx \\
&\sim \log \left(\frac{1}{\theta} \sum_{j=[\tau_i]}^{[\tau_i+\theta]} \exp(f_{i,j}) \Delta x \right)
\end{aligned}
$$

and $[\tau]$ denotes the index of τ in the spatial space x. The discretized observation $\vec{Y}_i = \vec{Y}(t_i)$ is symbolically denoted by

$$\vec{Y}_i = h(\vec{f}_i). \tag{5.24}$$

5.4 Convolution Particle Filter

In general, the classical filters such as sample importance resampling (SIR), and regularized particle filter (RPF), can handle the augmented state vector if a dynamic noise term is artificially added to the parameters. But, the drawback of this, is the reduction of the estimation performance of the filter. However, another promising approach which avoids adding extra artificial noises to the parameters is the particle convolution filter approach proposed by Rossi and Vila (2006). This approach is based on kernel estimation techniques and it is free of the analytical knowledge of both the state and observation variable distributions. Only the capability of simulating the state and observation noises is required. Moreover, it can handle the problem of small magnitude observation noise which is typical in financial data. To apply the convolution filter of Campillo and Rossi (2006), we suppose that the parameter Θ is a random variable with a given prior law $P_\Theta(\Theta)$. Now our system given in the previous section becomes

$$
\begin{cases}
\Theta_{i+1} = \Theta_i \\
\vec{f}_{i+1} = F(\vec{f}_i, \Delta \vec{W}_i, \Theta_{i+1}) \\
\tilde{\vec{Y}}_{i+1} = h(\vec{f}_{i+1}) - h(\vec{f}_i)
\end{cases}
\tag{5.25}
$$

where our observation mechanism is reset as $\tilde{\vec{Y}}_{i+1}$, because in the continuous case we always use the observation data as the differential form $d\vec{Y}(t)$. Furthermore, noting that the original initial condition $f(0, x)$ is given by

$$f(0, x) = \frac{1}{2}\sigma^2 \int_0^x q(z, z; c)dz + \xi(0) + \frac{\lambda_\kappa}{\kappa}(e^{-\kappa x} - 1) + \lambda_\xi x$$

$$+ h(x) + \frac{\sigma_\chi}{\kappa}(1 - e^{-2\kappa x}) + \frac{1}{2}\sigma_\xi^2 x + e^{-\kappa x}\chi(0)$$

$$+ \rho\frac{\sigma_\chi \sigma_\xi}{\kappa}(1 - e^{-\kappa x}), \quad (5.26)$$

the discrete version becomes

$$\vec{f}_0(\Theta) = \text{function of } \Theta. \quad (5.27)$$

The essence of the convolution particle filter is to combine the simulated observation data with the real observed data by using the Parzen-Rosenblatt kernel K. This kernel $K : \mathbb{R}^d \mapsto \mathbb{R}$ is a bounded positive symmetric function such that $\int K(x)\,dx = 1$. For example, the Gaussian kernel is $K(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^d e^{-\frac{\|x\|^2}{2}}$. The Parzen-Rosenblatt kernel generally satisfies $\|x\|^d K(x) \to 0$ as $\|x\| \to \infty$.

In this thesis, we only use this to construct the likelihood functional which is used to construct the particle filter. Now our convolution particle filter algorithm is:

- We assume that $P_\Theta(\Theta)$ is set as the uniformly distributed probability in some bounded set. Then we generate N i.i.d. particles $\Theta^{(1)}, \Theta^{(2)}, \cdots, \Theta^{(N)}$.

- The initial particles of $\vec{f}_0^{(i)}$ for $i = 1, 2, \cdots, N$ are automatically generated from (5.27).

- From (5.25) we can generate the sequence of $\vec{f}_{i+1}^{(j)}, \Theta_{i+1}^{(j)}, \tilde{\vec{Y}}_{i+1}^{(j)}$ for $j = 1, 2, \cdots, N$.

- At each time step $t = t_{i+1}$, we get the real observation data such that

$$\tilde{\mathbf{Y}}_{i+1} = \text{real data of } \vec{Y}(t_{i+1}) - \vec{Y}(t_i). \quad (5.28)$$

- Construct the likelihood by using a Parzen-Rosenblatt kernel for $j = 1, 2, \cdots, N$

$$\hat{p}^{(j)}(\vec{Y}(t_{i+1})|\vec{f}_i^{(j)}, \vec{Y}(t_i)) = \frac{1}{Nh^q}K_h(\tilde{\mathbf{Y}}_{i+1} - \tilde{\vec{Y}}_{i+1}^{(j)}),$$

where we used the following notation;

$$K_{h_n}(x) \triangleq \frac{1}{h_n^d}K\left(\frac{x}{h_n}\right) \tag{5.29}$$

and where $h_N > 0$ is the bandwidth parameter and N is the number of particles. The value of N, h_N and the kernel must be chosen by the user.

- We set the weight $\omega_{i+1}^{(j)}$ as

$$\omega_{i+1}^{(j)} = \hat{p}^{(j)}(\vec{Y}(t_{i+1})|\vec{f}_i^{(j)}, \vec{Y}(t_i)).$$

- Normalize the weight: for $j = 1, 2, \cdots, N$

$$\hat{\omega}_{i+1}^{(j)} = \frac{\omega_{i+1}^{(j)}}{\sum_{j=1}^{N}\omega_{i+1}^{(j)}}.$$

- The filtering densities are given by

$$p(\vec{f}_{i+1}|\vec{Y}(t_{i+1})) = \sum_{j=1}^{N}\hat{\omega}_{i+1}\delta_{\vec{f}_{i+1}}^{(j)}$$

$$p(\Theta_{i+1}|\vec{Y}(t_{i+1})) = \sum_{j=1}^{N}\hat{\omega}_{i+1}\delta_{\Theta}^{(j)}$$

- Resample for the state $\vec{f}_i^{(j)}$, if we need.

5.5 Simulation Studies

We start by simulation to the observation data, also from the quoted spot prices we identify the parameters of $h(t)$ by standard fitting

procedures, similar to what we have done in section 4.8. We set the system parameters as follows:

$$\kappa = 1.321, \ \lambda_\chi = 0.623, \ \sigma_\chi = 0.3,$$
$$\lambda_\xi = 0.04, \ \sigma_\xi = 0.05, \ \rho = 0.6$$

For the initial conditions of χ, ξ, we use

$$\chi(0) = 0.02, \xi(0) = 0.5.$$

We assume that the covariance kernel of $\sigma w(t, x)$ is given by

$$\sigma^2 q(x, y) = \sum_{k=1}^{100} \sigma^2 e^{-c(x+y)} \left(\frac{\sin(k\pi x)}{5} \right) \left(\frac{\sin(k\pi y)}{5} \right),$$

with $\sigma = 0.02, c = 0.2$. The simulated observation data of the futures is shown in Figure 5.1. For $T - T_0$ fixed to 1-month, the factor process $f(t, x)$ is also demonstrated in Figure 5.2.

Figure 5.1: Observation data

Figure 5.2: $f(t, x)$-process

For the simulation of the convolution particle filter, we assume that the unknown parameters $\{\kappa, \lambda_\chi, \sigma_\chi, \lambda_\xi, \sigma_\xi, \rho, \sigma, c\}$ are random constants where each follows a bounded uniform distribution. The upper and lower bounds are as follows:

$$1.00 \leq \kappa \leq 2.00 \quad , \quad 0.1 \leq \lambda_\chi \leq 1.00$$
$$0.28 \leq \sigma_\chi \leq 0.32 \quad , \quad 0.001 \leq \lambda_\xi \leq 0.06$$
$$0.02 \leq \sigma_\xi \leq 0.06 \quad , \quad 0.20 \leq \rho \leq 0.90$$
$$0.01 \leq \sigma \leq 0.03 \quad , \quad 0.1 \leq c \leq 0.3$$

The initial conditions of $\chi(0)$ and $\xi(0)$ are assumed to be $N(0.02, 0.005)$ and $N(0.9, 0.005)$, respectively.

Now we generate 500 particles for $(\kappa, \lambda_\chi, \sigma_\chi, \lambda_\xi, \sigma_\xi, \rho, \sigma, c)$ vector. Hence,we get the 500 initial conditions $f(0, x)$ from (5.26) and also 500 $\tilde{q}(x, x)$ functions form (5.11). The Parzen-Rosenblatt kernel is set as

$$K_h(\cdot) = \frac{1}{(2\pi h)^m} \exp\{-\frac{|| \cdot ||^2}{2h^2}\},$$

with $h^2 = 0.09$ and m is the dimension of $\vec{Y}$ ($m = 44$).

The estimated $f(t,x)$ process is shown in Figure 5.3. We performed five simulation studies. The estimates for $f(t,x)$ at $x = 0.79$ and $x = 3.19$ are shown in Figure 5.4 and Figure 5.5, respectively. The estimate of $S(t)$ is also shown in Figure 5.6.

Figure 5.3: Estimated $f(t,x)$ process

Figure 5.4: Estimated $f(t, 0.79)$

Figure 5.5: Estimated $f(t, 3.17)$

Figure 5.6: Estimated $S(t)$

Finally we demonstrated the estimation results for unknown parameters $\kappa, \sigma_\chi, \lambda_\chi, \lambda_\xi, \sigma_\xi, \rho, \sigma$ and c. From our estimation of function $\tilde{q}(x, x)$, see Figure 5.7, we noticed that even when some of the parameters are not fitted nicely, the function $\tilde{q}(x, x)$ as a whole is well identified. Notice that for the dynamics of the futures price, see equation (5.9), the value of the function $\tilde{q}(x, x)$ is important.

Figure 5.7: Estimated $\tilde{q}(x, x)$

Figure 5.8: Estimated κ

Figure 5.9: Estimated σ_χ

Figure 5.10: Estimated λ_ξ

True value of σ_ξ(blue) and its estimate

Figure 5.11: Estimated σ_ξ

True value of ρ(blue) and its estimate

Figure 5.12: Estimated ρ

Figure 5.13: Estimated σ

Figure 5.14: Estimated λ_χ

Figure 5.15: Estimated c

5.6 Concluding Remarks

In this chapter, we use the futures price formula with arithmetic average and do not use any approximation (geometric average) to its payoff structure as we have done in chapter 4. The price we have to pay for this is that the observation equation (futures) becomes nonlinear with respect to the states (factors). Hence, we can not use the Kalman filtering methodology and the quasi/maximum likelihood estimation procedure to estimate the parameters. To overcome this problem, we employ the particle filtering approach and in particular, the convolution particle filter. After explaining the merits of the algorithm in general, we recast our model within the same framework. Finally, we run a simulation study to test the feasibility of the proposed filter.

Chapter 6

Conclusions and Future Research

This chapter summarizes the main contributions of this thesis and presents some directions for further research.

6.1   Conclusions

- In Chapter 1 of this thesis we have studied various aspects of the energy markets where electricity and gas are clear examples. Since these markets and in particular the futures markets, form the basis of our subsequent works, a brief description of these markets is discussed in Chapter 1. Moreover, we discuss the Day Ahead (DA) spot price and its characteristic. We show that these markets contain big jumps, trend component and a seasonality component. The jumps considered as outliers which can be removed from the data set. The deterministic components related to the trend and seasonality can be fitted to the data using frequency domain analysis combined with standard fitting procedures. We give examples of implementing these procedures to three different markets namely, APX, EEX and UK-NBP Gas markets. The main objective of this chapter is to get a clean data ready to be used in modeling the dynamics of the uncertainty / stochastic part of the model

- Chapter 2 deals with the mathematical tools needed for the rest of the thesis. We present the Kalman filter in its finite and infinite form. In the finite form we state both the continuous and discrete version of the algorithm together with its corresponding likelihood. Also, since the particle filtering algorithm forms the basis of our subsequent works, a brief description of a generic particle filter algorithm is given in this Chapter. It consists of generating the samples sequentially according to some importance function, then updating the weights of those samples given the recent observation and finally, resampling step when necessary. The roles of both the importance sampling and resampling steps are also discussed. With regard the infinite dimensional filter, we present the algorithm in an intuitive way so that it resembles the finite dimensional version of the filter. Furthermore, we present the infinite dimensional Brownian motion and we discuss the observation mechanism with and without the additive noise in the measurements.

- Chapter 3 deals with the first implementation of the theoretical part of Chapter 2. Here, we illustrate the standard use of the discrete Kalman filter as it used in the literature and the use of MLE in estimating the parameters. After deriving a practical model for the futures price in the energy market which depends on a delivery period, we employ the geometric average approximation to the payoff of the futures price. Then present a sensitivity analysis of the parameters when using the MLE method. The simulations show that the likelihood function is almost flat and it is extremely hard to find the optimal parameters. The second part of this chapter employ the particle filtering method so that we can avoid the geometric approximation. Here also a sensitivity analysis of the parameters when using the MLE method is presented. Also in this case, the optimal parameters are hard to find.

- In Chapter 4, we have introduced our new model for the spot price dynamics. The model is constructed in such a way that

the observation noise is build in within the model and not an extra component just added for the sake of the implementation of the Kalman filtering equations. This exogenously imposed extra noise is used by many authors in the financial mathematics literature, and it was attributed to unconvincing arguments such as bid-ask spread, non synchronicity of the data, etc. However, in this chapter, we approach the modeling differently. In our setup, on one hand, the added measurement noise is built in within the model. On the other hand, the modeling of the correlation structure between the futures (observation) is a natural component of our formulation. The main idea is to assume that the term structure of futures prices on electricity given by Schwartz and Smith (2000) model is affected by an error term represented by a stochastic integral that generates infinite dimensional noise as it should depend on all time of, or to maturity. Hence, we do not need to add artificial noises to the observation equation in order to use the filtering methodology. In this model we employ the geometric approximation so that we can implement the infinite dimensional Kalman filtering algorithm. The factors are estimated as solutions of the resulting filtering problem. The new model has been tested using empirical study on the UK-NBP Gas market and the feasibility of the proposed filter is established.

- Chapter 5 addresses the issues of estimating the parameters of the new model without the use of any approximation to the payoff structure of the futures. As the observation given by the futures price which is found in Chapter 4 is highly non-linear, we propose to use a variant of the particle filter for the identification of the factors and the parameters of the system. This filter is based on the convolution kernel approximation techniques, and termed convolution particle filter, see Rossi and Vila (2006). we present our model in a form ready to be used in this framework. We employ the Bayesian framework where the augmented state variable which represents the state and the unknown parameters, is processed by a filtering procedure. The parameters are considered as random variables with a prescribed priori den-

sity functions. Then an extended state variable joining all the unknown quantities is considered, the posterior density of it is approximated using particle filter. For this reason, we present the approximation method to convert the infinite-dimensional system to the finite dimensional one. At the end of the chapter we present simulation work and studies, using real data to test the model.

6.2 Future Research

In this section we review some issues which can open the direction for future research.

- In Chapter 3, we have done a sensitivity analysis of the likelihood function with respect to the parameters when using the Kalman filter using the MLE method. In practice, the standard approach is to use the pricing formula of the futures as a tool in estimating the parameters through "calibration". This approach resembles the implied volatility of the stock in the Black-Scholes market. Although the second approach is essential for brokers and market-makers, because they have to follow the market, it is not consistent with the long-term investor as these implied parameters are inconsistent. An empirical study that investigate this point in terms of profit/loss and compare it with the results using filtering approach is a direction for further analysis.

- In Chapter 4, we use the geometric average approximation to our futures pricing formula. As a result the state space equations are linear. A new method which took a lot of attention is the use of the Expectation Maximization algorithm when estimating the parameters. However, this method in principle is not recursive and it needs a lot of computational power. Further research of optimizing this method and test it to our model is a direction for further research.

- In Chapter 5, we have used the particle filtering methodology when estimating the state and the parameters. Here, we used the convolution particle filter because our observation equation

has no extra noise term. However, within the convolution filter, the kernel density is arbitrary. Hence, looking at the optimal kernel density either theoretically or experimentally also worth the effort.

- An extensive empirical study to compare the results in Chapter 3, Chapter 4 and Chapter 5, that gives the Pros and Cons of each estimation method is important from practical point of view.

- Clearly, a new direction is to use our model for pricing options and other financial instruments. In our thesis, the focus was only on the futures contract.

Bibliography

Aihara, S., A. Bagchi and E. Imreizeeq (2009*a*). Identification of electricity spot models from futures prices. *Proc. of the 41st ISCIE International Symposium on Stochastic Systems Theory and Its Applications* pp. 217–222.

Aihara, S., A. Bagchi and E. Imreizeeq (2009*b*). Parameter estimation of electricity spot models form futures prices. *Proc. of SYSID09* pp. 1457–1462.

Aihara, S., A. Bagchi and E. Imreizeeq (2011). Identification of electricity spot models by using convolution particle filter. *Int. J. Innovative Computing, Information and Control* **7**(1), 61–72.

Aihara, S. and A. Bagchi (2010*a*). Identification of affine term structures from yield curve data. *International Journal of Theoretical and Applied Finance (IJTAF)* **13**(2), 259–283.

Aihara, S. and A. Bagchi (2010*b*). Parameter estimation of term structures modeled by stochastic hyperbolic systems. *Int. J. Innovative Computing, Information and Control* **6**(1), 171–182.

Arulampalam, S., S. Maskell, N. Gordon and T. Clapp (2002). A tutorial on particle filters for on-line nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing* **50**(2), 174–188.

Bagchi, A. (1975). Continuous time systems identification with unknown noise covariance. *Automatica* **11**, 533–536.

Bagchi, A. (1993). *Optimal Control of Stochastic Systems*. Prentice Hall. Singapore.

Bagchi, A. and V. Borkar (1984). Parameter identification in infinite dimensional linear systems. *Stochastics* **12**, 201–213.

Balakrishnan, A. V. (1969). Identification in automatic control systems. *Automatica* **5**, 817–829.

Balakrishnan, A. V. (1973). *Stochastic Differential Systems I.* Springer-Verlag.

Benth, F., J. Benth and S. Koekebakker (2008). *Stochastic modelling of Electricity and Related Markets.* World Scientific Publishing. UK.

Bølviken, E. and G. Storvik (2001). Deterministic and stochastic particle filters in state-space models. *A. Doucet, et al., (ed.), Sequential Monte Carlo methods in practice*, Chapter 5, Springer-Verlag.

Campillo, F. and V. Rossi (2006). Convolution particle filters for parameter estimation in general state-space models. Technical Report No. 5939. Institut National De Recherche en Informatique et en Automatique.

Clewlow, C. and C. Strickland (1999*a*). A multi-factor model for energy derivatives. Working Paper. School of Finance and Economics, University of Technology, Sydney.

Clewlow, C. and C. Strickland (1999*b*). Valuing energy options in a one factor model fitted to forward prices. Working Paper, School of Finance and Economics, University of Technology, Sydney.

Curtain, R. and A. Pritchard (1978). *Infinite dimensional linear systems theory.* Lecture notes in control and information sciences; 8. Springer-Verlag.

Da Prato, G. and J. Zabczyk (1992). *Stochastic Equations in Infinite Dimensions.* Cambridge University Press.

Doucet, A., S. Godsill and C. Andrieu (2000). On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing* **10**, 197–208.

Elliott, R. and C Hyndman (2007). Parameter estimation in commodity markets: A filtering approach. *Journal of Economic Dynamics and Control* **31**, 2350–2373.

Geman, H. and A. Roncoroni (2006). Understanding the fine structure of electricity prices. *Journal of Business* **79**(3), 1225–1262.

Gibson, R. and E. Schwartz (1990). Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance* **45**, 959–976.

Heath, D., R. Jarrow and A. Morton (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica* **60**, 77–105.

Imreizeeq, E. (2010). Estimation of two factor model of schwartz-smith using particle filter. Working Paper. Department of Applied Mathematics, University of Twente.

Jamshidian, F. (1991). Commodity option evaluation in the gaussian futures term structure model. *The Review of Futures Markets* **2**, 324–346.

Kallianpur, G. (1980). *Stochastic Filtering Theory.* Springer-Verlag. New York.

Kholopova, M. (2006). Estimating a two factor model for the forward rate of electricity. PhD thesis. University of Twente. The Netherland.

Manoliu, M. and S. Tompaidis (2002). Energy futures prices: term structure models with kalman filter estimation. *Applied Mathematical Finance* **9**(1), 21–43.

Moler, C. (2004). *Numerical Computing with MATLAB.* Electronic edition: The Mathworks: http://www.mathworks.com/moler.

Musiela, Marek and Marek Rutkowski (1997). *Martingale Methods in Financial Modelling.* Springer-Verlag.

Oksendal, Bernt (2003). *Stochastic Differential Equations: An Introduction with Applications.* Springer-Verlag.

Ristic, B., S. Arulampalam and N. Gordon (2004). *Beyond the Kalman Filter, Particle filters for tracking applications.* Artech House.

Rossi, V. and J.P. Vila (2006). Nonlinear filtering in discrete time: A particle convolution approach. *Ann. I.SU.P.* **50**(3), 71–102.

Schwartz, E. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance* **52**(3), 923–973.

Schwartz, E. and J. Smith (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science* **46**(7), 893–911.

Abstract

An important issue in the energy market as in the financial market is the parameter estimation of the models representing the dynamics of the spot and the futures. Because one is dealing with unobservable factors, a popular estimation method is the maximum likelihood estimation (MLE), under the assumption that observations are corrupted with additive Gaussian noise. In this framework, the state space representation is used together with the Kalman filtering techniques, and the parameter estimates are obtained through maximization of a likelihood functional. The additional noise in the observation is interpreted to take into account bid-ask spreads, price limits, etc. The argument is clearly forced and unconvincing. The problem is that there is no feedback of the observation noise to the spot price; this leads to a model that is not arbitrage free anymore. The goal of this thesis is two folds:

Starting from the two factor model of Schwartz-Smith (2000), we formulate and implement a new arbitrage free model for the futures prices of energy which can be used in a mathematically sound way when estimating the parameter of the model, using the method of maximum-likelihood. In our setup and by using a reverse engineering concept, a new model is developed by perturbing the futures price given by Schwartz-Smith by extra term that takes into account the uncertainties in both the time, and time of maturity, of the term structure of the futures. We ensure that the new model is arbitrage free. As a result of this formulation, the added measurement noise is built in within the model, and the model parameters can be calibrated through the derived likelihood functional without any ad hoc observation noise.

The second goal is to estimate the parameters of the new model without any modification to the nonlinear payoff of the futures. For that, we use the particle-filtering algorithm to estimate the parameters. On the empirical side, we identify the parameters of the model using real data from the European energy market.

Samenvatting

Een belangrijke vraag in de energie-markt zoals in de financile markt is de parameterschatting van de modellen die de dynamiek van zowel de spot als de futures weergeven. Vanwege het feit dat we met niet-waarneembare factoren werken, een populaire schattingsmethode is de maximum likelihood schatting (MLE), onder de aanname dat de waarnemingen zijn verstoord door additief Gaussische ruis. In dit kader wordt gebruik gemaakt van de toestandsruimte representatie samen met de Kalman filtering technieken, en de parameter schattingen worden verkregen door het maximalisering van een likelihood functioneel. De extra ruis in de waarneming wordt gerechtvaardigd door factoren als bid-ask spreads, prijslimieten, etc. Het argument is duidelijk geforceerd en niet overtuigend. Het probleem is dat er geen terugkoppeling is van de waarnemingsruis naar de spotprijs, en dit leidt tot een model dat niet meer arbitrage vrije is. Het doel van dit proefschrift is tweevouwdig:

Uitgaande van de twee factor model van Schwartz-Smith (2000), we formuleren en implementeren een nieuwe arbitrage vrije model voor de futures-prijzen van de energie dat kan gebruikt worden in een wiskundig verantwoorde manier bij het schatten van de parameter van het model. In onze setup en met behulp van een reverse-engineering concept, een nieuwe model is opgebouwd door het verstoren van de futures prijs gegeven door Schwartz-Smith die rekening houdt met de onzekerheden in zowel de tijd, en tijd van "maturity", van de term structuur van de futures. Wij zorgen ervoor dat het nieuwe model is arbitrage vrije. Als gevolg van deze formulering, is de toegevoegde meetruis ingebouwd bij het model, en de model parameters kunnen worden gekalibreerd door middel van de afgeleide likelihood functioneel zonder enige ad hoc waarnemingsruis.

Het tweede doel is om de parameters van het nieuwe model te schatten, zonder enige lineaire benadering van het niet-lineaire model van de futures. Daarvoor hebben we particle fltering algoritme gebruikt om de parameters te schatten. Aan de empirische kant, identificeren we de parameters van het model met behulp van gegevens uit de Europese energiemarkt.

Acknowledgements

I would like to express my gratitude to those who helped and assisted me in accomplishing this thesis.

First of all, I would like to express my sincere appreciation to my research supervisors, Prof. Arun Bagchi and Prof. Sinichi Aihara, for their kind help, guidance, support and encouragement.

I am grateful to all the members of my promotion committee: Dr. Dominique Dupont, Prof. Anton Stoorvogel, Prof. Michel Vellekoop, Dr. Hans van der Weide and Prof. Marc Wouters for their attentive reading of a draft of this thesis and subsequently supplying me with numerous helpful comments and suggestions.

I would like to thank Essent for partially supporting this research, particular thanks goes to Dr. Mahmoud Hamada.

I also thank all the members of the Systems, Signals and Control group for their pleasant company and to all the members of Business and Accounting group with whom I worked for two years.

A special thanks to all the friends in Enschede and beyond, for their moral support and encouragement.

Finally, I reserve infinite thanks to my family members, as well as to Mrs. A. Bagchi and her mother Mrs. P. Mookerjee. Special thanks go to my parents, my wife Khitam and my two sons (Ameer and Adnan). It would be impossible to accomplish this work without their endless love, patience, encouragement and support.

Enschede,   *Emad S. N. Imreizeeq*
March 2011