PARAMETER ESTIMATION IN A TWO-DIMENSIONAL COMMODITY MODEL

WENXI LIU

A THESIS

IN THE DEPARTMENT
OF
MATHEMATICS AND STATISTICS

PRESENTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE (MATHEMATICS) AT
CONCORDIA UNIVERSITY
MONTREAL, QUEBEC, CANADA
AUGUST 2011

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to verify that the thesis prepared

By: **Wenxi Liu**

Entitled: **Parameter Estimation in a Two-dimensional Commodity Model**

and submitted in partial fulfilment of requirements for the degree of

Master of Science (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. W. Sun ________________________________ Examiner

Dr. A. Sen ________________________________ Examiner

Dr. C. Hyndman ________________________ Thesis Supervisor

Approved by Dr. J. Garrido ________________________________
Chair of Department or Graduate Program Director

_______________2011_____ Dr. B. Lewis ________________________________
Dean of Faculty

ABSTRACT

Parameter Estimation in a Two-dimensional Commodity Model

Wenxi Liu

We consider the problem of estimating the parameters of an unobservable model for the spot price of a commodity. Using the observable time-series of the term-structure of futures prices and a filter-based implementation of the expectation maximization (EM) algorithm, we calculate the maximum likelihood parameter estimates (MLEs). New finite-dimensional filters are derived that allow the EM algorithm to be implemented without calculating Kalman smoother estimates. The method is applied to a two-factor commodity price model.

Keywords: Filter-based EM algorithm; Parameter estimation; Futures prices; Kalman filter

Acknowledgements

Foremost, I would like to express my sincere gratitude to my supervisor Dr. Cody Hyndman for his invaluable support, guidance and throughout my study at Concordia University.

I also would like to take the opportunity to thank for Dr. W. Sun and Dr. A. Sen serving on my thesis committee.

Besides, I would like to thank my colleagues in the department who taught and helped me a lot, not only in academic but also in life. Last but not least, I would like to thank my parents: Xiaoming Liu and Fengyu Lin for their love and understanding all these years.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Elliott and Hyndman (2007) demonstrated the main steps of a filtering approach to parameter estimation for a model of commodity spot prices using observed futures price data. The general method was illustrated with a one-dimensional model. In this thesis, we apply the approach to a higher-dimensional model, the Schwartz (1997) two-factor model.

To find the maximum likelihood parameter estimates, rather than direct maximization of the likelihood function, an iterative procedure, the expectation-maximization (EM) algorithm can be used. Each iteration consist of two steps: expectation (E-step) and maximization (M-step). We follow the work of Elliott and Krishnamurthy (1999) and Elliott and Hyndman (2007) to derive finite-dimensional filters required to obtain the maximum likelihood parameter estimates for the model via the EM algorithm. In order to use the filter-based EM algorithm in higher dimensional commodity models, we found it necessary to extend the results of Elliott and Hyndman (2007). Fontana and Runggaldier (2010) extended the method of Elliott and Hyndman (2007) to apply the filter-based EM algorithm to a credit risk model. We adapted some of the notation in Fontana and Runggaldier (2010) and further extend the method of Elliott and Hyndman (2007) in order to apply the filter-based EM algorithm to the two-factor commodity market model of Schwartz (1997).

The thesis is organized as follows. Chapter 2 reviews the Kalman filter, maximum likelihood parameter estimation (MLE) and the EM algorithm for state space model. Chapter 3 gives details on the application of the Kalman filter and the EM algorithm to the Schwartz (1997) two-factor model. In chapter 4, we present empirical results with simulated data and compare the filter-based approach to the EM algorithm developed in the thesis with the direct maximization of the likelihood function, to estimate model parameters. Conclusions and possible future work is presented in Chapter 5. Proofs of the EM algorithm updates of the model parameters and finite-dimensional filters are given in Appendix A and Appendix B, respectively.

Chapter 2

Kalman Filter and EM Algorithm

In this chapter, we review state space models, the Kalman filter, maximum likelihood estimation (MLE) and the expectation maximization (EM) algorithm. Our main reference for these topics are Anderson and Moore (1979), Harvey (1989), Elliott and Krishnamurthy (1999) and Elliott and Hyndman (2007).

2.1 Kalman Filter

Let θ be the parameter vector in some compact space Θ. A stochastic process $\{X_l\}_{l=0}^{N} \in \mathcal{R}^m$ is defined by the transition equation

$$X_l = c_l(\theta) + Q_l(\theta) X_{l-1} + G_l(\theta)\eta_l, \quad l = 1, \ldots, N. \tag{2.1}$$

We call X_l the state vector, as it is an $m \times 1$ vector and assume X_0 or its distribution is known. However, the sequence $\{X_l\}_{l=0}^{N}$ is not directly observable. In equation (2.1) $c_l(\theta)$ is an $m \times 1$ vector while $Q_l(\theta)$ and $G_l(\theta)$ are both $m \times m$ matrices. Randomness is incorporated by assuming η_l is an $m \times 1$ vector of serially uncorrelated normal random vectors with mean zero and covariance matrix I_m (an $m \times m$ identity matrix) with respect to a probability measure

$\mathcal{P}^{\theta}$.

The measurement equation

$$Y_l = d_l(\theta) + M_l(\theta) X_l + H_l(\theta) \varepsilon_l, \quad l = 0, \ldots, N, \tag{2.2}$$

connects every element of X_l to the $M \times 1$ vector of observed variables at time l, Y_l. In equation (2.1), $M_l(\theta)$ is an $M \times m$ matrix, $d_l(\theta)$ is an $M \times 1$ vector, and ε_l is an $M \times 1$ vector of serially uncorrelated normal random vectors with mean zero and covariance matrix I_M (an $M \times M$ identity matrix) with respect to the probability measure P_θ. The discrete-time Kalman filter is used to estimate the unobserved variables X_l based on observations Y_l. The best mean-square estimate of the state vector, μ_l, among all $\mathcal{Y}_l = \sigma(Y_0, Y_1, \ldots, Y_l)$-measurable functions, according to the definition of conditional expectation, is

$$\mu_l = E[X_l | \mathcal{Y}_l]$$

with covariance matrix

$$P_l = E\left[(X_l - \mu_l)(X_l - \mu_l)' | \mathcal{Y}_l\right]$$

One-step ahead predictions of those quantities are defined by

$$\mu_{l|l-1} = E[X_l | \mathcal{Y}_{l-1}]$$
$$P_{l|l-1} = E\left[\left(X_l - \mu_{l|l-1}\right)\left(X_l - \mu_{l|l-1}\right)' | \mathcal{Y}_{l-1}\right].$$

There are two steps in the implementation of Kalman filter, the prediction step and the updating step. The prediction step is given by the equations

$$\mu_{l|l-1} = c_l(\theta) + Q_l(\theta)\mu_{l-1}, \tag{2.3}$$

$$P_{l|l-1} = G_l(\theta) G_l(\theta)' + Q_l(\theta) P_{l-1} Q_l(\theta)' \tag{2.4}$$

for $l = 1, 2, \ldots, N$.

The updating steps are given by the equations

$$\mu_l = \mu_{l|l-1} + K_l \left(Y_l - M_l(\theta) \mu_{l|l-1} - d_l(\theta) \right), \tag{2.5}$$

$$P_l = P_{l|l-1} - K_l M_l(\theta) P_{l|l-1} \tag{2.6}$$

$$K_l = P_{l|l-1} M_l(\theta)' \left(M_l(\theta) P_{l|l-1} M_l(\theta)' + H_l(\theta) H_l(\theta)' \right)^{-1} \tag{2.7}$$

for $l = 1, 2, \ldots, N$. The quantities μ_0 and P_0 are needed to initialize the Kalman filter and are assumed known.

One way to estimate the parameter vector θ is to consider the observed likelihood. Suppose we have $(N+1)$ observations denoted as $\{Y_0, Y_1, \ldots, Y_N\}$. Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability measures on $(\Omega, \mathcal{F})$ all absolutely continuous with respect to a fixed probability measure P_0. The likelihood function for computing an estimate of the parameter θ based on the observations $\{Y_0, Y_1, \ldots, Y_N\}$ is

$$L(Y_0, \ldots, Y_N; \theta) = E_0 \left[\frac{dP_\theta}{dP_0} | \mathcal{Y}_N \right] \tag{2.8}$$

where $\mathcal{Y}_N = \sigma\{Y_0, Y_1, \ldots, Y_N\}$ is the filtration generated by the observations. Thus the maximum likelihood parameter estimate (MLE) of θ is

$$\hat{\theta} = argmax_{\theta \in \Theta} L(Y_0, \ldots, Y_N; \theta).$$

From the measurement equation (2.2), we can conclude that the conditional distribution of Y_i given $Y_0, \ldots, Y_{i-1}$ is normal with mean vector

$$\hat{Y}_i = d_i(\theta) + M_i(\theta) \mu_{i|i-1}$$

and covariance matrix

$$M_i(\theta) P_{i|i-1} M_i(\theta)' + H_i(\theta) H_i(\theta)'.$$

The log-likelihood function, $l(Y_0,\ldots,Y_N;\theta) = \log L(Y_0,\ldots,Y_N;\theta)$ can be written in *prediction error decomposition* form

$$
\begin{aligned}
&l(Y_0,\ldots,Y_N;\theta) \\
&= -\frac{MN}{2} log(2\pi) - \frac{1}{2}\sum_{l=0}^{N} log\,det\left(M_i(\theta)P_{i|i-1}M_i(\theta)' + H_i(\theta)H_i(\theta)'\right) \\
&-\frac{1}{2}\sum_{l=0}^{N}\left(Y_i-\hat{Y}_i\right)'\left[M_i(\theta)P_{i|i-1}M_i(\theta)' + H_i(\theta)H_i(\theta)'\right]^{-1}\left(Y_i-\hat{Y}_i\right). \qquad (2.9)
\end{aligned}
$$

Then numerical maximization techniques can be applied to equation (2.9) to find the MLE $\hat{\theta}$.

2.2 EM algorithm

An alternative to direct maximization of equation (2.9) is the Expectation Maximization (EM) algorithm. Let $\hat{\theta}_0$ be an initial parameter estimate. The EM algorithm produces a series of parameter estimates $\{\hat{\theta}_j\}_{j=1}^{\infty}$ which should converge to the MLE $\hat{\theta}$ provided certain conditions are satisfied as Wu (1983). Two steps are involved in the EM algorithm for each iteration. The first step, called the Expectation step (E-step) is to compute $Q\left(\hat{\theta}_j,\cdot\right)$, where

$$
Q\left(\hat{\theta}_j,\theta\right) = E_{\hat{\theta}_j}\left[log\frac{dP_\theta}{dP_{\hat{\theta}_j}}|\mathcal{Y}_N\right]. \qquad (2.10)
$$

It can be shown that

$$
\begin{aligned}
Q\left(\hat{\theta}_j,\theta\right) &= -\sum_{l=1}^{N}\ln|G_l(\theta)| - \sum_{l=0}^{N}\ln|H_l(\theta)| + E_{\hat{\theta}_j}\left[R\left(\hat{\theta}_j\right)|\mathcal{Y}_N\right] \\
&-\frac{1}{2}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}(X_l - Q_l(\theta)X_{l-1} - c_l(\theta))'G_l^{-2}(\theta)(X_l - Q_l(\theta)X_{l-1} - c_l(\theta))|\mathcal{Y}_N\right] \\
&-\frac{1}{2}E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}(Y_l - M_l(\theta)X_l - d_l(\theta))'H^{-2}(\theta)(Y_l - M_l(\theta)X_l - d_l(\theta))|\mathcal{Y}_N\right] \qquad (2.11)
\end{aligned}
$$

where $G_l^{-2}(\theta)$ is $\left(G_l^{-1}(\theta)\right)'\left(G_l^{-1}(\theta)\right)$,$H_l^{-2}(\theta)$ is $\left(H_l^{-1}(\theta)\right)'\left(H_l^{-1}(\theta)\right)$ and $R\left(\hat{\theta}_j\right)$ is for terms not depending on θ. In the next section, we give detail on how to derive equation

(2.11) using the reference probability measure technique which has been used extensively in filtering theory. The second step, called the Maximization step (M-step) is to find $\hat{\theta}_{j+1} = arg\ max_{\theta\in\Theta} Q\left(\hat{\theta}_k, \theta\right)$. In practice, we set the derivatives of Q_k as in equation (2.11) to zero and solve for θ.

2.3 Reference Probability Measure Technique

In order to compute the expectation in equation (2.10) and derive equation (2.11), we use a reference probability measure $\bar{P}$, which is easier to work with. Define the densities

$$\varphi(x) = (2\pi)^{-m/2} e^{-x'x/2} for x \in \mathcal{R}^m \tag{2.12}$$

$$\phi(y) = (2\pi)^{-M/2} e^{-y'y/2} for y \in \mathcal{R}^M. \tag{2.13}$$

Define the σ-fields $\mathcal{G}_k$=$\sigma\{X_0, \ldots, X_k, Y_0, Y_1, \ldots, Y_k\}$ and $\mathcal{G}_\infty = \vee_{k=1}^{\infty} \mathcal{G}_k$. We assume the existence of a probability measure $\bar{P}$ on $(\Omega, \mathcal{G}_\infty)$ such that under $\bar{P}$, the X_k are i.i.d with density in equation (2.12) and Y_k are i.i.d with density in equation (2.13). That is to say, $X_k \in \mathcal{R}^m$ are i.i.d Gaussian random variables and $Y_k \in \mathcal{R}^M$ are i.i.d Gaussian random variables. From the reference probability measure $\bar{P}$, we construct the real-world probability measure P_θ under which X_k and Y_k satisfy equation (2.1) and (2.2). Define

$$\lambda_0 = \frac{\phi(H_0(\theta)(Y_0 - M_0(\theta)X_0 - d_0(\theta)))}{|H_0(\theta)|\phi(Y_0)}.$$

For $l \geq 1$, define

$$\lambda_l = \frac{\phi(H_l(\theta)(Y_l - M_l(\theta)X_l - d_l(\theta)))}{|H_l(\theta)|\phi(Y_l)} \frac{\varphi\left(G_l^{-1}(\theta)(X_l - Q_l(\theta)X_{l-1} - c_l(\theta))\right)}{|G_l(\theta)|\varphi(X_l)}$$

For $k \geq 0$, define

$$\Lambda_k = \prod_{l=0}^{k} \lambda_l \tag{2.14}$$

Now we can define a new probability measure P_θ on $(\Omega, \mathcal{G}_\infty)$ by setting the $\mathcal{G}_k$-restriction of the Radon-Nikodym derivative of P_θ with respect to $\bar{P}$ to be

$$\left.\frac{dP_\theta}{d\bar{P}}\right|_{\mathcal{G}_k} = \Lambda_k.$$

Kolmogorov's extension theorem (Karatzas and Shreve (1991), p.50) shows the existence of P_θ on $\mathcal{G}_\infty$. Lemma 2.1 of Elliott and Hyndman (2007) showed that under the new measure P_θ,

$$\varepsilon_l = H_l^{-1}(\theta)\,(Y_l - M_l(\theta) - d_l(\theta)) \quad l = 0, 1, \ldots, N, \tag{2.15}$$

$$\eta_l = G_l^{-1}(\theta)\,(X_l - Q_l(\theta) X_{l-1} - c_l(\theta)) \quad l = 1, 2, \ldots, N \tag{2.16}$$

are sequences of i.i.d $N(0, I_m)$ and i.i.d $N(0, I_M)$ Gaussian random variables. So from the measure $\bar{P}$ under which the signal and observation are i.i.d Gaussian random variables, Lemma 2.1 of Elliott and Hyndman (2007) enables us to construct a measure P_θ under which the signal and observation satisfy the dynamics given by equation (2.1) and (2.2) with vectors and coefficient matrices $(c(\theta), d_l(\theta), Q(\theta), M_l(\theta), G(\theta), H(\theta))$. We should point out that the process of constructing P_θ based on the existence of $\bar{P}$ is invertible. So, starting with the real world probability measure P_θ, we can construct $\bar{P}$ directly without necessarily assuming the existence of it. We define $\bar{P}$ by

$$\left.\frac{d\bar{P}}{dP_\theta}\right|_{\mathcal{G}_k} = \Lambda_k^{-1}.$$

It can be shown that the signal and observation are mutually independent normal random variable under $\bar{P}$ using Bayes's theorem. That is, for any measurable test functions f, g

$$\bar{E}\left[f(X_k)\,g(Y_k)\,|\mathcal{G}_{k-1}\right] = \int_{\mathcal{R}^m} f(x)\,\varphi(x)\,dx \int_{\mathcal{R}^M} g(y)\,\phi(y)\,dy.$$

Further details on deriving equation (2.11) can be found in Section 2.4 of Elliott and Hyndman (2007).

Chapter 3

Application to a Commodity Model

Elliott and Hyndman (2007) implemented the EM algorithm to estimate the parameters of a commodity price model based on geometric Brownian motion using simulated data and compared this to direct maximization of the likelihood.

However, the geometric Brownian motion model doesn't capture the observed empirical features of a commodity price model since it assumes the instantaneous convenience yield and volatility are both constant. The two-factor model of Schwartz (1997) assumes, the spot price of the commodity, S, and the instantaneous convenience yield, δ to follow the joint stochastic process:

$$dS_t = (\mu - \delta_t) S_t dt + \sigma_1 S_t dB_t^1 \tag{3.1}$$

$$d\delta_t = \kappa(\alpha - \delta_t) dt + \sigma_2 dB_t^2 \tag{3.2}$$

where B_t^1 and B_t^2 are two correlated standard Brownian motion with

$$d < B_\cdot^1, B_\cdot^2 >_t = \rho dt$$

In equation (3.2) the instantaneous convenience yield follows a Ornstein-Uhlenbeck stochastic process. The parameter κ is the speed of mean-reversion and α is the long-term mean of instantaneous convenience yield.

Define $Z_t = lnS_t$, thus by Ito's lemma, the log-price of the commodity follows the process:

$$dZ_t = \left(\mu - \delta_t - \frac{1}{2}\sigma_1^2\right) dt + \sigma_1 dB_t^1 \tag{3.3}$$

Under an equivalent martingale measure Q, the joint process for the spot commodity price and the stochastic instantaneous convenience yield is as follows:

$$dS_t = (r - \delta_t) S_t dt + \sigma_1 S_t dB_t^{*1}$$

$$d\delta_t = [\kappa(\alpha - \delta_t) - \lambda] dt + \sigma_2 dB_t^{*2}$$

$$d < B_.^{*1}, B_.^{*2} >_t = \rho dt$$

where λ is the market price of the convenience yield risk and is assumed constant. The observations used to estimate the model parameters are futures prices. A futures contract is an agreement between two parties to exchange a specified asset of standardized quantity and quality for a price agreed today at a certain point in the future. In the market, people use the mechanism of marking to market so that the value of the contract is zero at any time. Motivated by the features of a futures contract, we now define the futures price.

Definition 3.1. *The futures price at time t for maturity T of the $\mathcal{F}_T$-measurable underlying asset S_T is $F(S,t,T) = E_Q[S_T|\mathcal{F}_t]$, where $E_Q[\,\cdot\,]$ is the expectation with respect to the risk-neutral measure Q.*

Jamishidian and Fein (1990) and Bjerksund (1995) have shown that the futures price at time t with time to maturity $\tau = T - t$ is:

$$F(S, \delta, \tau) = S_t e^{-\delta \frac{1-e^{-\kappa\tau}}{\kappa} + A(\tau)} \tag{3.4}$$

where

$$A(\tau) = \left(r - \alpha + \frac{\lambda}{\kappa} + \frac{1}{2}\frac{\sigma_2^2}{\kappa^2} - \frac{\sigma_1\sigma_2\rho}{\kappa} \right)\tau + \frac{1}{4}\sigma_2^2\frac{1-e^{-2\kappa\tau}}{\kappa^3}$$
$$+ \left(\alpha\kappa - \lambda + \sigma_1\sigma_2\rho - \frac{\sigma_2^2}{\kappa} \right)\frac{1-e^{-\kappa\tau}}{\kappa^2}. \quad (3.5)$$

Thus, the log-futures price is

$$lnF(S,\delta,\tau) = lnS - \delta\frac{1-e^{-\kappa\tau}}{\kappa} + A(\tau). \quad (3.6)$$

Equation (3.4)-(3.5) can be derived by considering the expectation directly or using the method of stochastic flows as in Hyndman (2007).

3.1 Empirical Model

To implement the model, we discretize equation (3.2) and equation (3.3) using an Euler scheme as is done in Schwartz (1997). An empirical model based on the Kalman filter can then be used to estimate the parameters of (3.2)-(3.3) using observed futures prices. Define $X_l = [Z_l, \delta_l]'$. From equations (3.2)-(3.3), the transition equation can be written as

$$X_l = c + QX_{l-1} + G\eta_l,\ l = 1,\ldots,N \quad (3.7)$$

where

$$c = \begin{bmatrix} \left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t \\ \kappa\alpha\Delta t \end{bmatrix},$$

$$Q = \begin{bmatrix} 1 & -\Delta t \\ 0 & 1-\kappa\Delta t \end{bmatrix},$$

$$G = \begin{bmatrix} \sigma_1\sqrt{\Delta t} & 0 \\ \sigma_2\rho\sqrt{\Delta t} & \sigma_2\sqrt{\Delta t(1-\rho^2)} \end{bmatrix},$$

and η_l, $l=1,\ldots,N$ are serially uncorrelated random vectors with

$$E(\eta_l)=0 \quad \text{and} \quad Var(\eta_l)=I_{m\times m}.$$

From equation (3.5) and equation (3.6), the measurement equation can be obtained as,

$$Y_l(\theta)=d_l(\theta)+M_l(\theta)X_l+H(\theta)\varepsilon_l, \quad l=0,\ldots,N \tag{3.8}$$

where for $i=1,\ldots,M$, $\tau_l^{(i)}$ is the maturity of the i-th futures contract at step $l=0,\ldots,N$. Then we assume that the observed futures prices given by the market, $\hat{F}(\tau_l^{(i)})$ are the theoretical futures prices given equation (3.4) plus an additional noise term. That is

$$Y_l=\begin{bmatrix} \ln\hat{F}\left(\tau_l^{(1)}\right) \\ \vdots \\ \ln\hat{F}\left(\tau_l^{(M)}\right) \end{bmatrix},$$

is given by equation (3.6),

$$d_l(\theta)=\begin{bmatrix} A\left(\tau_l^{(1)}\right) \\ \vdots \\ A\left(\tau_l^{(M)}\right) \end{bmatrix},$$

is given by equation (3.5) and

$$M_l(\theta)=\begin{bmatrix} 1, & -\frac{1-e^{-\kappa\tau_l^{(1)}}}{\kappa} \\ \vdots, & \vdots \\ 1, & -\frac{1-e^{-\kappa\tau_l^{(M)}}}{\kappa} \end{bmatrix},$$

is an $M\times 2$ matrix with first column equal to the vector $I_{M\times 1}$ with each of its elements 1. Also, ε_l, $l=1,\ldots,N$ are $M\times 1$ vector of serially uncorrelated normal random vectors with

$$E(\varepsilon_l)=0 \quad \text{and} \quad Var(\varepsilon_l)=I_{M\times M}$$

and H_l is a constant $M\times M$ diagonal matrix. Substituting the particular state spate formulation for the model in equation (2.11), taking derivatives and solving for θ gives

Theorem 3.1. *The revised EM parameter estimate* $\hat{\theta}_{j+1} = \left(\mu, \kappa, \alpha, \sigma_1, \sigma_2, \rho, \lambda, H^2\right)$ *at iteration* $j+1$ *satisfies*

$$\mu = \frac{e_1' G^{-2} \left(\hat{L}_N^{(2)} - Q_l \hat{L}_N^{(1)} + \frac{1}{2}\sigma_1^2 N \Delta t e_1 - \kappa\alpha\Delta t N e_2 \right)}{N \Delta t e_1' G^{-2} e_1} \tag{3.9}$$

where

$$\hat{L}_N^{(1)} = E_{\hat{\theta}_j} \left[\sum_{l=1}^{N} X_{l-1} | \mathcal{Y}_N \right] \tag{3.10}$$

$$\hat{L}_N^{(2)} = E_{\hat{\theta}_j} \left[\sum_{l=1}^{N} X_l | \mathcal{Y}_N \right]. \tag{3.11}$$

Theorem 3.2. *The revised EM parameter estimate* $\hat{\theta}_{j+1} = \left(\mu, \kappa, \alpha, \sigma_1, \sigma_2, \rho, \lambda, H^2\right)$ *at iteration* $j+1$ *satisfies*

$$\lambda = \left[\sum_{l=0}^{N} \left(A_l^\lambda\right)' H^{-2} A_l^\lambda \right]^{-1} \times \left[\sum_{l=0}^{N} \left(A_l^\lambda\right)' H^{-2} \left(Y_l - B_l^\lambda\right) - \hat{K}_N \left(A_l^\lambda\right) \right] \tag{3.12}$$

where

$$\begin{aligned} A_l^\lambda &= \frac{\tau_l}{\kappa} - \frac{1 - e^{-\kappa\tau_l}}{\kappa^2} \\ B_l^\lambda &= \left(r - \alpha + \frac{1}{2}\frac{\sigma_2^2}{\kappa^2} - \frac{\sigma_1\sigma_2\rho}{\kappa} \right)\tau_l + \frac{1}{4}\frac{\sigma_2^2\left(1 - e^{-2\kappa\tau_l}\right)}{\kappa^3} + \left(\alpha\kappa + \sigma_1\sigma_2\rho - \frac{\sigma_2^2}{\kappa} \right)\frac{1 - e^{-\kappa\tau_l}}{\kappa^2} \\ \hat{K}_N\left(A_l^\lambda\right) &= E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} \left(A_l^\lambda\right)' H^{-2} M_l X_l | \mathcal{Y}_N \right]. \end{aligned} \tag{3.13}$$

Theorem 3.3. *The revised EM parameter estimate* $\hat{\theta}_{j+1} = \left(\mu, \kappa, \alpha, \sigma_1, \sigma_2, \rho, \lambda, H^2\right)$*at iteration* $j+1$ *satisfies*

$$\alpha = \frac{\kappa\Delta t e_2' G^{-2} \left(\hat{L}_N^{(2)} - Q_l \hat{L}_N^{(1)} - \left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t e_1 N \right) + \sum_{l=0}^{N} \left(A_l^\alpha\right)' H^{-2} \left(Y_l - B_l^\alpha\right) - \hat{K}_N\left(A_l^\alpha\right)}{\kappa^2 \Delta t^2 N e_2' G^{-2} e_2 + \sum_{l=0}^{N} \left(A_l^\alpha\right)' H^{-2} A_l^\alpha} \tag{3.14}$$

where

$$A_l^{\alpha} = -\tau_l + \frac{1-e^{-\kappa\tau_l}}{\kappa}$$
$$B_l^{\alpha} = \left(r + \frac{\lambda}{\kappa} + \frac{1}{2}\frac{\sigma_2^2}{\kappa^2} - \frac{\sigma_1\sigma_2\rho}{\kappa}\right)\tau_l + \frac{1}{4}\frac{\sigma_2^2\left(1-e^{-2\kappa\tau_l}\right)}{\kappa^3} + \left(-\lambda + \sigma_1\sigma_2\rho - \frac{\sigma_2^2}{\kappa}\right)\frac{1-e^{-\kappa\tau_l}}{\kappa^{-2}}$$
$$\hat{K}_N\left(A_l^{\alpha}\right) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(A_l^{\alpha}\right)' H^{-2} M_l X_l | \mathcal{Y}_N\right].$$

Theorem 3.4. *The revised EM parameter estimate* $\hat{\theta}_{j+1} = \left(\mu, \kappa, \alpha, \sigma_1, \sigma_2, \rho, \lambda, H^2\right)$ *at iteration* $j+1$ *satisfies the polynomial equation*

$$a_4^{\sigma_1}\sigma_1^4 + a_3^{\sigma_1}\sigma_1^3 + a_2^{\sigma_1}\sigma_1^2 + a_1^{\sigma_1}\sigma_1 + a_0^{\sigma_1} = 0 \tag{3.15}$$

with

$$a_4^{\sigma_1} = -\sum_{l=0}^{N}\left(A_l^{\sigma_1}\right)' H^{-2} A_l^{\sigma_1} + \frac{N\Delta t}{4\left(1-\rho^2\right)}$$
$$a_3^{\sigma_1} = -\frac{\rho}{2\sigma_2\left(1-\rho^2\right)} \times \left(e_2'\hat{L}_N^{(2)} - (1-\kappa\Delta t)\, e_2'\hat{L}_N^{(1)} - \kappa\alpha\Delta t N\right)$$
$$+ \sum_{l=0}^{N}\left(A_l^{\sigma_1}\right)' H^{-2}\left(Y_l - B_l^{\sigma_1}\right) - \hat{K}_N\left(A_l^{\sigma_1}\right)$$
$$a_2^{\sigma_1} = -N + \frac{e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - N\mu\Delta t}{\left(1-\rho^2\right)}$$
$$a_1^{\sigma_1} = -\frac{\rho}{\sigma_2\left(1-\rho^2\right)\Delta t}\left(e_2'\hat{H}_N^{(3)} e_1\right.$$
$$-e_2'\hat{H}_N^{(1)}\left(e_1 - \Delta t e_2\right) - (1-\kappa\Delta t)\, e_1'\hat{H}_N^{(1)} e_2 + (1-\kappa\Delta t)\, e_2' H_N^{(2)}\left(e_1 - \Delta t e_2\right)$$
$$\left. -\mu\Delta t\left(e_2'\hat{L}_N^{(2)} - (1-\kappa\Delta t)\hat{L}_N^{(1)}\right) - \kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - \mu\Delta t N\right)\right)$$
$$a_0^{\sigma_1} = \frac{1}{\left(1-\rho^2\right)\Delta t}\left(e_1'\hat{H}_N^{(3)} e_1 - 2e_1'\hat{H}_N^{(1)}\left(e_1 - \Delta t e_2\right)\right.$$
$$\left. + \left(e_1' - \Delta t e_2'\right)\hat{H}_N^{(2)}\left(e_1 - \Delta t e_2\right) + \mu^2\Delta t^2 N - 2\mu\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)}\right)\right)$$

where

$$
\begin{aligned}
A_l^{\sigma_1} &= -\frac{\sigma_2\rho\tau_l}{\kappa} + \frac{\sigma_2\rho\left(1-e^{-\kappa\tau_l}\right)}{\kappa^2} \\
B_l^{\sigma_1} &= \left(r-\alpha+\frac{\lambda}{\kappa}+\frac{1}{2}\frac{\sigma_2^2}{\kappa^2}\right)\tau_l + \frac{1}{4}\frac{\sigma_2^2\left(1-e^{-2\kappa\tau_l}\right)}{\kappa^3} + \left(\alpha\kappa-\lambda-\frac{\sigma_2^2}{\kappa}\right)\frac{1-e^{-\kappa\tau_l}}{\kappa^2} \\
\hat{K}_N\left(A_l^{\sigma_1}\right) &= E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(A_l^{\sigma_1}\right)' H^{-2} M_l X_l | \mathcal{Y}_N\right] \\
\hat{H}_N^{(1)} &= E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} X_l X_{l-1}' | \mathcal{Y}_N\right] \qquad (3.16) \\
\hat{H}_N^{(2)} &= E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} X_{l-1} X_{l-1}' | \mathcal{Y}_N\right] \qquad (3.17) \\
\hat{H}_N^{(3)} &= E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} X_l X_l' | \mathcal{Y}_N\right] \qquad (3.18)
\end{aligned}
$$

and $\hat{L}_N^{(1)}$ *and* $\hat{L}_N^{(2)}$ *are defined as equation* (3.10) *and* (3.11) *respectively.*

Theorem 3.5. *The revised EM parameter estimate* $\hat{\theta}_{j+1} = \left(\mu,\kappa,\alpha,\sigma_1,\sigma_2,\rho,\lambda,H^2\right)$ *at iteration* $j+1$ *satisfies the polynomial equation*

$$
a_6^{\sigma_2}\sigma_2^6 + a_5^{\sigma_2}\sigma_2^5 + a_4^{\sigma_2}\sigma_2^4 + a_3^{\sigma_2}\sigma_2^3 + a_2^{\sigma_2}\sigma_2^2 + a_1^{\sigma_2}\sigma_2 + a_0^{\sigma_2} = 0 \qquad (3.19)
$$

with

$$
\begin{aligned}
a_6^{\sigma_2} &= -\frac{1}{2}\sum_{l=0}^{N}\left(A_l^{\sigma_2}\right)' H^{-2} A_l^{\sigma_2} \\
a_5^{\sigma_2} &= -\sum_{l=0}^{N}\left(A_l^{\sigma_2}\right)' \hat{H}^{-2} B_l^{\sigma_2} - \frac{1}{2}\sum_{l=0}^{N}\left(B_l^{\sigma_2}\right)' \hat{H}^{-2} A_l^{\sigma_2} \\
a_4^{\sigma_2} &= \sum_{l=0}^{N}\left(A_l^{\sigma_2}\right)' H^{-2}\left(Y_l - D_l^{\sigma_2}\right) - \hat{K}_N\left(A_l^{\sigma_2}\right) - \sum_{l=0}^{N}\left(B_l^{\sigma_2}\right)' \hat{H}^{-2} B_l^{\sigma_2} \\
a_3^{\sigma_2} &= \sum_{l=0}^{N}\left(B_l^{\sigma_2}\right)' H^{-2}\left(Y_l - D_l^{\sigma_2}\right) - \hat{K}_N\left(B_l^{\sigma_2}\right) \\
a_2^{\sigma_2} &= -N
\end{aligned}
$$

$$
\begin{aligned}
a_1^{\sigma_2} =& -\frac{\rho}{\sigma_1\left(1-\rho^2\right)\Delta t}\Big(e_2'\hat{H}_N^{(3)}e_1 - e_2'\hat{H}_N^{(1)}\left(e_1-\Delta t e_2\right) \\
& -\left(1-\kappa\Delta t\right)e_1'\hat{H}_N^{(1)}e_2 + \left(1-\kappa\Delta t\right)e_1'\hat{H}_N^{(2)}\left(e_1-\Delta t e_2\right) \\
& -\left(\mu-\frac{1}{2}\sigma_1^2\right)\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(1-\kappa\Delta t\right)e_2'\hat{L}_N^{(1)}\right) \\
& -\kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - N\left(\mu-\frac{1}{2}\sigma_1^2\right)\Delta t\right)\Big) \\
a_0^{\sigma_2} =& \frac{1}{\Delta t\left(1-\rho^2\right)}\Big(e_2'\hat{H}_N^{(3)}e_2 - 2e_2'\hat{H}^{(1)}\left(e_1-\Delta t e_2\right) \\
& +\left(1-\kappa\Delta t\right)^2 e_2'\hat{H}_N^{(2)}e_2 - 2\kappa\alpha\Delta t\left(e_2'\hat{L}_N^{(2)} - \left(1-\kappa\Delta t\right)e_2'\hat{L}_N^{(1)}\right) + N\kappa^2\alpha^2\Delta t^2\Big)
\end{aligned}
$$

where

$$
\begin{aligned}
A_l^{\sigma_2} &= \frac{\tau_l}{\kappa^2} + 0.5\frac{1-e^{-2\kappa\tau_l}}{\kappa^3} - 2\frac{1-e^{-\kappa\tau_l}}{\kappa^3} \\
B_l^{\sigma_2} &= \sigma_1\rho\left(\frac{\tau}{\kappa} - \frac{1-e^{-\kappa\tau_l}}{\kappa^2}\right) \\
D_l^{\sigma_2} &= \left(r-\alpha+\frac{\lambda}{\kappa}\right)\tau_l + \left(\alpha\kappa-\lambda\right)\frac{1-e^{\kappa\tau_l}}{\kappa^2} \\
\hat{K}_N\left(A_l^{\sigma_2}\right) &= E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(A_l^{\sigma_2}\right)' H^{-2} M_l X_l | \mathcal{Y}_N\right] && (3.20) \\
\hat{K}_N\left(B_l^{\sigma_2}\right) &= E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(B_l^{\sigma_2}\right)' H^{-2} M_l X_l | \mathcal{Y}_N\right] && (3.21)
\end{aligned}
$$

and $\hat{L}_N^{(1)}$, $\hat{L}_N^{(2)}$, $\hat{H}_N^{(1)}$, $\hat{H}_N^{(2)}$, $\hat{H}_N^{(3)}$ *are calculated using equation* (3.10)-(3.18).

Theorem 3.6. *The revised EM parameter estimate* $\hat{\theta}_{j+1} = \left(\mu,\kappa,\alpha,\sigma_1,\sigma_2,\rho,\lambda,H^2\right)$ *at iteration* $j+1$ *satisfies the polynomial equation*

$$
a_5^{\rho}\rho^5 + a_4^{\rho}\rho^4 + a_3^{\rho}\rho^3 + a_2^{\rho}\rho^2 + a_1^{\rho}\rho + a_0^{\rho} = 0 \qquad (3.22)
$$

with

$$
\begin{aligned}
a_5^{\rho} =& -\sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^{-2} A_l^{\rho} \\
a_4^{\rho} =& \sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^{-2} \left(Y_l - B_l^{\rho}\right) - \hat{K}_N \left(A_l^{\rho}\right) \\
a_3^{\rho} =& 2\sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^{-2} A_l^{\rho} - N \\
a_2^{\rho} =& -2\left(\sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^{-2} \left(Y_l - B_l^{\rho}\right) - \hat{K}_N^{(6)}\right) + \frac{1}{\sigma_1\sigma_2\Delta t\left(1-\rho^2\right)^2}\left(e_2'\hat{H}_N^{(3)} e_1\right. \\
& - e_2'\hat{H}_N^{(1)}\left(e_1 - \Delta t e_2\right) - (1-\kappa\Delta t)\, e_1'\hat{H}_N^{(1)} e_2 + (1-\kappa\Delta t)\, e_2' H_N^{(2)}\left(e_1 - \Delta t e_2\right) \\
& - \left(\mu - \frac{1}{2}\sigma_1^2\right)\left(e_2'\hat{L}^{(2)} - (1-\kappa\Delta t)\, e_2'\hat{L}_N^{(1)}\right) \\
& \left. - \kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - N\left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t\right)\right) \\
a_1^{\rho} =& N - \sum_{l=0}^{N} \left(A_l^{\rho}\right)' \hat{H}^{-2} A_l^{\rho} - \frac{\rho}{\sigma_1^2\Delta t\left(1-\rho^2\right)^2}\left(e_1'\hat{H}_N^{(3)} e_1 - 2 e_1'\hat{H}_N^{(1)}\left(e_1 - \Delta t e_2\right)\right. \\
& + \left(e_1' - \Delta t e_2'\right)\hat{H}_N^{(2)}\left(e_1 - \Delta t e_2\right) - 2\left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)}\right) \\
& \left. + \left(\mu - \frac{1}{2}\sigma_1^2\right)^2 \Delta t^2 N\right) - \frac{\rho}{\sigma_2^2\Delta t\left(1-\rho^2\right)^2}\left(e_2'\hat{H}_N^{(3)} e_2 - 2 e_2'\hat{H}^{(1)}\left(e_1 - \Delta t e_2\right)\right. \\
& \left. + (1-\kappa\Delta t)^2 e_2'\hat{H}_N^{(2)} e_2 - 2\kappa\alpha\Delta t\left(e_2'\hat{L}_N^{(2)} - (1-\kappa\Delta t)\, e_2'\hat{L}_N^{(1)}\right) + N\kappa^2\alpha^2\Delta t^2\right) \\
a_0^{\rho} =& \frac{1}{\sigma_1\sigma_2\Delta t\left(1-\rho^2\right)^2}\left(e_2'\hat{H}_N^{(3)} e_1 - e_2'\hat{H}_N^{(1)}\left(e_1 - \Delta t e_2\right)\right. \\
& - (1-\kappa\Delta t)\, e_1'\hat{H}_N^{(1)} e_2 + (1-\kappa\Delta t)\, e_2'\hat{H}_N^{(2)}\left(e_1 - \Delta t e_2\right) \\
& - \left(\mu - \frac{1}{2}\sigma_1^2\right)\left(e_2'\hat{L}_N^{(2)} - (1-\kappa\Delta t)\, e_2'\hat{L}_N^{(1)}\right) \\
& \left. - \kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - N\left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t\right)\right) \\
& + \sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^{-2} \left(Y_l - B_l^{\rho}\right) - \hat{K}_N \left(A_l^{\rho}\right)
\end{aligned}
$$

where

$$
\begin{aligned}
A_l^{\rho} &= -\frac{\sigma_1\sigma_2\tau_l}{\kappa} + \frac{\sigma_1\sigma_2\left(1-e^{-\kappa\tau_l}\right)}{\kappa^2} \\
B_l^{\rho} &= \left(r - \alpha + \frac{\lambda}{\kappa} + \frac{1}{2}\frac{\sigma_2^2}{\kappa^2}\right)\tau_l + \frac{1}{4}\sigma_2^2\frac{1-e^{-2\kappa\tau_l}}{\kappa^3} + \left(\alpha\kappa - \lambda - \frac{\sigma_2^2}{\kappa}\right)\frac{1-e^{-\kappa\tau_l}}{\kappa^2} \\
\hat{K}_N\left(A_l^{\rho}\right) &= E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(A_l^{\rho}\right)' H^{-2} M_l X_l | \mathcal{Y}_N\right] \qquad (3.23)
\end{aligned}
$$

and $\hat{L}_N^{(1)}$, $\hat{L}_N^{(2)}$, $\hat{H}_N^{(1)}$, $\hat{H}_N^{(2)}$, $\hat{H}_N^{(3)}$ *are calculated using equation* (3.10)-(3.18).

Theorem 3.7. *The revised EM parameter estimate* $\hat{\theta}_{j+1} = \left(\mu, \kappa, \alpha, \sigma_1, \sigma_2, \rho, \lambda, H^2\right)$ *at iteration* $j+1$ *satisfies the polynomial equation*

$$
\begin{aligned}
H^2 &= \frac{1}{N+1}E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(Y_l - M_l X_l - d_l\right)\left(Y_l - M_l X_l - d_l\right)' | Y_N\right] \\
&= \frac{1}{N+1}\left[\sum_{l=0}^{N} Y_l Y_l' - \sum_{l=0}^{N} Y_l d_l' - \hat{J}_N - \left(\hat{J}_N\right)' + \hat{U}_N - \sum_{l=0}^{N} d_l Y_l' + \sum_{l=0}^{N} d_l d_l'\right] \qquad (3.24)
\end{aligned}
$$

where

$$
\hat{J}_N = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} M_l X_l H^{-2}\left(Y_l - d_l\right)' | \mathcal{Y}_N\right] \qquad (3.25)
$$

$$
\hat{U}_N = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} M_l X_l X_l' M_l' | \mathcal{Y}_N\right]. \qquad (3.26)
$$

For the parameter κ, due to the terms of $e^{-\kappa\tau}$ and $e^{-2\kappa\tau}$, if we take the derivative of equation (2.11) with respect to κ and set it to zero directly, we can not obtain an equation that must be satisfied and easily implemented. Therefore we use an approximation after taking the derivative of equation (2.11) with respect to κ. That is, we obtain an approximate solution of $\frac{\partial Q}{\partial \kappa} = 0$, which gives an approximate update of κ for the EM algorithm.

Theorem 3.8. *The revised EM parameter estimate* $\hat{\theta}_{j+1} = \left(\mu, \kappa, \alpha, \sigma_1, \sigma_2, \rho, \lambda, H^2\right)$ *at iteration* $j+1$ *is approximated by the polynomial equation*

$$
a_9^{\kappa}\kappa^9 + a_8^{\kappa}\kappa^8 + a_7^{\kappa}\kappa^7 + a_6^{\kappa}\kappa^6 + a_5^{\kappa}\kappa^5 + a_4^{\kappa}\kappa^4 + a_3^{\kappa}\kappa^3 + a_2^{\kappa}\kappa^2 + a_1^{\kappa}\kappa + a_0^{\kappa} = 0 \qquad (3.27)
$$

with

$$a_9^\kappa = -\hat{H}_N^{(0)}\left(h_l^{1(i)}, m_l^{1(i)}\right) - \hat{K}_N\left(s_l^{1(i)}, h_l^{1(i)}\right) - \hat{K}_N\left(p_l^{1(i)}, m_l^{1(i)}\right) - \sum_{l=0}^{N} p_l^1 H^{-2} d_l^1$$

$$a_8^\kappa = \sum_{l=0}^{N} \left(p_l^1\right)' H^{-2} Y_l + \Bigg(-\Delta t^2 e_2' G^{-2} e_2 e_2' \hat{H}_N^{(2)} e_2 + \alpha \Delta t^2 e_2' G^{-2} e_2 \left(\hat{L}_N^{(1)}\right)' e_2 \Bigg)$$

$$+ \hat{K}_N\left(h_l^{1(i)}\right) - \hat{H}_N^{(0)}\left(h_l^{1(i)}\right) - \hat{H}_N^{(0)}\left(s_l^{0(i)}, m_l^{1(i)}\right) - \hat{H}_N^{(0)}\left(s_l^{1(i)}, m_l^{0(i)}\right) - \hat{K}_N\left(s_l^{0(i)}, s_l^{1(i)}\right)$$

$$- \hat{K}_N\left(s_l^{1(i)}, s_l^{0(i)}\right) - \hat{K}_N\left(p_l^{1(i)}\right) - \hat{K}_N\left(p_l^{0(i)}, m_l^{1(i)}\right) - \hat{K}_N\left(p_l^{1(i)}, m_l^{0(i)}\right)$$

$$- \sum_{l=0}^{N} p_l^0 H^{-2} d_l^1 - \sum_{l=0}^{N} p_l^1 H^{-2} s_l^0 - N\alpha^2 \Delta t^2 e_2' G^{-2} e_2$$

$$- \sum_{n=-2}^{1} \left(\sum_{l=0}^{N} \left(p_l^{-n-1}\right)' H^{-2} s_l^n \right)$$

$$a_7^\kappa = \sum_{l=0}^{N} \left(p_l^0\right)' H^{-2} Y_l - \Delta t e_2' G^{-2} \left(\hat{H}_N^{(1)} e_2 - \left(e_1 e_1' - \Delta t e_1 e_2' + e_2 e_2'\right) \hat{H}_N^{(2)} e_2 \right.$$

$$\left. - \left(\mu - \frac{1}{2}\sigma_1^2\right) \Delta t e_1 \left(\hat{L}_N^{(1)}\right)' e_2 \right)$$

$$+ \hat{K}_N\left(h_l^{0(i)}\right) - \hat{H}_N^{(0)}\left(h_l^{0(i)}\right) - \hat{H}_N^{(0)}\left(h_l^{-1(i)}, m_l^{1(i)}\right) - \hat{H}_N^{(0)}\left(h_l^{0(i)}, m_l^{0(i)}\right)$$

$$- \hat{H}_N^{(0)}\left(h_l^{1(i)}, m_l^{-1(i)}\right) - \hat{K}_N\left(s^{-1(i)}, h_l^{1(i)}\right) - \hat{K}_N\left(s_l^{0(i)}, h_l^{0(i)}\right)$$

$$- \hat{K}_N\left(s_l^{1(i)}, h_l^{-1(i)}\right) - \hat{K}_N\left(p_l^{0(i)}\right) - \hat{K}_N\left(p_l^{-1(i)}, m_l^{1(i)}\right) - \hat{K}_N\left(p_l^{0(i)}, m_l^{0(i)}\right)$$

$$- \hat{K}_N\left(p_l^{1(i)}, m_l^{-1(i)}\right) - \sum_{l=0}^{N} p_l^{-1} H^{-2} d_l^1 - \sum_{l=0}^{N} p_l^0 H^{-2} d_l^0 - \sum_{l=0}^{N} p_l^1 H^{-2} d_l^{-1}$$

$$+ \alpha \Delta t e_2' G^{-2} \left(\hat{L}_N^{(2)} - Q \hat{L}_N^{(1)} - N \left(\mu - \frac{1}{2}\sigma_1^2\right) \Delta t e_1 \right)$$

$$a_6^\kappa = \sum_{l=0}^{N} \left(p_l^{-1}\right)' H^{-2} Y_l + \hat{K}_N\left(h_l^{-1(i)}\right) - \hat{H}_N^{(0)}\left(h_l^{-1(i)}\right) - \hat{H}_N^{(0)}\left(h_l^{-2(i)}, m_l^{1(i)}\right)$$

$$- \hat{H}_N^{(0)}\left(h_l^{-1(i)}, m_l^{0(i)}\right) - \hat{H}_N^{(0)}\left(h_l^{0(i)}, m^{-1(i)}\right) - \hat{K}_N\left(s_l^{-2(i)}, h_l^{1(i)}\right) - \hat{K}_N\left(s_l^{-1(i)}, h_l^{0(i)}\right)$$

$$- \hat{K}_N\left(s_l^{0(i)}, h_l^{-1(i)}\right) - \hat{K}_N\left(s_l^{1(i)}, h_l^{-2(i)}\right) - \hat{K}_N\left(p_l^{-1(i)}\right) - \hat{K}_N\left(p_l^{-2(i)}, m_l^{1(i)}\right)$$

$$- \hat{K}_N\left(p_l^{-1(i)}, m_l^{0(i)}\right) - \hat{K}_N\left(p_l^{0(i)}, m_l^{-1(i)}\right)$$

$$a_5^{\kappa} = \sum_{l=0}^{N} \left(p_l^{-2}\right)' H^{-2} Y_l + \hat{K}_N\left(h_l^{-2(i)}\right) - \hat{H}_N^{(0)}\left(h_l^{-2(i)}\right) - \hat{H}_N^{(0)}\left(h_l^{-2(i)}, m_l^{0(i)}\right)$$

$$- \hat{H}_N^{(0)}\left(h_l^{-1(i)}, m_l^{-1(i)}\right) - \hat{K}_N\left(s_l^{-3(i)}, h^1\right) - \hat{K}_N\left(s_l^{-2(i)}, h_l^{0(i)}\right) - \hat{K}_N\left(s_l^{-1(i)}, h_l^{-1(i)}\right)$$

$$- \hat{K}_N\left(s_l^{0(i)}, h_l^{-2(i)}\right) - \hat{K}_N\left(p_l^{-2(i)}\right) - \hat{K}_N\left(p_l^{-3(i)}, m_l^{1(i)}\right)$$

$$- \hat{K}_N\left(p_l^{-2(i)}, m_l^{0(i)}\right) - \hat{K}_N\left(p_l^{-1(i)}, m_l^{-1(i)}\right) - \sum_{n=-3}^{1}\left(\sum_{l=0}^{N} p_l^{-2-n} H^{-2} d_l^{n}\right)$$

$$a_4^{\kappa} = \sum_{l=0}^{N} \left(p_l^{-3}\right)' H^{-2} Y_l - \hat{H}_N^{(0)}\left(dM_{-2}, m_{-1}\right) - \hat{K}_N\left(d_{-3}, dM_0\right) - \hat{K}_N\left(d_{-2}, dM_{-1}\right)$$

$$- \hat{K}_N\left(d_{-1}, dM_{-2}\right) - \hat{K}_N\left(p_{-3}\right) - \hat{K}_N\left(p_{-4}, m_1\right) - \hat{K}_N\left(p_{-3}, m_0\right) - \hat{K}_N\left(p_{-2}, m_{-1}\right)$$

$$- \sum_{n=-3}^{1}\left(\sum_{l=0}^{N} p_{-3-n} H^{-2} d_n\right)$$

$$a_3^{\kappa} = \sum_{l=0}^{N} \left(p_l^{-4(i)}\right)' H^{-2} Y_l - \hat{K}_N\left(s_l^{-3(i)}, h_l^{-1(i)}\right) - \hat{K}_N\left(s_l^{-2(i)}, h_l^{-2(i)}\right) - \hat{K}_N\left(p_l^{-4(i)}\right)$$

$$- \hat{K}_N\left(p_l^{-4(i)}, m_l^{0(i)}\right)$$

$$- \hat{K}_N\left(p_l^{-3(i)}, m_l^{-1(i)}\right) - \sum_{n=0}^{3}\left(\sum_{l=0}^{N} p_l^{n-4} H^{-2} d_l^{-n}\right)$$

$$a_2^{\kappa} = -\hat{K}_N\left(s_l^{-3(i)}, h_l^{-2(i)}\right) - \hat{K}_N\left(p_l^{-4(i)}, m_l^{-1(i)}\right) - \sum_{l=0}^{N} p_l^{-4(i)} H^{-2} s_l^{-1(i)} - \sum_{l=0}^{N} p_l^{-3(i)} H^{-2} s_l^{-2(i)}$$

$$- \sum_{l=0}^{N} p_l^{-2(i)} H^{-2} s_l^{-3(i)}$$

$$a_1^{\kappa} = -\sum_{l=0}^{N} p_l^{-4} H^{-2} d_l^{-2} - \sum_{l=0}^{N} p_l^{-3} H^{-2} d_l^{-3}$$

$$a_0^{\kappa} = -\sum_{l=0}^{N} p_l^{-4} H^{-2} d_l^{-3}$$

where

$$\hat{K}_N(h_l^m) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} Y_l' H^{-2}[\mathbf{0}_{M\times 1}, h_l^m] X_l | \mathcal{Y}_N\right] \tag{3.28}$$

$$m = -2, -1, 0, 1$$

$$\hat{K}_N\left(s_l^k, h_l^m\right) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} \left(s_l^k\right)' H^{-2}[\mathbf{0}_{M\times 1}, h_l^m] X_l | \mathcal{Y}_N\right] \tag{3.29}$$

$$k = -3, -2, -1, 0, 1, m = -2, -1, 0, 1$$

$$\hat{K}_N(p_l^m) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} (p_l^m)' H^{-2}[\mathbf{I}_{M\times 1}, \mathbf{0}_{M\times 1}] X_l | \mathcal{Y}_N\right] \tag{3.30}$$

$$m = -4, -3, -2, -1, 0, 1$$

$$\hat{K}_N(p_l^m, m_l^n) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} (p_l^m)' H^{-2}[\mathbf{0}_{M\times 1}, m_l^n] X_l | \mathcal{Y}_N\right] \tag{3.31}$$

$$m = -4, -3, -2, -1, 0, 1, n = -1, 0, 1$$

$$\hat{H}_N^{(0)}(h_l^m) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} (h_l^m)' H^{-2} \mathbf{I}_{M\times 1} X_l^{(2)} X_l^{(1)} | \mathcal{Y}_N\right] \tag{3.32}$$

$$m = -2, -1, 0, 1$$

$$\hat{H}_N^{(0)}(h_l^m, m_l^n) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} (h_l^m)' H^{-2} m_l^n X_l^{(2)} X_l^{(2)} | \mathcal{Y}_N\right] \tag{3.33}$$

$$m = -2, -1, 0, 1, n_l = -1, 0, 1$$

and $\mathbf{0}_{M\times 1}$ *is an* $M \times 1$ *vector with* 0 *in each position,* $\mathbf{I}_{M\times 1}$ *is an* $M \times 1$ *vector with* 1 *in each position and* $X_l^{(i)}$*,* $i = 1, 2$ *is the i-th element of process* X_l*.*

Proof. See Appendix A. □

In Elliott and Hyndman (2007), it was shown that a polynomial for the single volatility parameter has exactly one positive root. In this work, in order to update σ_1, σ_2, ρ, κ, we simply perform one iteration of Newton-Rhapson method for them in each iteration of the EM algorithm.

In the next section, we derive finite-dimensional filters necessary to calculate the conditional expectations of equation (3.10)-(3.11), (3.13), (3.16), (3.17), (3.18), (3.20), (3.21), (3.23), (3.25), (3.26), and (3.28)-(3.33).

3.2 Finite-dimensional filters

The E-step of the EM algorithm involves the computation of various conditional expectations. These conditional expectations can be obtained using the Kalman smoother which needs a forward and backward pass through the data. Alternatively, we derive finite-dimensional filters which only need a forward pass when implementing the E-step of the EM algorithm. Define

$$L_k^{(1)} = \sum_{l=1}^{k} X_{l-1}$$

$$L_k^{(2)} = \sum_{l=1}^{k} X_l$$

$$H_k^{(1)} = \sum_{l=1}^{k} X_l X_{l-1}'$$

$$H_k^{(2)} = \sum_{l=1}^{k} X_{l-1} X_{l-1}'$$

$$H_k^{(3)} = \sum_{l=1}^{k} X_l X_l$$

$$H_k^{(0)} = H_k^{ij}\left(\bar{f}_l, \bar{g}_l\right) = \sum_{l=0}^{k} \bar{f}_l' H^{-2} \bar{g}_l X_l^{(i)} X_l^{(j)}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, m$$

$$K_k\left(f_l, g_l\right) = \sum_{l=0}^{k} f_l' H^{-2} g_l X_l$$

$$J_k = \sum_{l=0}^{k} M_l X_l \left(Y_l - d_l\right)'$$

$$U_k = \sum_{l=0}^{k} M_l X_l X_l' M_l'$$

where $f_l \in \mathcal{R}^M$, $\overline{f_l} \in \mathcal{R}^M$, $Y_l \in \mathcal{R}^M$,$d_l \in \mathcal{R}^M$, $X_l \in R^m$ and g_l and $\overline{g_l}$ are both $M \times m$ matrices.

For $i, j \in 1, \ldots, m$, we have the scalar processes

$$
\begin{aligned}
L_k^{i(1)} &= \sum_{l=1}^{k} X_{l-1}^{(i)} \\
L_k^{i(2)} &= \sum_{l=1}^{k} X_l^{(i)} \\
H_k^{ij(1)} &= \sum_{l=1}^{k} X_l^{(i)} X_{l-1}^{(j)} \\
H_k^{ij(2)} &= \sum_{l=1}^{k} X_{l-1}^{(i)} X_{l-1}^{(j)} \\
H_k^{ij(3)} &= \sum_{l=1}^{k} X_l^{(i)} X_l^{(j)} \\
J_k^{ij} &= \sum_{l=0}^{k} M_l^{(i)} X_l \left(Y_l^{(j)} - d_l^{(j)} \right) \\
&= \sum_{l=0}^{k} M^{(i,1)} X_l^{(1)} \left(Y_l^{(j)} - d_l^{(j)} \right) + \sum_{l=0}^{k} M^{(i,2)} X_l^{(2)} \left(Y_l^{(j)} - d_l^{(j)} \right) \\
U_k^{ij} &= \sum_{l=0}^{k} M_l^{(i)} X_l X_l' M_l^{(j)'} \\
&= \sum_{l=0}^{k} M_l^{(i,1)} X_l^{(1)} X_l^{(1)} M_l^{(j,1)} + \sum_{l=0}^{k} M_l^{(i,1)} X_l^{(1)} X_l^{(2)} M_l^{(j,2)} \\
&\quad + \sum_{l=0}^{k} M_l^{(i,2)} X_l^{(2)} X_l^{(1)} M_l^{(j,1)} + \sum_{l=0}^{k} M_l^{(i,2)} X_l^{(2)} X_l^{(2)} M_l^{(j,2)}
\end{aligned}
$$

where $M_l^{(m,n)}$ represents the element in the m-th row and n-th column of M_l, $Y_l^{(n)}$ is the n-th element of Y_l and $d_l^{(n)}$ is the n-th element of d_l. Let e_n be the unit column vector with 1 in the

n-th position. We can rewrite $L_k^{i(n)}, n = 1, 2$, $H_k^{ij(n)}, n = 1, 2, 3$, J_k^{ij} and U_k^{ij} as

$$
\begin{aligned}
L_k^{i(1)} &= \sum_{l=1}^{k} < X_{l-1}, e_i > \\
L_k^{i(2)} &= \sum_{l=1}^{k} < X_l, e_i > \\
H_k^{ij(1)} &= \sum_{l=1}^{k} < X_l, e_i >< X_{l-1}, e_j > \\
H_k^{ij(2)} &= \sum_{l=1}^{k} < X_{l-1}, e_i >< X_{l-1}, e_j > \\
H_k^{ij(3)} &= \sum_{l=1}^{k} < X_l, e_i >< X_l, e_j > \\
J_k^{ij} &= \sum_{l=0}^{k} < X_l, e_1 M_l^{(i,1)} \left(Y_l^{(j)} - d_l^{(j)}\right) > + \sum_{l=0}^{k} < X_l, e_2 M_l^{(i,2)} \left(Y_l^{(j)} - d_l^{(j)}\right) > \\
U_k^{ij} &= \sum_{l=0}^{k} < X_l, e_1 M_l^{(i,1)} >< X_l, e_1 M_l^{(j,1)} > + \sum_{l=0}^{k} < X_l, e_1 M_l^{(i,1)} >< X_l, e_2 M_l^{(j,2)} > \\
&\quad + \sum_{l=0}^{k} < X_l, e_2 M_l^{(i,2)} >< X_l, e_1 M_l^{(j,1)} > + \sum_{l=0}^{k} < X_l, e_2 M_l^{(i,2)} >< X_l, e_2 M_l^{(j,2)} >
\end{aligned}
$$

Denote

$$
\begin{aligned}
J_k^{ij(1)} &= \sum_{l=0}^{k} < X_l, e_1 M_l^{(i,1)} \left(Y_l^{(j)} - d_l^{(j)}\right) > \\
J_k^{ij(2)} &= \sum_{l=0}^{k} < X_l, e_2 M_l^{(i,2)} \left(Y_l^{(j)} - d_l^{(j)}\right) > \\
U_k^{ij(1)} &= \sum_{l=0}^{k} < X_l, e_1 M_l^{(i,1)} >< X_l, e_1 M_l^{(j,1)} > \\
U_k^{ij(2)} &= \sum_{l=0}^{k} < X_l, e_1 M_l^{(i,1)} >< X_l, e_2 M_l^{(j,2)} > \\
U_k^{ij(3)} &= \sum_{l=0}^{k} < X_l, e_2 M_l^{(i,2)} >< X_l, e_1 M_l^{(j,1)} > \\
U_k^{ij(4)} &= \sum_{l=0}^{k} < X_l, e_2 M_l^{(i,2)} >< X_l, e_2 M_l^{(j,2)} > .
\end{aligned}
$$

Theorem 3.9. *Finite-dimensional filters for* $\hat{L}_k^{i(n)}, n = 1, \ldots, m$, K_k, $\hat{H}_k^{ij(n)}, n = 0, 1, 2, 3$, $J_k^{ij(m)}, m =$

$1,2$ *and* $U_k^{ij(n)}, n = 1,2,3,4$ *are given by*

$$\hat{L}_k^{i(n)} = E_{\hat{\theta}_j}\left[L_k^{i(n)}|\mathcal{Y}_k\right] = u_k^{i(n)} + \left(v_k^{i(n)}\right)^{'}\mu_k, \tag{3.34}$$

$$\hat{H}_k^{ij(n)} = E_{\hat{\theta}_j}\left[H_k^{ij(n)}|\mathcal{Y}_k\right] = \bar{u}_k^{ij(n)} + \left(\bar{v}_k^{i(n)}\right)'\mu_k + Tr\left[\bar{d}_k^{ij(n)}P_k\right] + \mu_k^{'}\bar{d}_k^{ij(n)}\mu_k, \tag{3.35}$$

$$\hat{K}_k = E_{\hat{\theta}_j}\left[K_k|\mathcal{Y}_k\right] = \bar{r}_k + \bar{s}_k^{'}\mu_k \tag{3.36}$$

$$\hat{J}_k^{ij(m)} = E_{\hat{\theta}_j}\left[J_k^{ij(m)}|\mathcal{Y}_k\right] = \bar{a}_k^{ij(m)} + \bar{b}_k^{ij(m)'}\mu_k \tag{3.37}$$

$$\hat{U}_k^{ij(n)} = E_{\hat{\theta}_j}\left[U_k^{ij(n)}|\mathcal{Y}_k\right] = a_k^{ij(n)} + b_k^{ij(n)'}\mu_k + Tr\left[d_k^{ij(n)}P_k^{-1}\right] + \mu_k^{'}d_k^{ij(n)}\mu_k \tag{3.38}$$

where μ_k *and* P_k *are the conditional mean and covariance of the Kalman filter,respectively. The recursions for sufficient statistics* $a,b,d,\bar{a},\bar{b},\bar{r},\bar{s},u,v,\bar{u},\bar{v},\bar{d}$ *of finite-dimensional filters are shown and proved in Appendix B.*

Proof. See Appendix B. □

Chapter 4

Application to Simulated Data

This chapter provides the results from a numerical implementation of the filter-based EM algorithm using simulated data.

From equation (3.7) and (3.8), we generated log-spot prices , convenience yields, and log-futures prices with disturbances. We assumed that futures prices were observed approximately weekly, $\Delta t = \frac{1}{48}$, at maturities of 1, 3, 6, 9 and 12 months for 10 years, giving 480 observations. We assumed that matrix H was diagonal.

We then applied the filter-based EM algorithm to estimate the parameters of the empirical model. We simulated the data using true parameter values, obtain the maximum likelihood parameter estimates by direct maximization of the likelihood function and the estimates by the filter-based EM algorithm which are presented in Table 4.1.

The procedure used to find the MLE directly was MATLAB's *fminsearch* routine (Nelder-Mead simplex direct search). In both of the direct MLE approach and the filter-based EM approach, the stoping criteria used to stop the algorithm if the maximum of the absolute differences between the parameters from one step to the next was less than 10^{-10}. In the

Table 4.1: True parameter values, MLEs calculated by direct maximization, and MLEs calculated via filter-based EM algorithm for simulated data

parameter	True value	direct MLE	filter-based EM
μ	0.14	0.126058079726224	0.104211787509002
κ	1.8	1.614571441144134	1.422333684869484
α	0.12	0.149930278665568	0.149771455353883
σ_1	0.4	0.431573239773677	0.391350599321044
σ_2	0.53	0.660039947685030	0.535489394505664
ρ	0.77	0.900933854538372	0.864687895637105
λ	0.2	0.241878708046353	0.197791958943542
H_{11}^2	0.25	0.245843333800211	0.239020495456221
H_{22}^2	0.25	0.240751824894155	0.235102064525230
H_{33}^2	0.25	0.235414924742002	0.230963742472480
H_{44}^2	0.25	0.258241063058967	0.253469948351176
H_{55}^2	0.25	0.236981492986396	0.229651394525523
log-likelihood		-1759.690802091415	-1760.198517611300
iterations		2262	1097

filter-based EM approach, the algorithm was stopped if

$$max\left(|\hat{\mu}_k-\hat{\mu}_{k-1},|\hat{\kappa}_k-\hat{\kappa}_{k-1}|,|\hat{\alpha}_k-\hat{\alpha}_{k-1}|,|\hat{\sigma}_{1k}-\hat{\sigma}_{1k-1}|,|\hat{\sigma}_{2k}-\hat{\sigma}_{2k-1}|,|\hat{\rho}_k-\hat{\rho}_{k-1}|,\right.$$
$$\left.|\hat{\lambda}_k-\hat{\lambda}_{k-1}|,|\left(\hat{H}_k^2\right)^{(11)}-\left(\hat{H}_{k-1}^2\right)^{(11)}|,\ldots,|\left(\hat{H}_k^2\right)^{(55)}-\left(\hat{H}_{k-1}^2\right)^{(55)}|\right)<10^{-10}$$

We took the true parameter values by referring to the result of the oil market for the two-factor model in Schwartz (1997). Particularly, we referred to the result from 10-year market data. This was done in order to make the true parameter value sensible rather than taking random number. It is worth observing that the elements of H^2 can not be too small since the implementation involves the inverse, H^{-2}. We examine the likelihood function value for the EM algorithm to see if the the value of likelihood function is increasing after each iteration of the EM algorithm. To do so, we update only one parameter at a time (holding the others fixed). This is called a Generalized EM algorithm.

We found that for some starting values, using the EM algorithm, the log-likelihood function value tended to decrease after some iterations. For example, when updating μ only using starting value 0.22, the value of log-likelihood function decreased after the second iteration (See Figure 4.1).

Similar problems occurred when updating other parameters. It happened because the precision of the software we use is not high enough to distinguish between two numbers with decimals more than the software can store. In order to avoid this, we set the criteria which stopped the iteration of the EM algorithm if the new likelihood function is smaller then the previous and moved on to the next update algorithm.

The spreads between MLE estimates and EM estimates for σ_1, σ_2, ρ and κ are expected since for those parameters, we use one iteration of Newton-Rhapson method to solve for the polynomials in each iteration of the EM algorithm. The nature of Newton-Rhapson method causes

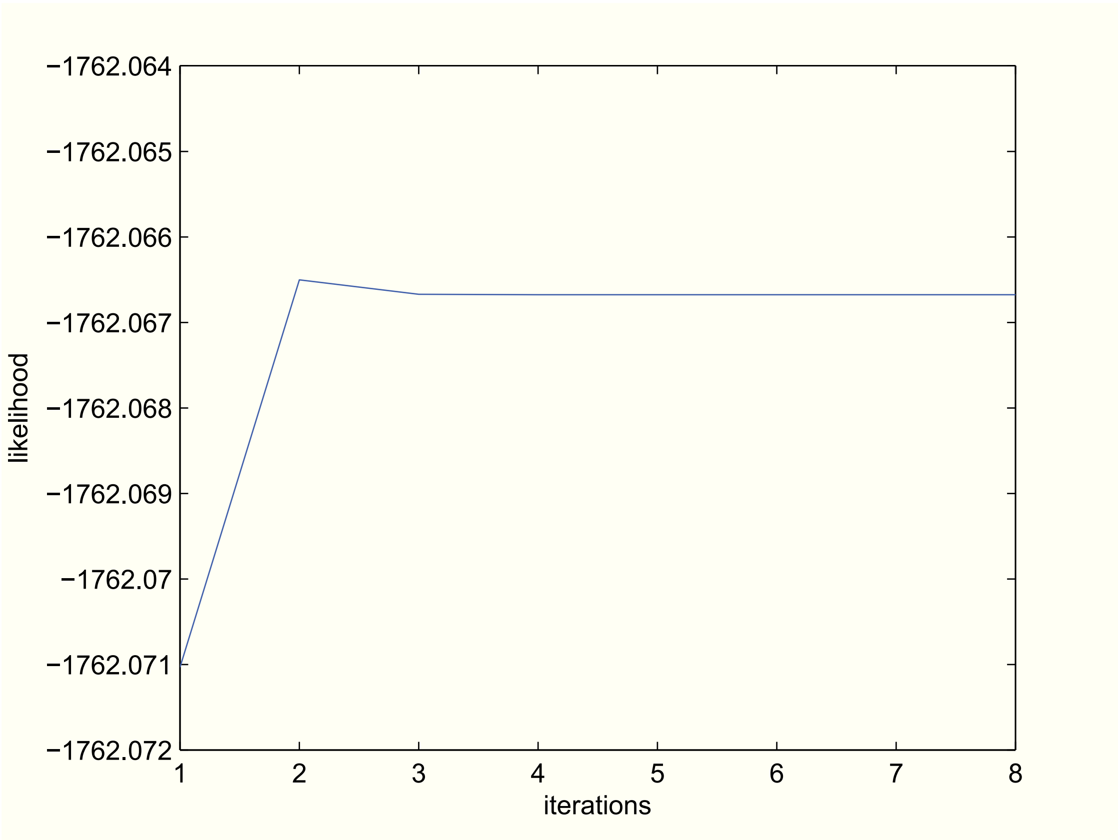

Figure 4.1: log-likelihood function value when updating μ at a starting point 0.22 while holding other parameter values fixed.

inaccuracy in each iteration of the EM update for the those parameters by only obtaining an approximate solution for the polynomials in Theorem 3.4-3.8. Another reason for the spreads is the precision of the software we used. Since MATLAB can only store numbers with limited decimals rounding errors are expected in the implementation. Particularly for the parameter κ, the polynomial in Theorem 3.8 is an approximation of the equation that κ should satisfy. That is to say, including the Newton-Rhapson method to solve for the polynomial, we have two levels of approximation when updating the parameter κ. We also should point out that even though, the two estimates (direct and EM) for each parameter are not exactly equal, the log-likelihood function value are relatively close. This may be due to the flatness of the log-likelihood function around the MLE estimate values for each parameter, numerical round-off error, and the well-known slow convergence property of the EM algorithm. We use σ_1 as a example to explain. From Figure 4.2, we can see that the likelihood is quite flat between

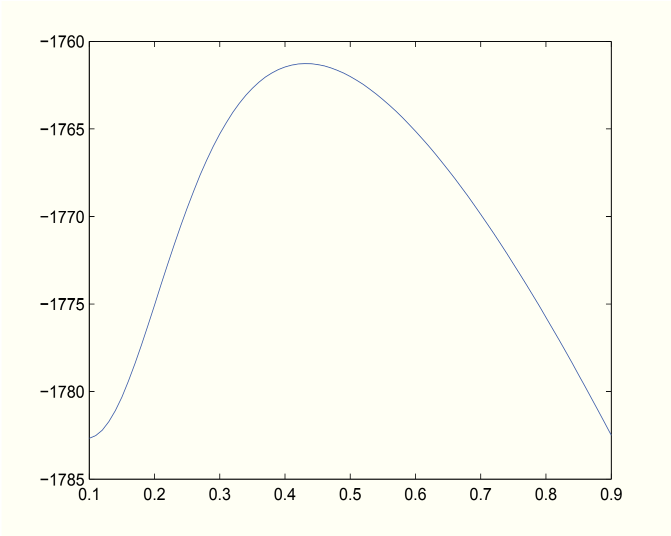

Figure 4.2: Log-likelihood function value about σ_1 when holding other parameter values fixed.

0.3 and 0.4. Wu (1983) proved that under suitable conditions, the sequence of parameter estimates generated by the EM algorithm converges to the MLE estimate. We can see in the Figure 4 that parameter estimates of EM algorithm converged very fast at the beginning and then it went slow but closer to the MLE estimates for each parameter.

The Kalman filter estimates of the log-spot price using the EM parameter values fit the simulated data fairly well as shown in left plot in Figure 4.4. However, the Kalman filter estimates of the convenience yield using the EM parameter estimates do not track the simulated convenience yield as well. One possible reason is the difficulty of estimating the mean reversion parameter κ. We also examed how the EM estimate performs in pricing futures contracts compared to the futures prices simulated with true parameter value. In Figure 4.5, the differ-

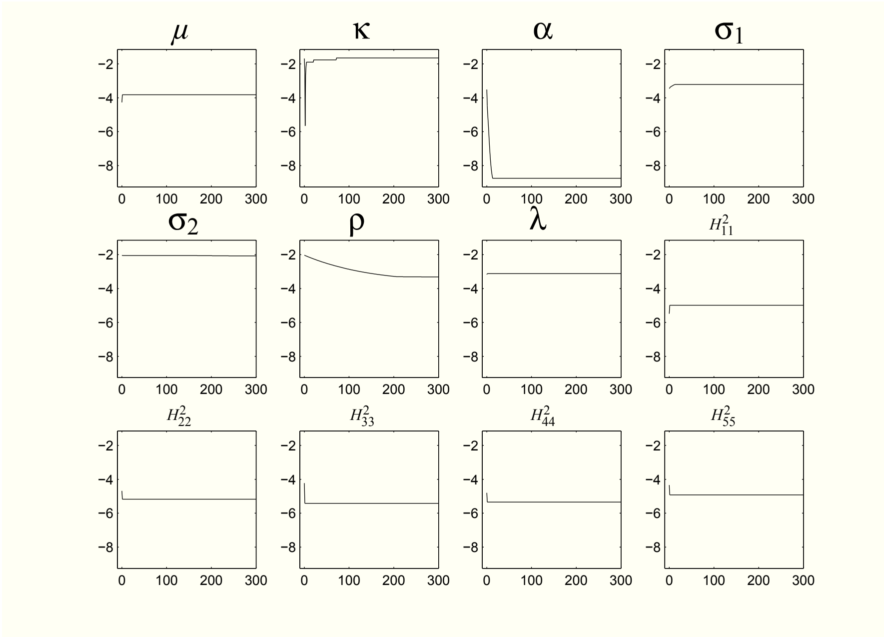

Figure 4.3: Logarithm of absolute errors of the j-th iteration of EM algorithm and MLE estimates for $\mu, \kappa, \alpha, \sigma_1, \sigma_2, \rho, \lambda, H^2_{11}, H^2_{22}, H^2_{33}, H^2_{44}, H^2_{55}$, from left to right and from top to bottom.

ence in futures prices between the two sets of the parameter value is small in a short term and gets bigger as time passes by. That is to say, the EM estimate prices futures contracts better in a short term than in a long term.

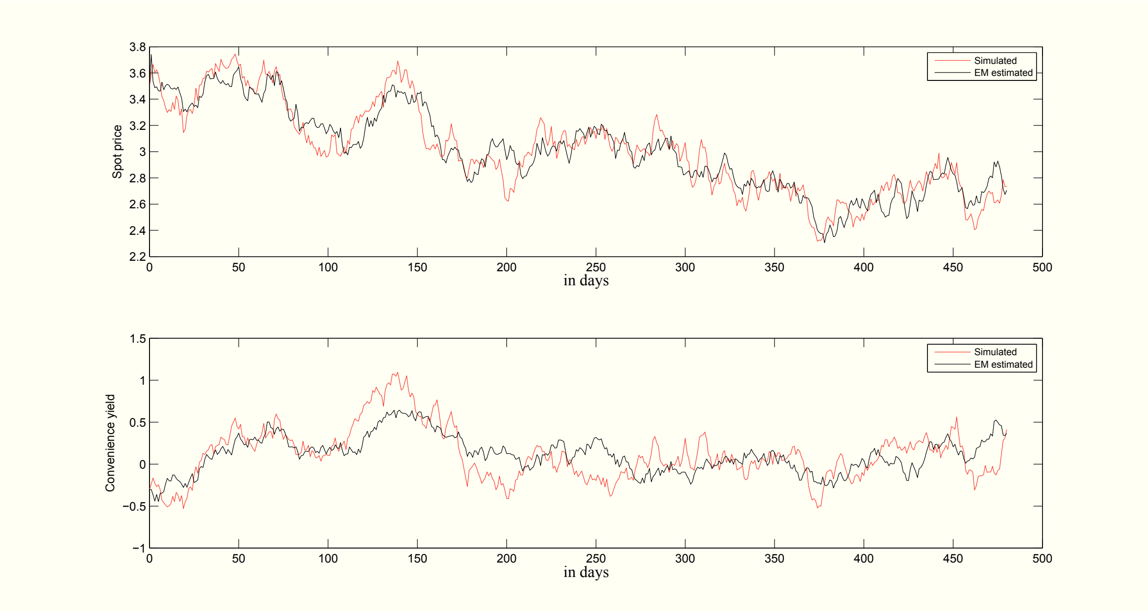

Figure 4.4: spot price and convenience yield with true parameter value and EM estimate

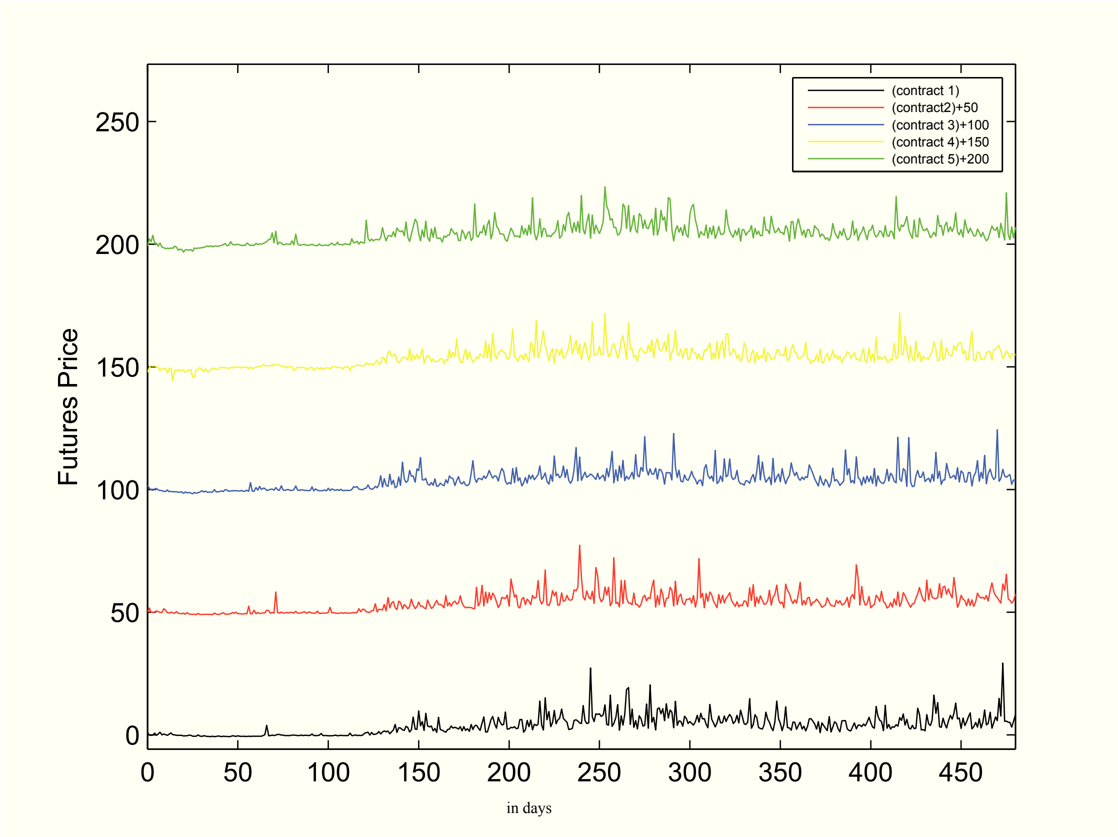

Figure 4.5: Futures prices difference between true parameter value and EM estimate for 5 futures contracts

Chapter 5

Conclusion and future work

5.1 Conclusion

We find the filtering approach appealing method by which we can estimate unobservable signals using the observations. We extended the work of Elliott and Krishnamurthy (1999) and Elliott and Hyndman (2007) to derive a filter-based EM algorithm to estimate the parameters in the Schwartz (1997) two-factor model. In general, other commodity price models and general Gaussian state-space models can be considered using a similar procedure to estimate the model parameters.

Elliott and Krishnamurthy (1999) outlined the advantages of the filter-based EM algorithm over the standard smoother-based EM algorithm which should be applicable here. First, the memory costs are significantly reduced compared to a smoother-based algorithm. Secondly, the filter-based algorithm is well suited to parallel implementation on a multi-processor system because the filters are decoupled. Moreover, a filter-based algorithm will be at least twice as fast as existing smoother-based EM-algorithms since only a forward pass is required for the filter-based algorithm. In addition, steady state properties of the Kalman filter could speed up a filter-based EM algorithm which can't done with a smoother-based algorithm.

5.2 Future work

Considering the problems we came with during this work, implementing the algorithm with higher numerical precision is a way to reduce the round-off errors introduced many during the numerical calculation required for the implementation. Since we used approximation for the updates of EM algorithm at each iteration, we can think about a better numerical procedure to approximately solve the polynomials in Theorem 3.1-3.8. Also, in the first approximation for updating κ, we could approximate $e^{-\kappa_l \tau_l}$ and $e^{-2\kappa_l \tau_l}$ using more precise Taylor series expansions by adding the third or higher terms in the Taylor series of $e^{-\kappa_l \tau_l}$ and $e^{-2\kappa_l \tau_l}$. It would be interesting to apply the procedure to different commodities in different markets and also to different models. Particularly, it is interesting to investigate how the filter-based algorithm approach works with higher-dimensional models.

Appendix A

Proof of Theorem 3.1-3.8

In this appendix, we give the detailed proofs of the expressions for the M-step parameter updates given in Theorems 3.1- 3.8.

In the E-step of EM algorithm, we have function $Q\left(\hat{\theta}_j, \theta\right)$

$$\begin{aligned} Q\left(\hat{\theta}_j, \theta\right) = & -\sum_{l=1}^{N} \ln |G_l| - \sum_{l=0}^{N} \ln |H| + E_{\hat{\theta}_j}\left[R\left(\hat{\theta}_j\right) |Y_N\right] \\ & -\frac{1}{2} E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} \left(X_l - Q_l X_{l-1} - c_l\right)' G^{-2} \left(X_l - Q_l X_{l-1} - c_l\right) |Y_N\right] \\ & -\frac{1}{2} E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} \left(Y_l - M_l X_l - d_l\right)' H^{-2} \left(Y_l - M_l X_l - d_l\right) |Y_N\right] \end{aligned}$$

The main idea is to take first derivative with respect to each parameter then set them to zero. We consider each parameter separately.

A.1 Proof of Theorem 3.1

Taking the derivative of Q equation (2.11) with respect to μ obtain

$$\begin{aligned}
\frac{\partial Q\left(\hat{\theta}_j, \theta\right)}{\partial \mu} &= -E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(-\frac{\partial c_l}{\partial \mu}\right)' G^{-2}\left(X_l - Q_l X_{l-1} - c_l\right)|Y_N\right] \\
&= E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} \Delta t e_1' G^{-2}\left(X_l - Q_l X_{l-1} - c_l\right)|Y_N\right] \\
&= \Delta t e_1{}' G^{-2} E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(X_l - Q_l X_{l-1} - \mu \Delta t e_1 + \frac{1}{2}\sigma_1^2 \Delta t e_1 - \kappa\alpha\Delta t e_2\right)|Y_N\right] \qquad \text{(A.1)}
\end{aligned}$$

Set equation (A.1) to zero to find

$$\Delta t e_1' G^{-2}\left(E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(X_l - Q_l X_{l-1}\right)|Y_N\right] - N\mu\Delta t e_1 + \frac{1}{2}\sigma_1^2 N \Delta t e_1 - \kappa\alpha\Delta t N e_2\right) = 0 \qquad \text{(A.2)}$$

Simplify Equation (A.2) and obtain Equation (3.9)

$$\mu = \frac{e_1' G^{-2}\left(\hat{L}_N^{(2)} - Q_l \hat{L}_N^{(1)} + \frac{1}{2}\sigma_1^2 N \Delta t e_1 - \kappa\alpha\Delta t N e_2\right)}{N \Delta t e_1' G^{-2} e_1}$$

where $\hat{L}_N^{(1)}$ and $\hat{L}_N^{(2)}$ are defined as in equation (3.10) and (3.11).

A.2 Proof of Theorem 3.2

By taking the derivative of $Q\left(\hat{\theta}_j, \theta\right)$ with respect to λ, we obtain

$$\begin{aligned}
\frac{\partial Q\left(\hat{\theta}_j, \theta\right)}{\partial \lambda} &= -E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(-\frac{\partial d_l}{\partial \lambda}\right)' H^{-2}\left(Y_l - M_l X_l - d_l\right)|Y_N\right] \\
&= E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(A_l^{\lambda}\right)' H^{-2}\left(Y_l - M_l X_l - \lambda A_l^{\lambda} - B_l^{\lambda}\right)|Y_N\right] \\
&= E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(A_l^{\lambda}\right)' H^{-2}\left(Y_l - B_l^{\lambda}\right) - \left(A_l^{\lambda}\right)' H^{-2} M_l X_l - \left(A_l^{\lambda}\right)' H^{-2} A_l^{\lambda}|Y_N\right] \\
&= \sum_{l=0}^{N}\left(A_l^{\lambda}\right)' H^{-2}\left(Y_l - B_l^{\lambda}\right) - \hat{K}_N\left(A_l^{\lambda}\right) - \lambda\sum_{l=0}^{N}\left(A_l^{\lambda}\right)' H^{-2} A_l^{\lambda} \qquad \text{(A.3)}
\end{aligned}$$

Set equation (A.3) to zero and simplify it. Then we can obtain Equation (3.12)

$$\lambda = \left[\sum_{l=0}^{N} \left(A_l^{\lambda}\right)' H^{-2} A_l^{\lambda}\right]^{-1} \times \left[\sum_{l=0}^{N} \left(A_l^{\lambda}\right)' H^{-2} \left(Y_l - B_l^{\lambda}\right) - \hat{K}_N \left(A_l^{\lambda}\right)\right]$$

where $\hat{K}_N \left(A_l^{\lambda}\right)$ are defined as in equation (**??**).

A.3 Proof of Theorem 3.3

Similarly with the previous two proofs, we take the derivative of $Q\left(\hat{\theta}_j, \theta\right)$ with respect to α

$$\begin{aligned}
\frac{\partial Q\left(\hat{\theta}_j, \theta\right)}{\partial \alpha} =& E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} \left(\frac{\partial c_l}{\partial \alpha}\right)' G^{-2} \left(X_l - Q_l X_{l-1} - c_l\right) | Y_N\right] \\
&+ E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} \left(\frac{\partial d_l}{\partial \alpha}\right)' H^{-2} \left(Y_l - M_l X_l - d_l\right) | Y_N\right] \\
=& E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} \kappa \Delta t e_2' G^{-2} \left(X_l - Q_l X_{l-1} - c_l\right) | Y_N\right] \\
&+ E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} (A_l^{\alpha})' H^{-2} \left(Y_l - M_l X_l - \alpha A_l^{\alpha} - B_l^{\alpha}\right) | Y_N\right] \\
=& \kappa \Delta t e_2' G^{-2} \left(\hat{L}_N^{(2)} - Q_l \hat{L}_N^{(1)} - \left(\mu - \frac{1}{2}\sigma_1^2\right) \Delta t e_1 N - \kappa \alpha \Delta t N e_2\right) \\
&+ \sum_{l=0}^{N} (A_l^{\alpha})' H^{-2} \left(Y_l - B_l^{\alpha}\right) - \hat{K}_N \left(A_l^{\alpha}\right) - \alpha \sum_{l=0}^{N} (A_l^{\alpha})' H^{-2} A_l^{\alpha} \qquad (A.4)
\end{aligned}$$

Setting Equation (A.4) to zero and simplifying it gets us Equation (3.14)

$$\alpha = \frac{\kappa \Delta t e_2' G^{-2} \left(\hat{L}_N^{(2)} - Q_l \hat{L}_N^{(1)} - \left(\mu - \frac{1}{2}\sigma_1^2\right) \Delta t e_1 N\right) + \sum_{l=0}^{N} \left(A_l^{\alpha}\right)' H^{-2} \left(Y_l - B_l^{\alpha}\right) - \hat{K}_N \left(A_l^{\alpha}\right)}{\kappa \Delta t e_2' G^{-2} \kappa \Delta t N e_2 + \sum_{l=0}^{N} \left(A_l^{\alpha}\right)' H^{-2} A_l^{\alpha}}$$

A.4 Proof of Theorem 3.4

First, take the derivative of $Q\left(\hat{\theta}_j, \theta\right)$ with respect to σ_1

$$\begin{aligned}
\frac{\partial Q\left(\hat{\theta}_j, \theta\right)}{\partial \sigma_1} = & -\frac{N}{\sigma_1} - \frac{1}{2} E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} \left(X_l - Q_l X_{l-1} - c_l\right)' \frac{\partial G^{-2}}{\partial \sigma_1} \left(X_l - Q_l X_{l-1} - c_l\right) | Y_N\right] \\
& + E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} \left(\frac{\partial c_l}{\partial \sigma_1}\right)' G^{-2} \left(X_l - Q_l X_{l-1} - c_l\right) | Y_N\right] \\
& + E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} \left(\frac{\partial d_l}{\partial \sigma_1}\right)' H^{-2} \left(Y_l - M_l X_l - d_l\right) | Y_N\right] \\
& = -\frac{N}{\sigma_1} - \frac{1}{2} E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} \left(X_l - Q_l X_{l-1} - c_l\right)' \begin{bmatrix} -\frac{2}{\sigma_1^3 \Delta t \left(1-\rho^2\right)} & \frac{\rho}{\sigma_1^2 \sigma_2 \Delta t \left(1-\rho^2\right)} \\ \frac{\rho}{\sigma_1^2 \sigma_2 \Delta t \left(1-\rho^2\right)} & 0 \end{bmatrix}\right. \\
& \times \left(X_l - Q_l X_{l-1} - c_l\right) | Y_N] \\
& + E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} e_1^{'} \left(-\sigma_1 \Delta t\right) G^{-2} \left(X_l - Q_l X_{l-1} - \Delta t \mu e_1 - \Delta t \kappa \alpha e_2 + \frac{1}{2} \sigma_1^2 \Delta e_1\right) | Y_N\right] \\
& + E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} \left(A_l^{\sigma_1}\right)' H^{-2} \left(Y_l - M_l X_l - B_l^{\sigma_1} - \sigma_1 A_1^{\sigma_1}\right) | Y_N\right] \qquad \text{(A.5)}
\end{aligned}$$

Now we define two new processes V_l and Z_l

$$V_l = e_1^{'} X_l - \left(e_1^{'} - \Delta t e_2^{'}\right) X_{l-1} - \mu \Delta t + \frac{1}{2} \sigma_1^2 \Delta t$$

$$Z_l = e_2^{'} X_l - \left(1 - \kappa \Delta t\right) e_2^{'} X_{l-1} - \kappa \alpha \Delta t$$

So $(X_l - M_l X_{l-1} - c_l)$ can be written as $[V_l, Z_l]'$. The second term of the Equation (A.5) can be written as

$$-\frac{1}{2}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}(X_l - Q_l X_{l-1} - c_l)'\begin{bmatrix} -\frac{2}{\sigma_1^3\Delta t(1-\rho^2)} & \frac{\rho}{\sigma_1^2\sigma_2\Delta t(1-\rho^2)} \\ \frac{\rho}{\sigma_1^2\sigma_2\Delta t(1-\rho^2)} & 0 \end{bmatrix}\times(X_l - Q_l X_{l-1} - c_l)\,|Y_N\right]$$

$$= -\frac{1}{2(1-\rho^2)\Delta t}E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\begin{bmatrix} V \\ Z \end{bmatrix}'\begin{bmatrix} -\frac{2}{\sigma_1^3} & \frac{\rho}{\sigma_1^2\sigma_2} \\ \frac{\rho}{\sigma_1^2\sigma_2} & 0 \end{bmatrix}\begin{bmatrix} V \\ Z \end{bmatrix}|Y_N\right]$$

$$= -\frac{1}{2(1-\rho^2)\Delta t}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\begin{bmatrix} -\frac{2}{\sigma_1^3}V + \frac{\rho}{\sigma_1^2\sigma_2}Z \\ \frac{\rho}{\sigma_1^2\sigma_2}V \end{bmatrix}'\begin{bmatrix} V \\ Z \end{bmatrix}|Y_N\right]$$

$$= \frac{1}{\sigma_1^3(1-\rho^2)\Delta t}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(e_1'X_l - \left(e_1' - \Delta t e_2'\right)X_{l-1} - \mu\Delta t + \frac{1}{2}\sigma_1^2\Delta t\right)^2|Y_N\right]$$

$$- \frac{\rho}{\sigma_1^2\sigma_2(1-\rho^2)\Delta t}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(e_1'X_l - \left(e_1' - \Delta t e_2'\right)X_{l-1} - \mu\Delta t + \frac{1}{2}\sigma_1^2\Delta t\right)\right.$$

$$\left.\times\left(e_2'X_l - (1-\kappa\Delta t)e_2'X_{l-1} - \kappa\alpha\Delta t\right)|Y_N\right]$$

$$= \frac{1}{\sigma_1^3(1-\rho^2)\Delta t}\left(e_1'\hat{H}_N^{(3)}e_1 - 2\hat{H}_N^{(1)}(e_1 - \Delta t e_2) + \left(e_1' - \Delta t e_2'\right)\hat{H}_N^{(2)}(e_1 - \Delta t e_2)\right.$$

$$-2\mu\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)}\right) + \mu^2\Delta t^2 N + \frac{1}{4}\sigma_1^4\Delta t^2 N$$

$$\left.+\sigma_1^2\Delta t\left(e_1^1\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - N\mu\Delta t\right)\right)$$

$$- \frac{\rho}{\sigma_1^2\sigma_2(1-\rho^2)\Delta t}\left(e_2'\hat{H}_N^{(3)}e_1 - e_2'\hat{H}_N^{(1)}(e_1 - \Delta t e_2) - (1-\kappa\Delta t)e_1'\hat{H}_N^{(1)}e_2\right.$$

$$+(1-\kappa\Delta t)H_N^{(2)}(e_1 - \Delta t e_2) - \mu\Delta t\left(e_2'\hat{L}_N^{(2)} - e_2'(1-\kappa\Delta t)\hat{L}_N^{(1)}\right)$$

$$\left.-\kappa\Delta t\alpha\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - N\mu\Delta\right) + \frac{1}{2}\sigma_1^2\Delta t\left(e_2'\hat{L}_N^{(2)} - (1-\kappa\Delta t)e_2'\hat{L}_N^{(1)} - \kappa\alpha\Delta t N\right)\right)$$

$$
\begin{aligned}
&\frac{\partial Q\left(\hat{\theta}_j,\theta\right)}{\partial\sigma_1} = \\
&-\frac{N}{\sigma_1}+\frac{1}{\sigma_1^3\left(1-\rho^2\right)\Delta t}\left(e_1'\hat{H}_N^{(3)}e_1-2e_1'\hat{H}_N^{(1)}\left(e_1-\Delta t e_2\right)\right. \\
&+\left(e_1'-\Delta t e_2'\right)\hat{H}_N^{(2)}\left(e_1-\Delta t e_2\right)+\mu^2\Delta t^2 N-2\mu\Delta t\left(e_1'\hat{L}_N^{(2)}-\left(e_1'-\Delta t e_2'\right)'\hat{L}_N^{(1)}\right)\Bigg)+\frac{\sigma_1\Delta t N}{4\left(1-\rho^2\right)} \\
&+\frac{e_1'\hat{L}_N^{(2)}-\left(e_1'-\Delta t e_2'\right)\hat{L}_N^{(1)}-N\mu\Delta t}{\sigma_1\left(1-\rho^2\right)}-\frac{\rho}{\sigma_1^2\sigma_2\left(1-\rho^2\right)\Delta t}\left(e_2'\hat{H}_N^{(3)}e_1\right. \\
&-e_2'\hat{H}_N^{(1)}\left(e_1-\Delta t e_2\right)-\left(1-\kappa\Delta t\right)e_1'\hat{H}_N^{(1)}e_2+\left(1-\kappa\Delta t\right)e_2'H_N^{(2)}\left(e_1-\Delta t e_2\right) \\
&-\mu\Delta t\left(e_2'\hat{L}_N^{(2)}-\left(1-\kappa\Delta t\right)\hat{L}_N^{(1)}\right)-\kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)}-\left(e_1'-\Delta t e_2'\right)\hat{L}_N^{(1)}-\mu\Delta t N\right)\Bigg) \\
&-\frac{\rho}{2\sigma_2\left(1-\rho^2\right)}\times\left(e_2'\hat{L}_N^{(2)}-\left(1-\kappa\Delta t\right)e_2'\hat{L}_N^{(1)}-\kappa\alpha\Delta t N\right) \\
&-\sigma_1\Delta t e_1' G^{-2}\left(\hat{L}_N^{(2)}-Q_l\hat{L}_N^{(1)}-N\Delta t\mu e_1-N\Delta t\kappa\alpha e_2\right)-\frac{1}{2}\sigma_1^3\Delta t^2 e_1' G^{-2}e_1 N \\
&+\sigma_1\sum_{l=0}^{N}\left(A_l^{\sigma_1}\right)'H^{-2}A_l^{\sigma_1}+\sum_{l=0}^{N}\left(A_l^{\sigma_1}\right)'H^{-2}\left(Y_l-B_l^{\sigma_1}\right)-\hat{K}_N\left(A_l^{\sigma_1}\right)
\end{aligned}
$$

Setting equation (3.4) to zero and multiplying by σ_1^3 gives

$$
a_4^{\sigma_1}\sigma_1^4+a_3^{\sigma_1}\sigma_1^3+a_2^{\sigma_1}\sigma_1^2+a_1^{\sigma_1}\sigma_1+a_0^{\sigma_1}=0
$$

with

$$
\begin{aligned}
a_4^{\sigma_1} =& -\sum_{l=0}^{N}\left(A_l^{\sigma_1}\right)' H^{-2} A_l^{\sigma_1} + \frac{N\Delta t}{4\left(1-\rho^2\right)} \\
a_3^{\sigma_1} =& -\frac{\rho}{2\sigma_2\left(1-\rho^2\right)} \times \left(e_2'\hat{L}_N^{(2)} - (1-\kappa\Delta t)\, e_2'\hat{L}_N^{(1)} - \kappa\alpha\Delta t N\right) \\
& + \sum_{l=0}^{N}\left(A_l^{\sigma_1}\right)' H^{-2}\left(Y_l - B_l^{\sigma_1}\right) - \hat{K}_N\left(A_l^{\sigma_1}\right) \\
a_2^{\sigma_1} =& -N + \frac{e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - N\mu\Delta t}{\left(1-\rho^2\right)} \\
a_1^{\sigma_1} =& -\frac{\rho}{\sigma_2\left(1-\rho^2\right)\Delta t}\left(e_2'\hat{H}_N^{(3)} e_1\right. \\
& - e_2'\hat{H}_N^{(1)}\left(e_1 - \Delta t e_2\right) - (1-\kappa\Delta t)\, e_1'\hat{H}_N^{(1)} e_2 + (1-\kappa\Delta t)\, e_2' H_N^{(2)}\left(e_1 - \Delta t e_2\right) \\
& \left. - \mu\Delta t\left(e_2'\hat{L}_N^{(2)} - (1-\kappa\Delta t)\,\hat{L}_N^{(1)}\right) - \kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - \mu\Delta t N\right)\right) \\
a_0^{\sigma_1} =& \frac{1}{\left(1-\rho^2\right)\Delta t}\left(e_1'\hat{H}_N^{(3)} e_1 - 2e_1'\hat{H}_N^{(1)}\left(e_1 - \Delta t e_2\right)\right. \\
& \left. + \left(e_1' - \Delta t e_2'\right)\hat{H}_N^{(2)}\left(e_1 - \Delta t e_2\right) + \mu^2\Delta t^2 N - 2\mu\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)}\right)\right)
\end{aligned}
$$

A.5 Proof of Theorem 3.5

Taking the derivative of $Q\left(\hat{\theta}_j, \theta\right)$ with respect to σ_2, we get

$$
\begin{aligned}
\frac{\partial Q\left(\hat{\theta}_j,\theta\right)}{\partial\sigma_2} =& -\frac{N}{\sigma_2} - \frac{1}{2}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(X_l - Q_l X_{l-1} - c_l\right)' \frac{\partial G^{-2}}{\partial\sigma_2}\left(X_l - Q_l X_{l-1} - c_l\right)|Y_N\right] \\
& + E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(\frac{\partial d_l}{\partial\sigma_2}\right)' H^{-2}\left(Y_l - M_l X_l - d_l\right)|Y_N\right] \\
=& -\frac{N}{\sigma_2} - \frac{1}{2}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(X_l - Q_l X_{l-1} - c_l\right)' \begin{bmatrix} 0 & \frac{\rho}{\sigma_2^2\sigma_1\left(1-\rho^2\right)\Delta t} \\ \frac{\rho}{\sigma_2^2\sigma_1\left(1-\rho^2\right)\Delta t} & \frac{-2}{\sigma_2^3\Delta t\left(1-\rho^2\right)} \end{bmatrix}\right. \\
& \left. \times\left(X_l - Q_l X_{l-1} - c_l\right)|Y_N\right] + E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(\sigma_2 A_l^{\sigma_2} + B_l^{\sigma_2}\right)' H^{-2}\right. \\
& \left. \times\left(Y_l - M_l X_l - \frac{1}{2}\sigma_2^2 A_l^{\sigma_2} - \sigma_2 B_l^{\sigma_2} - D_l^{\sigma_2}\right)|Y_N\right].
\end{aligned}
$$

Expanding the matrix multiplication in the second term of last equation as follows

$$-\frac{1}{2\left(1-\rho^{2}\right)\Delta t}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\begin{bmatrix}V\\Z\end{bmatrix}'\begin{bmatrix}0 & \frac{\rho}{\sigma_1\sigma_2^2}\\ \frac{\rho}{\sigma_1\sigma_2^2} & -\frac{2}{\sigma_2^3}\end{bmatrix}\begin{bmatrix}V\\Z\end{bmatrix}|Y_N\right]$$

$$=-\frac{1}{2\left(1-\rho^{2}\right)\Delta t}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\begin{bmatrix}-\frac{\rho}{\sigma_1\sigma_2^2}Z\\ \frac{\rho}{\sigma_1\sigma_2^2}V-\frac{2Z}{\sigma_2^3}\end{bmatrix}'\begin{bmatrix}V\\Z\end{bmatrix}|Y_N\right]$$

$$=-\frac{1}{2\left(1-\rho^{2}\right)\Delta t}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(2VZ\frac{\rho}{\sigma_1\sigma_2^2}-\frac{2Z^2}{\sigma_2^3}\right)|Y_N\right]$$

$$=-\frac{\rho}{\sigma_1\sigma_2^2\left(1-\rho^{2}\right)\Delta t}\left(e_2'\hat{H}_N^{(3)}e_1-e_2'\hat{H}_N^{(1)}\left(e_1-\Delta t e_2\right)\right.$$

$$-\left(1-\kappa\Delta t\right)e_1'\hat{H}_N^{(1)}e_2+\left(1-\kappa\Delta t\right)e_2'H_N^{(2)}\left(e_1-\Delta t e_2\right)$$

$$-\left(\mu-\frac{1}{2}\sigma_1^2\right)\Delta t\left(e_2'\hat{L}_N^{(2)}-\left(1-\kappa\Delta t\right)e_2'\hat{L}_N^{(1)}\right)$$

$$\left.-\kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)}-\left(e_1'-\Delta t e_2'\right)\hat{L}_N^{(1)}-N\left(\mu-\frac{1}{2}\sigma_1^2\right)\Delta t\right)\right)$$

$$+\frac{1}{\sigma_2^3\left(1-\rho^{2}\right)\Delta t}\left(e_2'\hat{H}_N^{(3)}e_2-2\left(1-\kappa\Delta t\right)e_2'\hat{H}^{(1)}e_2\right.$$

$$\left.+\left(1-\kappa\Delta t\right)^2e_2'\hat{H}_N^{(2)}e_2-2\kappa\alpha\Delta t\left(e_2'\hat{L}_N^{(2)}-\left(1-\kappa\Delta t\right)e_2'\hat{L}_N^{(1)}\right)+N\kappa^2\alpha^2\Delta t^2\right)$$

$$
\begin{aligned}
&\frac{\partial Q\left(\hat{\theta}_j,\theta\right)}{\partial\sigma_2}\\
&=-\frac{N}{\sigma_2}-\frac{\rho}{\sigma_1\sigma_2^2\left(1-\rho^2\right)\Delta t}\left(e_2'\hat{H}_N^{(3)}e_1-e_2'\hat{H}_N^{(1)}\left(e_1-\Delta te_2\right)\right.\\
&-\left(1-\kappa\Delta t\right)e_1'\hat{H}_N^{(1)}e_2+\left(1-\kappa\Delta t\right)e_2'H_N^{(2)}\left(e_1-\Delta te_2\right)\\
&-\left(\mu-\frac{1}{2}\sigma_1^2\right)\Delta t\left(e_2'\hat{L}_N^{(2)}-\left(1-\kappa\Delta t\right)e_2'\hat{L}_N^{(1)}\right)\\
&\left.-\kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)}-\left(e_1'-\Delta te_2'\right)\hat{L}_N^{(1)}-N\left(\mu-\frac{1}{2}\sigma_1^2\right)\Delta t\right)\right)\\
&+\frac{1}{\sigma_2^3\left(1-\rho^2\right)\Delta t}\left(e_2'\hat{H}_N^{(3)}e_2-2\left(1-\kappa\Delta t\right)e_2'\hat{H}^{(1)}e_2\right.\\
&\left.+\left(1-\kappa\Delta t\right)^2e_2'\hat{H}_N^{(2)}e_2-2\kappa\alpha\Delta t\left(e_2'\hat{L}_N^{(2)}-\left(1-\kappa\Delta t\right)e_2'\hat{L}_N^{(1)}\right)+N\kappa^2\alpha^2\Delta t^2\right)\\
&-\frac{1}{2}\sigma_2^3\sum_{l=0}^{N}\left(A_l^{\sigma_2}\right)'H^{-2}A_l^{\sigma_2}-\sigma_2^2\left(\sum_{l=0}^{N}\left(A_l^{\sigma_2}\right)'H^{-2}B_l^{\sigma_2}+\frac{1}{2}\sum_{l=0}^{N}\left(B_l^{\sigma_2}\right)'H^{-2}A_l^{\sigma_2}\right)\\
&+\sigma_2\left(\sum_{l=0}^{N}\left(A_l^{\sigma_2}\right)'H^{-2}\left(Y_l-D_l^{\sigma_2}\right)-\hat{K}_N\left(A_l^{\sigma_2}\right)-\sum_{l=0}^{N}\left(B_l^{\sigma_2}\right)'H^{-2}B_l^{\sigma_2}\right)\\
&+\sum_{l=0}^{N}\left(B_l^{\sigma_2}\right)'H^{-2}\left(Y_l-D_l^{\sigma_2}\right)-\hat{K}_N\left(B_l^{\sigma_2}\right)
\end{aligned}
$$

Setting equation $\frac{\partial Q(\hat{\theta}_j,\theta)}{\partial\sigma_2}$ to zero and multiplying by σ_2^3 gives

$$
a_6^{\sigma_2}\sigma_2^6+a_5^{\sigma_2}\sigma_2^5+a_4^{\sigma_2}\sigma_2^4+a_3^{\sigma_2}\sigma_2^3+a_2^{\sigma_2}\sigma_2^2+a_1^{\sigma_2}\sigma_2+a_0^{\sigma_2}=0
$$

with

$$
\begin{aligned}
a_6^{\sigma_2} &= -\frac{1}{2}\sum_{l=0}^{N}\left(A_l^{\sigma_2}\right)' H^{-2} A_l^{\sigma_2} \\
a_5^{\sigma_2} &= -\sum_{l=0}^{N}\left(A_l^{\sigma_2}\right)' H^{-2} B_l^{\sigma_2} - \frac{1}{2}\sum_{l=0}^{N}\left(B_l^{\sigma_2}\right)' H^{-2} A_l^{\sigma_2} \\
a_4^{\sigma_2} &= \sum_{l=0}^{N}\left(A_l^{\sigma_2}\right)' H^{-2}\left(Y_l - D_l^{\sigma_2}\right) - \hat{K}_N\left(A_l^{\sigma_2}\right) - \sum_{l=0}^{N}\left(B_l^{\sigma_2}\right)' H^{-2} B_l^{\sigma_2} \\
a_3^{\sigma_2} &= \sum_{l=0}^{N}\left(B_l^{\sigma_2}\right)' H^{-2}\left(Y_l - D_l^{\sigma_2}\right) - \hat{K}_N\left(B_l^{\sigma_2}\right) \\
a_2^{\sigma_2} &= -N \\
a_1^{\sigma_2} &= -\frac{\rho}{\sigma_1\left(1-\rho^2\right)\Delta t}\left(e_2'\hat{H}_N^{(3)} e_1 - e_2'\hat{H}_N^{(1)}\left(e_1 - \Delta t e_2\right)\right. \\
&\quad - \left(1-\kappa\Delta t\right) e_1'\hat{H}_N^{(1)} e_2 + \left(1-\kappa\Delta t\right) e_2' H_N^{(2)}\left(e_1 - \Delta t e_2\right) \\
&\quad - \left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t\left(e_2'\hat{L}_N^{(2)} - \left(1-\kappa\Delta t\right) e_2'\hat{L}_N^{(1)}\right) \\
&\quad \left. -\kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - N\left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t\right)\right) \\
a_0^{\sigma_2} &= \frac{1}{\left(1-\rho^2\right)\Delta t}\left(e_2'\hat{H}_N^{(3)} e_2 - 2\left(1-\kappa\Delta t\right) e_2'\hat{H}^{(1)} e_2\right. \\
&\quad \left. + \left(1-\kappa\Delta t\right)^2 e_2'\hat{H}_N^{(2)} e_2 - 2\kappa\alpha\Delta t\left(e_2'\hat{L}_N^{(2)} - e_2'\left(1-\kappa\Delta t\right)\hat{L}_N^{(1)}\right) + N\kappa^2\alpha^2\Delta t^2\right)
\end{aligned}
$$

A.6 Proof of Theorem 3.6

Take the derivative of $Q\left(\hat{\theta}_j, \theta\right)$ with respect to ρ to find

$$
\begin{aligned}
\frac{\partial Q\left(\hat{\theta}_j, \theta\right)}{\partial \rho} &= \frac{N\rho}{1-\rho^2} - \frac{1}{2} E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} \left(X_l - Q_l X_{l-1} - c_l\right)' \frac{\partial G^{-2}}{\partial \rho} \left(X_l - Q_l X_{l-1} - c_l\right) | Y_N\right] \\
&- E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} \left(-\frac{\partial d_l}{\partial \rho}\right)' H^{-2} \left(Y_l - M_l X_l - d_l\right) | Y_N\right] \\
&= \frac{N\rho}{1-\rho^2} - \frac{1}{2\Delta t \left(1-\rho^2\right)^2} E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} \left(X_l - Q_l X_{l-1} - c_l\right)' \begin{bmatrix} \frac{2\rho}{\sigma_1^2} & -\frac{1+\rho^2}{\sigma_1\sigma_2} \\ -\frac{1+\rho^2}{\sigma_1\sigma_2} & \frac{2\rho}{\sigma_2^2} \end{bmatrix}\right. \\
&\left.\times \left(X_l - Q_l X_{l-1} - c_l\right) | Y_N\right] + E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^2 \left(Y_l - M_l X_l - \rho A_l^{\rho} - B_l^{\rho}\right) | Y_N\right].
\end{aligned}
$$

Expanding the matrix multiplication in the second term of last equation gives

$$
\begin{aligned}
&-\frac{\rho}{\sigma_1^2\Delta t\left(1-\rho^2\right)^2}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}V^2|Y_N\right]+\frac{1+\rho^2}{\sigma_1\sigma_2\Delta t\left(1-\rho^2\right)^2}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}VZ|Y_N\right]\\
&-\frac{\rho}{\sigma_2^2\Delta t\left(1-\rho^2\right)^2}E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}Z^2|Y_N\right]\\
&=-\frac{\rho}{\sigma_1^2\Delta t\left(1-\rho^2\right)^2}\left(e_1'\hat{H}_N^{(3)}e_1-2e_1'\hat{H}_N^{(1)}\left(e_1-\Delta te_2\right)\right.\\
&+\left(e_1'-\Delta te_2'\right)\hat{H}_N^{(2)}\left(e_1-\Delta te_2\right)-2\left(\mu-\frac{1}{2}\sigma_1^2\right)\Delta t\left(e_1'\hat{L}_N^{(2)}-\left(e_1'-\Delta te_2'\right)\hat{L}_N^{(1)}\right)\\
&\left.+\left(\mu-\frac{1}{2}\sigma_1^2\right)^2\Delta t^2N\right)+\frac{1+\rho^2}{\sigma_1\sigma_2\Delta t\left(1-\rho^2\right)^2}\left(e_2'\hat{H}_N^{(3)}e_1\right.\\
&-e_2'\hat{H}_N^{(1)}\left(e_1-\Delta te_2\right)-\left(1-\kappa\Delta t\right)e_1'\hat{H}_N^{(1)}e_2+\left(1-\kappa\Delta t\right)e_2'\hat{H}_N^{(2)}\left(e_1-\Delta te_2\right)\\
&-\left(\mu-\frac{1}{2}\sigma_1^2\right)\left(e_2'\hat{L}^{(2)}-\left(1-\kappa\Delta t\right)e_2'\hat{L}_N^{(1)}\right)\\
&\left.-\kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)}-\left(e_1'-\Delta te_2'\right)\hat{L}_N^{(1)}-N\left(\mu-\frac{1}{2}\sigma_1^2\right)\Delta t\right)\right)\\
&-\frac{\rho}{\sigma_2^2\Delta t\left(1-\rho^2\right)^2}\left(e_2'\hat{H}_N^{(3)}e_2-2\left(1-\kappa\Delta t\right)e_2'\hat{H}^{(1)}e_2\right.\\
&\left.+\left(1-\kappa\Delta t\right)^2e_2'\hat{H}_N^{(2)}e_2-2\kappa\alpha\Delta t\left(e_2'\hat{L}_N^{(2)}-\left(1-\kappa\Delta t\right)e_2'\hat{L}_N^{(1)}\right)+N\kappa^2\alpha^2\Delta t^2\right)
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial Q\left(\hat{\theta}_j,\theta\right)}{\partial\rho} \\
&=\frac{N\rho}{1-\rho^2}-\frac{\rho}{\sigma_1^2\Delta t\left(1-\rho^2\right)^2}\left(e_1'\hat{H}_N^{(3)}e_1-2e_1'\hat{H}_N^{(1)}\left(e_1-\Delta te_2\right)\right. \\
&+\left(e_1'-\Delta te_2'\right)\hat{H}_N^{(2)}\left(e_1-\Delta te_2\right)-2\left(\mu-\frac{1}{2}\sigma_1^2\right)\Delta t\left(e_1'\hat{L}_N^{(2)}-\left(e_1'-\Delta te_2'\right)\hat{L}_N^{(1)}\right) \\
&\left.+\left(\mu-\frac{1}{2}\sigma_1^2\right)^2\Delta t^2N\right)+\frac{1+\rho^2}{\sigma_1\sigma_2\Delta t\left(1-\rho^2\right)^2}\left(e_2'\hat{H}_N^{(3)}e_1\right. \\
&-e_2'\hat{H}_N^{(1)}\left(e_1-\Delta te_2\right)-\left(1-\kappa\Delta t\right)e_1'\hat{H}_N^{(1)}e_2+\left(1-\kappa\Delta t\right)e_2'H_N^{(2)}\left(e_1-\Delta te_2\right) \\
&-\left(\mu-\frac{1}{2}\sigma_1^2\right)\left(e_2'\hat{L}^{(2)}-\left(1-\kappa\Delta t\right)e_2'\hat{L}_N^{(1)}\right) \\
&\left.-\kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)}-\left(e_1'-\Delta te_2'\right)\hat{L}_N^{(1)}-N\left(\mu-\frac{1}{2}\sigma_1^2\right)\Delta t\right)\right) \\
&-\frac{\rho}{\sigma_2^2\Delta t\left(1-\rho^2\right)^2}\left(e_2'\hat{H}_N^{(3)}e_2-2\left(1-\kappa\Delta t\right)e_2'\hat{H}^{(1)}e_2\right. \\
&\left.+\left(1-\kappa\Delta t\right)^2e_2'\hat{H}_N^{(2)}e_2-2\kappa\alpha\Delta t\left(e_2'\hat{L}_N^{(2)}-\left(1-\kappa\Delta t\right)e_2'\hat{L}_N^{(1)}\right)+N\kappa^2\alpha^2\Delta t^2\right) \\
&+\sum_{l=0}^{N}\left(A_l^{\rho}\right)'H^2\left(Y_l-B_l^{\rho}\right)-\hat{K}_N\left(A_l^{\rho}\right)-\rho\sum_{l=0}^{N}\left(A_l^{\rho}\right)'H^{-2}A_l^{\rho}. \qquad \text{(A.6)}
\end{aligned}
$$

Setting equation (A.6) and multiplying by ρ^3 gives

$$
a_5^{\rho}\rho^5+a_4^{\rho}\rho^4+a_3^{\rho}\rho^3+a_2^{\rho}\rho^2+a_1^{\rho}\rho+a_0^{\rho}=0
$$

where

$$
\begin{aligned}
a_5^{\rho} =& -\sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^{-2} A_l^{\rho} \\
a_4^{\rho} =& \sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^{-2} \left(Y_l - B_l^{\rho}\right) - \hat{K}_N \left(A_l^{\rho}\right) \\
a_3^{\rho} =& 2\sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^{-2} A_l^{\rho} - N \\
a_2^{\rho} =& -2\left(\sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^{-2} \left(Y_l - B_l^{\rho}\right) - \hat{K}_N^{(6)}\right) + \frac{1}{\sigma_1\sigma_2\Delta t\left(1-\rho^2\right)^2}\left(e_2'\hat{H}_N^{(3)} e_1\right. \\
& - e_2'\hat{H}_N^{(1)}\left(e_1 - \Delta t e_2\right) - \left(1-\kappa\Delta t\right) e_1'\hat{H}_N^{(1)} e_2 + \left(1-\kappa\Delta t\right) e_2'\hat{H}_N^{(2)}\left(e_1 - \Delta t e_2\right) \\
& - \left(\mu - \frac{1}{2}\sigma_1^2\right)\left(e_2'\hat{L}^{(2)} - e_2'\left(1-\kappa\Delta t\right)\hat{L}_N^{(1)}\right) \\
& \left. - \kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - N\left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t\right)\right) \\
a_1^{\rho} =& N - \sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^{-2} A_l^{\rho} - \frac{\rho}{\sigma_1^2\Delta t\left(1-\rho^2\right)^2}\left(e_1'\hat{H}_N^{(3)} e_1 - 2e_1'\hat{H}_N^{(1)}\left(e_1 - \Delta t e_2\right)\right. \\
& + \left(e_1' - \Delta t e_2'\right)\hat{H}_N^{(2)}\left(e_1 - \Delta t e_2\right) - 2\left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)}\right) \\
& \left. + \left(\mu - \frac{1}{2}\sigma_1^2\right)^2 \Delta t^2 N\right) - \frac{\rho}{\sigma_2^2\Delta t\left(1-\rho^2\right)^2}\left(e_2'\hat{H}_N^{(3)} e_2 - 2\left(1-\kappa\Delta t\right) e_2'\hat{H}^{(1)} e_2\right. \\
& \left. + \left(1-\kappa\Delta t\right)^2 e_2'\hat{H}_N^{(2)} e_2 - 2\kappa\alpha\Delta t\left(e_2'\hat{L}_N^{(2)} - \left(1-\kappa\Delta t\right) e_2'\hat{L}_N^{(1)}\right) + N\kappa^2\alpha^2\Delta t^2\right) \\
a_0^{\rho} =& \frac{1}{\sigma_1\sigma_2\Delta t\left(1-\rho^2\right)^2}\left(e_2'\hat{H}_N^{(3)} e_1 - e_2'\hat{H}_N^{(1)}\left(e_1 - \Delta t e_2\right)\right. \\
& - \left(1-\kappa\Delta t\right) e_1'\hat{H}_N^{(1)} e_2 + \left(1-\kappa\Delta t\right) e_2' H_N^{(2)}\left(e_1 - \Delta t e_2\right) \\
& - \left(\mu - \frac{1}{2}\sigma_1^2\right)\left(e_2'\hat{L}^{(2)} - \left(1-\kappa\Delta t\right) e_2'\hat{L}_N^{(1)}\right) \\
& \left. - \kappa\alpha\Delta t\left(e_1'\hat{L}_N^{(2)} - \left(e_1' - \Delta t e_2'\right)\hat{L}_N^{(1)} - N\left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t\right)\right) \\
& + \sum_{l=0}^{N} \left(A_l^{\rho}\right)' H^{-2} \left(Y_l - B_l^{\rho}\right) - \hat{K}_N \left(A_l^{\rho}\right)
\end{aligned}
$$

A.7 Proof of Theorem 3.7

Take derivative of $Q\left(\hat{\theta}_j, \theta\right)$ with respect to H^2 by matrix calculus and get

$$\frac{\partial Q\left(\hat{\theta}_j,\theta\right)}{\partial H^2} = -\frac{(N+1)}{2}H^{-2} + \frac{1}{2}H^{-2}E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}(Y_l - M_lX_l - d_l)(Y_l - M_lX_l - d_l)'\,|Y_N\right]H^{-2} \tag{A.7}$$

Set Equation (A.7) to zero and simplify. So we can get Equation (3.24)

$$\begin{aligned} H^2 &= \frac{1}{N+1}E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}(Y_l - M_lX_l - d_l)(Y_l - M_lX_l - d_l)'\,|\mathcal{Y}_N\right] \\ &= \frac{1}{N+1}\left(\sum_{l=0}^{N}Y_lY_l' - \sum_{l=0}^{N}Y_ld_l' - \hat{J}_N - \left(\hat{J}_N\right)' + \hat{U}_N - \sum_{l=0}^{N}d_lY_l' + \sum_{l=0}^{N}d_ld_l'\right) \end{aligned}$$

where $\hat{J}_N$ and $\hat{U}_N$ are defined as in equation (3.25) and (3.26) respectively.

A.8 Proof of Theorem 3.8

Taking derivative of $Q\left(\hat{\theta}_j, \theta\right)$ with respect to κ gives

$$\begin{aligned} \frac{\partial Q\left(\hat{\theta}_j,\theta\right)}{\partial \kappa} =& E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(\frac{\partial c_l}{\partial \kappa} + \frac{\partial Q_l}{\partial \kappa}X_{l-1}\right)'G^{-2}(X_l - Q_lX_{l-1} - c_l)\,|\mathcal{Y}_N\right] \\ &+ E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(\frac{\partial M_l}{\partial \kappa}X_l + \frac{\partial d_l}{\partial \kappa}\right)'H^{-2}(Y_l - M_lX_l - d_l)\,|\mathcal{Y}_N\right] \\ &= E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(\frac{\partial c_l}{\partial \kappa}\right)'G^{-2}(X_l - Q_lX_{l-1} - c_l)\,|Y_N\right] \\ &+ E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}X_{l-1}'\left(\frac{\partial Q_l}{\partial \kappa}\right)'G^{-2}(X_l - Q_lX_{l-1} - c_l)\,|\mathcal{Y}_N\right] \\ &+ E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}X_l'\left(\frac{\partial M_l}{\partial \kappa}\right)'H^{-2}(Y_l - M_lX_l - d_l)\,|\mathcal{Y}_N\right] \\ &+ E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(\frac{\partial d_l}{\partial \kappa}\right)'H^{-2}(Y_l - M_lX_l - d_l)\,|\mathcal{Y}_N\right] \end{aligned} \tag{A.8}$$

Let us first calculate $\frac{\partial M_l}{\partial\kappa}$, $\frac{\partial Q_l}{\partial\kappa}$ and $\frac{\partial d_l}{\partial\kappa}$

$$
\begin{aligned}
\frac{\partial M_l}{\partial\kappa} &= \left[\mathbf{0}_{M\times 1}, \frac{-\tau_l e^{-\kappa\tau_l}}{\kappa} + \frac{1-e^{-\kappa\tau_l}}{\kappa^2}\right] \in \mathcal{R}^{M\times 2} \\
\frac{\partial Q_l}{\partial\kappa} &= -\Delta t e_2 e_2' \in \mathcal{R}^{2\times 2} \\
\frac{\partial d_l}{\partial\kappa} &= \left(\frac{-\lambda}{\kappa^2} - \frac{\sigma_2^2}{\kappa^3} + \frac{\sigma_1\sigma_2\rho}{\kappa^2}\right)\tau_l + \frac{1}{2}\frac{\sigma_2^2\tau_l e^{-2\kappa\tau}}{\kappa^3} - \frac{3}{4}\frac{\sigma_2^2\left(1-e^{-2\kappa\tau_l}\right)}{\kappa^4} \\
&\quad + \frac{\left(\alpha+\frac{\sigma_2^2}{\kappa^2}\right)\left(1-e^{-\kappa\tau_l}\right)}{\kappa^2} + \frac{\left(\alpha\kappa-\lambda+\sigma_1\sigma_2\rho-\frac{\sigma_2^2}{\kappa}\right)\tau_l e^{-\kappa\tau_l}}{\kappa^2} \\
&\quad - \frac{2\left(\alpha\kappa-\lambda+\sigma_1\sigma_2\rho-\frac{\sigma_2^2}{\kappa}\right)\left(1-e^{-\kappa\tau_l}\right)}{\kappa^3} \in \mathcal{R}^{M\times 1}
\end{aligned}
$$

If we set equation (A.8) to zero, we find that it is impossible to solve for κ or even obtain an elementary equation that must be statisfied. Therefore, we use an approximation of $e^{-\kappa\tau_l}$ and $e^{-2\kappa\tau_l}$. Using Taylor's theorem, approximate them around $\hat{\kappa}_j$ (the j-th estimate of κ), to obtain

$$
e^{-\kappa\tau_l^{(i)}} \approx e^{-\hat{\kappa}_j\tau_l^{(i)}} - \tau_l^{(i)} e^{-\hat{\kappa}_j\tau_l^{(i)}}\left(\kappa-\hat{\kappa}_j\right) + \frac{1}{2}\left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j\tau_l^{(i)}}\left(\kappa-\hat{\kappa}_j\right)^2 \tag{A.9}
$$

$$
e^{-2\kappa\tau_l^{(i)}} \approx e^{-2\hat{\kappa}_j\tau_l^{(i)}} - 2\tau_l^{(i)} e^{-2\hat{\kappa}_j\tau_l^{(i)}}\left(\kappa-\hat{\kappa}_j\right) + 2\left(\tau_l^{(i)}\right)^2 e^{-2\hat{\kappa}_j\tau_l^{(i)}}\left(\kappa-\hat{\kappa}_j\right)^2 \tag{A.10}
$$

where $l = 1,\ldots,N$ and $j = 1,\ldots,M$ by which $\tau_l^{(j)}$ is the term structure of the time for the l-th observation of j-th futures price.

Substitute $e^{-\kappa\tau_l^{(i)}}$ and $e^{-2\kappa\tau_l^{(i)}}$ by Equation (A.9) and (A.10) respectively. Therefore, each element of M_l, d_l, $\frac{\partial M_l}{\partial\kappa}$ and $\frac{\partial d_l}{\partial\kappa}$ can be approximated by Equation (A.11) to (A.14)

$$
\begin{aligned}
M_l^{(i)} &\approx \left[1, -\frac{1}{\kappa} + \frac{e^{-\hat{\kappa}_j\tau_l^{(i)}}}{\kappa} - \tau_l^{(i)} e^{-\hat{\kappa}_j\tau_l^{(i)}} + \frac{\hat{\kappa}_j\tau_l^{(i)} e^{-\hat{\kappa}_j\tau_l^{(i)}}}{\kappa} + \frac{1}{2}\left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j\tau_l^{(i)}}\kappa \right. \\
&\qquad \left. -\hat{\kappa}_j\left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} + \frac{1}{2}\frac{\hat{\kappa}_j^2\left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j\tau_l^{(i)}}}{\kappa}\right] \\
&= \left[1, \sum_{n=-1}^{1}\kappa^n m_l^{n(i)}\right]
\end{aligned} \tag{A.11}
$$

where

$$
\begin{aligned}
m_l^{-1(i)} =& -1 + e^{-\hat{\kappa}_j \tau_l^{(i)}} + \hat{\kappa}_j \tau_l^{(i)} e^{-\hat{\kappa}_j \tau_l^{(i)}} + \frac{1}{2}\hat{\kappa}_j^2 \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} && \in \mathcal{R} \\
m_l^{0(i)} =& -\tau_l^{(i)} e^{\hat{\kappa}_j \tau_l^{(i)}} - \hat{\kappa}_j \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} && \in \mathcal{R} \\
m_l^{1(i)} =& \frac{1}{2}\left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} \in \mathcal{R}
\end{aligned}
$$

and $m_l^{n(i)}$, $n = -1, 0, 1$ represent the coefficients of κ^n respectively.

Then substituting

$$
\begin{aligned}
\frac{\partial M_l^{(i)}}{\partial \kappa} \approx & \left[0, \frac{1}{2}\left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} - \frac{1}{2}\left(\tau_l^{(i)}\right)^3 e^{-\hat{\kappa}_j \tau_l^{(i)}} \kappa + \hat{\kappa}_j \left(\tau_l^{(i)}\right)^3 e^{-\hat{\kappa}_j \tau_l^{(i)}} \right. \\
& \left. -\frac{1}{2}\frac{\hat{\kappa}_j^2 \left(\tau_l^{(i)}\right)^3 e^{-\hat{\kappa}_j \tau_l^{(i)}}}{\kappa} + \frac{1}{\kappa^2} - \frac{e^{-\hat{\kappa}_j \tau_l^{(i)}}}{\kappa^2} - \frac{\hat{\kappa}_j \tau_l^{(i)} e^{-\hat{\kappa}_j \tau_l^{(i)}}}{\kappa^2} - \frac{1}{2}\frac{\hat{\kappa}_j^2 \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}}}{\kappa^2}\right] \\
= & \left[0, \sum_{n=-2}^{1} \kappa^n h_l^{n(i)}\right] \in \mathcal{R}^{1\times 2}
\end{aligned}
\tag{A.12}
$$

where

$$
\begin{aligned}
h_l^{-2(i)} =& 1 - e^{-\hat{\kappa}_j \tau_l^{(i)}} - \hat{\kappa}_j \tau_l^{(i)} e^{-\hat{\kappa}_j \tau_l^{(i)}} - \frac{1}{2}\hat{\kappa}_j^2 \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} \\
h_l^{-1(i)} =& -\frac{1}{2}\hat{\kappa}_j^2 \left(\tau_l^{(i)}\right)^3 e^{-\hat{\kappa}_j \tau_l^{(i)}} \\
h_l^{0(i)} =& \frac{1}{2}\left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} + \hat{\kappa}_j \left(\tau_l^{(i)}\right)^3 e^{-\hat{\kappa}_j \tau_l^{(i)}} \\
h_l^{1(i)} =& -\frac{1}{2}\left(\tau_l^{(i)}\right)^3 e^{-\hat{\kappa}_j \tau_l^{(i)}}
\end{aligned}
$$

and $h_l^{n(i)}$, $n = -2, -1, 0, 1$ represent the coefficients of κ^n respectively.

$$
d_l \approx \sum_{n=-3}^{1} \kappa^n s_l^{n(i)} \tag{A.13}
$$

where

$$
\begin{aligned}
s_l^{-3(i)} =& \sigma_2^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} - \frac{1}{2}\sigma_2^2 \tau_l^{(i)} \hat{\kappa}_j e^{-2\hat{\kappa}_j \tau_l^{(i)}} - \frac{1}{2}\sigma_2^2 \hat{\kappa}_j^2 \left(\tau_l^{(i)}\right)^2 e^{-2\hat{\kappa}_j \tau_l^{(j)}} + \sigma_2^2 \hat{\kappa}_j \tau_l^{(i)} e^{-\hat{\kappa}_j \tau_l^{(i)}} \\
& + \frac{1}{2}\sigma_2^2 \left(\tau_l^{(j)}\right)^2 \hat{\kappa}_j^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} - \frac{3}{4}\sigma_2^2 - \frac{1}{4}\sigma_2^2 e^{-2\hat{\kappa}_j \tau_l^{(i)}} \\
s_l^{-2(i)} =& -\frac{1}{2}\sigma_1\sigma_2\rho\hat{\kappa}_j^2 \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} - \sigma_1\sigma_2\rho\hat{\kappa}_j \tau_l^{(i)} e^{-\hat{\kappa}_j \tau_l^{(i)}} + \frac{1}{2}\sigma_2^2 \tau_l^{(i)} - \lambda + \sigma_2^2 \hat{\kappa}_j \left(\tau_l^{(i)}\right)^2 e^{-2\hat{\kappa}_j \tau_l^{(i)}} \\
& - \sigma_2^2 \left(\tau_l^{(i)}\right)^2 \hat{\kappa}_j e^{-\hat{\kappa}_j \tau_l^{(j)}} + \lambda \tau_l^{(j)} \hat{\kappa}_j e^{-\hat{\kappa}_j \tau_l^{(j)}} + \frac{1}{2}\lambda \tau_l^{(j)} 2\hat{\kappa}_j^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} - \sigma_1\sigma_2\rho e^{-\hat{\kappa}_j \tau_l^{(j)}} + \lambda e^{-\hat{\kappa}_j \tau_l^{(i)}} \\
& + \frac{1}{2}\sigma_2^2 \tau_l^{(i)} e^{-2\hat{\kappa}_j \tau_l^{(i)}} + \sigma_1\sigma_2\rho - \sigma_2^2 \tau_l^{(i)} e^{-\hat{\kappa}_j \tau_l^{(i)}} \\
s_l^{-1(i)} =& \sigma_1\sigma_2\rho\hat{\kappa}_j \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} + \lambda \tau_l^{(i)} - \alpha e^{-\hat{\kappa}_j \tau_l^{(i)}} + \alpha - \tau_l^{(i)} \sigma_1\sigma_2\rho - \alpha \tau_l^{(i)} \hat{\kappa}_j e^{-\hat{\kappa}_j \tau_l^{(i)}} \\
& - \frac{1}{2}\alpha\hat{\kappa}_j^2 \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} - \lambda\hat{\kappa}_j \left(\tau_l^{(j)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} + \sigma_1\sigma_2\rho\tau_l^{(i)} e^{-\hat{\kappa}_j \tau_l^{(i)}} - \frac{1}{2}\sigma_2^2 \left(\tau_l^{(i)}\right)^2 e^{-2\hat{\kappa}_j \tau_l^{(i)}} \\
& - \lambda \tau_l^{(i)} e^{-\hat{\kappa}_j \tau_l^{(i)}} + \frac{1}{2}\sigma_2^2 \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa} \tau_l^{(i)}} \\
s_l^{0(i)} =& -\frac{1}{2}\sigma_1\sigma_2\rho \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} + \tau_l^{(i)} r - \tau_l^{(i)} \alpha + \alpha\hat{\kappa}_j \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} + \alpha \tau_l^{(i)} e^{-\hat{\kappa}_j \tau_l^{(i)}} \\
& + \frac{1}{2}\lambda \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}} \\
s_l^{1(i)} =& -\frac{1}{2}\alpha \left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j \tau_l^{(i)}}
\end{aligned}
$$

and $s_l^{n(i)}$, $n = -3, -2, -1, 0, 1$ represent the coefficients of κ^n respectively.

$$
\frac{\partial d_l^{(i)}}{\partial \kappa} \approx \sum_{n=-4}^{1} \kappa^n p_l^{n(i)} \tag{A.14}
$$

where

$$
\begin{aligned}
p_l^{-4(i)} =& \frac{3}{4}\sigma_2^2 e^{-2\hat{\kappa}_j\tau_l^{(i)}} - 3\sigma_2^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} + \frac{9}{4}\sigma_2^2 + \frac{3}{2}\sigma_2^2\tau_l^{(i)}\hat{\kappa}_j e^{-2\hat{\kappa}_j\tau_l^{(i)}} \\
&+ \frac{3}{2}\sigma_2^2\left(\tau_l^{(i)}\right)^2\hat{\kappa}_j^2 e^{-2\hat{\kappa}_j\tau_l^{(i)}} - 3\sigma_2^2\tau_l^{(i)}\hat{\kappa}_j e^{-\hat{\kappa}_j\tau_l^{(i)}} - \frac{3}{2}\sigma_2^2\left(\tau_l^{(i)}\right)^2\hat{\kappa}_j^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} \\
p_l^{-3(i)} =& 2\sigma_1\sigma_2\rho\tau_l^{(i)}\hat{\kappa}_j e^{-\hat{\kappa}_j\tau_l^{(i)}} + \sigma_1\sigma_2\rho\left(\tau_l^{(i)}\right)^2\hat{\kappa}_j^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} - \sigma_2^2\tau - \sigma_2^2\tau_l^{(i)} e^{-2\hat{\kappa}_j\tau_l^{(i)}} \\
&+ 2\sigma_2^2\tau_l^{(i)} e^{-\hat{\kappa}_j\tau_l^{(i)}} - 2\sigma_1\sigma_2\rho - 2\lambda e^{-\hat{\kappa}_j\tau_l^{(i)}} + 2\lambda - 2\sigma_2^2\left(\tau_l^{(i)}\right)^2\hat{\kappa}_j e^{-2\hat{\kappa}_j\tau_l^{(i)}} \\
&+ \sigma_2^2\left(\tau_l^{(i)}\right)^3\hat{\kappa}_j^2 e^{-2\hat{\kappa}_j\tau_l^{(i)}} + 2\sigma_2^2\left(\tau_l^{(i)}\right)^2\hat{\kappa}_j e^{-\hat{\kappa}_j\tau_l^{(i)}} - \frac{1}{2}\left(\tau_l^{(i)}\right)^3\sigma_2^2\hat{\kappa}_j^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} \\
&- 2\lambda\tau_l^{(i)}\hat{\kappa}_j e^{-\hat{\kappa}_j\tau_l^{(i)}} - \lambda\left(\tau_l^{(i)}\right)^2\hat{\kappa}_j^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} + 2\sigma_1\sigma_2\rho e^{-\hat{\kappa}_j\tau_l^{(i)}} \\
p_l^{-2(i)} =& -\tau_l^{(i)}\sigma_1\sigma_2\rho e^{-\hat{\kappa}_j\tau_l^{(i)}} + \frac{1}{2}\sigma_2^2\left(\tau_l^{(i)}\right)^2 e^{-2\hat{\kappa}_j\tau_l^{(i)}} - \frac{1}{2}\sigma_2^2\left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} + \lambda\tau_l^{(i)} e^{-\hat{\kappa}_j\tau_l^{(i)}} \\
&+ \sigma_1\sigma_2\rho\tau_l^{(i)} + \alpha e^{-\hat{\kappa}_j\tau_l^{(i)}} - \alpha - 2\sigma_2^2\left(\tau_l^{(i)}\right)^3\hat{\kappa}_j e^{-2\hat{\kappa}_j\tau_l^{(i)}} + \alpha\tau_l^{(i)}\hat{\kappa}_j e^{-\hat{\kappa}_j\tau_l^{(i)}} \\
&+ \frac{1}{2}\alpha\left(\tau_l^{(i)}\right)^2\hat{\kappa}_j^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} + \left(\tau_l^{(i)}\right)^2\lambda\hat{\kappa}_j e^{-\hat{\kappa}_j\tau_l^{(i)}} - \frac{1}{2}\left(\tau_l^{(i)}\right)^3\lambda\hat{\kappa}_j^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} \\
&+ \left(\tau_l^{(i)}\right)^3\sigma_2^2\hat{\kappa}_j e^{-\hat{\kappa}_j\tau_l^{(i)}} - \lambda\tau_l^{(i)} - \left(\tau_l^{(i)}\right)^2\sigma_1\sigma_2\rho\hat{\kappa}_j e^{-\hat{\kappa}_j\tau_l^{(i)}} + \frac{1}{2}\left(\tau_l^{(i)}\right)^3\sigma_1\sigma_2\rho\hat{\kappa}_j^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} \\
p_l^{-1(i)} =& \sigma_2^2\left(\tau_l^{(i)}\right)^3 e^{-2\hat{\kappa}_j\tau_l^{(i)}} - \frac{1}{2}\sigma_2^2\left(\tau_l^{(i)}\right)^3 e^{-\hat{\kappa}_j\tau_l^{(i)}} + \frac{1}{2}\left(\tau_l^{(i)}\right)^3\alpha\hat{\kappa}_j^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} \\
&+ \left(\tau_l^{(i)}\right)^3\lambda\hat{\kappa}_j e^{-\hat{\kappa}_j\tau_l^{(i)}} - \left(\tau_l^{(i)}\right)^3\sigma_1\sigma_2\rho\hat{\kappa}_j e^{-\hat{\kappa}_j\tau_l^{(i)}} \\
p_l^{0(i)} =& \frac{1}{2}\left(\tau_l^{(i)}\right)^3\sigma_1\sigma_2\rho e^{-\hat{\kappa}_j\tau_l^{(i)}} - \frac{1}{2}\alpha\left(\tau_l^{(i)}\right)^2 e^{-\hat{\kappa}_j\tau_l^{(i)}} - \frac{1}{2}\lambda\left(\tau_l^{(i)}\right)^3 e^{-\hat{\kappa}_j\tau_l^{(i)}} - \left(\tau_l^{(i)}\right)^3\alpha\hat{\kappa}_j e^{-\hat{\kappa}_j\tau_l^{(i)}} \\
p_l^{1(i)} =& \frac{1}{2}\left(\tau_l^{(i)}\right)^3\alpha e^{-\hat{\kappa}_j\tau_l^{(i)}}
\end{aligned}
$$

p_{-4}, $n = -1, 0, 1$ represent the coefficients of κ^n respectively.

Now focus on Equation (A.8). We will calculate each term of it respectively. First calculate

$E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(\frac{\partial c_l}{\partial \kappa}\right)' G^{-2}\left(X_l - Q_l X_{l-1} - c_l\right)|\mathcal{Y}_N\right]$.

$$
\begin{aligned}
& E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(\frac{\partial c_l}{\partial \kappa}\right)' G^{-2}\left(X_l - Q_l X_{l-1} - c_l\right)|\mathcal{Y}_N\right] \\
=& E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\alpha\Delta t e_2' G^{-2}\left(X_l - Q_l X_{l-1} - \left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t e_1 - \kappa\alpha\Delta t e_2\right)|\mathcal{Y}_N\right] \\
=& \alpha\Delta t e_2' G^{-2}\left(\hat{L}_N^{(2)} - Q_l \hat{L}_N^{(1)} - N\left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t e_1\right) - \kappa N \alpha^2 \Delta t^2 e_2' G^{-2} e_2
\end{aligned}
$$

Next we consider $E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} X_{l-1}'\left(\frac{\partial Q_l}{\partial \kappa}\right)' G^{-2}\left(X_l - Q_l X_{l-1} - c_l\right)|\mathcal{Y}_N\right]$.

$$
\begin{aligned}
& E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} X_{l-1}'\left(\frac{\partial Q_l}{\partial \kappa}\right)' G^{-2}\left(X_l - Q_l X_{l-1} - c_l\right)|\mathcal{Y}_N\right] \\
=& E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} X_{l-1}'\left(-\Delta t\right) e_2 e_2' G^{-2}\left(X_l - Q_l X_{l-1} - c_l\right)|\mathcal{Y}_N\right] \\
=& -\Delta t e_2' G^{-2} E_{\hat{\theta}_j}\left[\sum_{l=1}^{N} X_{l-1}' e_2\left(X_l - Q_l X_{l-1} - c_l\right)|Y_N\right] \\
=& -\Delta t e_2' G^{-2} E_{\hat{\theta}_j}\left[\sum_{l=1}^{N}\left(X_l X_{l-1}' e_2 - Q_l X_{l-1} X_{l-1}' e_2 - c_l X_{l-1}' e_2\right)|Y_N\right] \\
=& -\Delta t e_2' G^{-2}\left(\hat{H}_N^{(1)} e_2 - \left(e_1 e_1' - \Delta t e_1 e_2' + e_2 e_2' - \kappa\Delta t e_2 e_2'\right)\hat{H}_N^{(2)} e_2 - \left(\left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t e_1\right.\right. \\
& \left.\left. + \kappa\alpha\Delta t e_2\right)\left(\hat{L}_N^{(1)}\right)' e_2\right) \\
=& -\Delta t e_2' G^{-2}\left(\hat{H}_N^{(1)} e_2 - \left(e_1 e_1' - \Delta t e_1 e_2' + e_2 e_2'\right)\hat{H}_N^{(2)} e_2 - \left(\mu - \frac{1}{2}\sigma_1^2\right)\Delta t e_1\left(\hat{L}_N^{(1)}\right)' e_2\right) \\
& + \kappa\left(-\Delta t^2 e_2' G^{-2} e_2 e_2' \hat{H}_N^{(2)} e_2 + \alpha\Delta t^2 e_2' G^{-2} e_2\left(\hat{L}_N^{(1)}\right)' e_2\right)
\end{aligned}
$$

Before calculating $E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} X_l'\left(\frac{\partial M_l}{\partial \kappa}\right)' H^{-2}\left(Y_l - M_l X_l - d_l\right)|\mathcal{Y}_N\right]$ and $E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(\frac{\partial d_l}{\partial \kappa}\right)' H^{-2}\left(Y_l - M_l X_l - d_l\right)|\mathcal{Y}_N\right]$, first we define some finite-dimensional filters

$$\hat{K}_N(h_m) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} H^{-2}\left[\mathbf{0}_{M\times 1},\, h_m\right] X_l | \mathcal{Y}_N\right]$$

$$m = -2, -1, 0, 1$$

$$\hat{K}_N(d_k, h_m) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} d_k' H^{-2}\left[\mathbf{0}_{M\times 1}, h_m\right] X_l | \mathcal{Y}_N\right]$$

$$k = -3, -2, -1, 0, 1, m = -2, -1, 0, 1$$

$$\hat{K}_N(p_m) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} p_m' H^{-2}\left[\mathbf{I}_{M\times 1}, \mathbf{0}_{M\times 1}\right] X_l | \mathcal{Y}_N\right]$$

$$m = -4, -3, -2, -1, 0, 1$$

$$\hat{K}_N(p_m, m_n) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} p_m' H^{-2}\left[\mathbf{0}_{M\times 1}, m_n\right] X_l | \mathcal{Y}_N\right]$$

$$m = -4, -3, -2, -1, 0, 1, n = -1, 0, 1$$

$$\hat{H}_N(dM_m) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} dM_m' H^{-2} \mathbf{I}_{M\times 1} X_l^{(2)} X_l^{(1)} | \mathcal{Y}_N\right]$$

$$m = -2, -1, 0, 1$$

$$\hat{H}_N(dM_m, m_n) = E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} dM_m' H^{-2} m_n X_l^{(2)} X_l^{(2)} | \mathcal{Y}_N\right]$$

$$m = -2, -1, 0, 1, n_l = -1, 0, 1$$

where $\mathbf{0}_{M\times1}$ is an $M\times1$ vector with 0 in each position, $\mathbf{I}_{M\times1}$ is an $M\times1$ vector with 1 in each position and $X_l^{(i)}$, $i=1,2$ is the i-th element of process X_l. Then

$$
\begin{aligned}
&E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} X_l'\left(\frac{\partial M_l}{\partial\kappa}\right)' H^{-2}\left(Y_l - M_l X_l - d_l\right)|\mathcal{Y}_N\right]\\
=&E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} X_l'\left(\frac{\partial M_l}{\partial\kappa}\right)' H^{-2} Y_l|\mathcal{Y}_N\right] - E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} X_l'\left(\frac{\partial M_l}{\partial\kappa}\right)' H^{-2} M_l X_l|\mathcal{Y}_N\right]\\
&-E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} X_l'\left(\frac{\partial M_l}{\partial\kappa}\right)' H^{-2} d_l|\mathcal{Y}_N\right]\\
=&E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} Y_l' H^{-2}\left[\mathbf{0}_{M\times1}, \sum_{n=-2}^{1}\kappa^n dM_n\right] X_l|\mathcal{Y}_N\right]\\
&-E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} X_l'\left[\mathbf{0}_{M\times1}, \sum_{n=-2}^{1}\kappa^n dM_n\right] H^{-2}\left[\mathbf{I}_{M\times1}, \mathbf{0}_{M\times1}\right] X_l|\mathcal{Y}_N\right]\\
&-E_{\hat{\theta}_j}\left[\sum_{l=0}^{N} X_l'\left[\mathbf{0}_{M\times1}, \sum_{n=-2}^{1}\kappa^n dM_n\right] H^{-2}\left[\mathbf{0}_{M\times1}, \sum_{n=-1}^{1}\kappa^n m_n\right]|\mathcal{Y}_N\right]\\
&-E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(\sum_{n=-3}^{1}\kappa^n d_n\right) H^{-2}\left[\mathbf{0}_{M\times1}, \sum_{n=-2}^{1}\kappa^n dM_n\right] X_l|\mathcal{Y}_N\right] \qquad \text{(A.15)}
\end{aligned}
$$

Thus equation (A.15) can be written as

$$
\begin{aligned}
&\sum_{n=-2}^{1}\kappa^n \hat{K}_N\left(dM_n\right) - \sum_{n=-2}^{1}\kappa^n \hat{H}_N\left(dM_n\right) - \sum_{m=-1}^{1}\sum_{n=-2}^{1}\kappa^{m+n}\hat{H}_N\left(dM_n, m_m\right)\\
&-\sum_{m=-3}^{1}\sum_{n=-2}^{1}\kappa^{m+n}\hat{K}_N\left(d_m, dM_n\right)
\end{aligned}
$$

Now continue to calculate $E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(\frac{\partial d_l}{\partial \kappa}\right)' H^{-2}\left(Y_l - M_l X_l - d_l\right)|\mathcal{Y}_N\right]$

$$
\begin{aligned}
& E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(\frac{\partial d_l}{\partial \kappa}\right)' H^{-2}\left(Y_l - M_l X_l - d_l\right)|\mathcal{Y}_N\right] \\
=& \sum_{l=0}^{N}\left(\frac{\partial d_l}{\partial \kappa}\right)' H^{-2} Y_l - E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(\frac{\partial d_l}{\partial \kappa}\right)' H^{-2} M_l X_l|\mathcal{Y}_N\right] - \sum_{l=0}^{N}\left(\frac{\partial d_l}{\partial \kappa}\right)' H^{-2} d_l \\
=& \sum_{l=0}^{N}\left(\sum_{n=-4}^{1}\kappa^n p_n\right)' H^{-2} Y_l \\
& - E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(\sum_{n=-4}^{1}\kappa^n p_n\right)' H^{-2}\left[\mathbf{0}_{M\times 1}, \sum_{n=-1}^{1}\kappa^n m_n\right] X_l|\mathcal{Y}_N\right] \\
& - E_{\hat{\theta}_j}\left[\sum_{l=0}^{N}\left(\sum_{n=-4}^{1}\kappa^n p_n\right)' H^{-2}\left[I_{M\times 1}, \mathbf{0}_{M\times 1}\right] X_l|\mathcal{Y}_N\right] \\
& - \sum_{l=0}^{N}\left(\sum_{n=-4}^{1}\kappa^n p_n\right)' H^{-2}\left(\sum_{n=3}^{1}\kappa^n d_n\right)
\end{aligned} \tag{A.16}
$$

Therefore, Equation (A.16) can be written as

$$
\begin{aligned}
& \sum_{n=-4}^{1}\left(\kappa^n \sum_{l=0}^{N} p_n' H^{-2} Y_l\right) - \sum_{n=-4}^{1}\left(\kappa^n \hat{K}_N\left(p_n\right)\right) - \sum_{m=-4}^{1}\sum_{n=-1}^{1}\kappa^{m+n}\hat{K}_N\left(p_m, m_n\right) \\
& - \sum_{m=-4}^{1}\sum_{n=-3}^{1}\kappa^{m+n} p_m' H^{-2} d_n
\end{aligned}
$$

From the calculation above,the $(j+1)$-th estimate of κ should satisfy the equation

$$
a_9^\kappa \kappa^9 + a_8^\kappa \kappa^8 + a_7^\kappa \kappa^7 + a_6^\kappa \kappa^6 + a_5^\kappa \kappa^5 + a_4^\kappa \kappa^4 + a_3^\kappa \kappa^3 + a_2^\kappa \kappa^2 + a_1^\kappa \kappa + a_0^\kappa = 0
$$

with

$$a_9^\kappa = -\hat{H}(dM_1, m_1) - \hat{K}_N(d_1, dM_1) - \hat{K}_N(p_1, m_1) - \sum_{l=0}^{N} p_1 H^{-2} d_1$$

$$a_8^\kappa = \sum_{l=0}^{N} p_1' H^{-2} Y_l + \left(-\Delta t^2 e_2' G^{-2} e_2 e_2' \hat{H}_N^{(2)} e_2 + \alpha \Delta t^2 e_2' G^{-2} e_2 \left(\hat{L}_N^{(1)}\right)' e_2\right)$$

$$+\hat{K}_N(dM_1) - \hat{H}(dM_1) - \hat{H}(dM_0, m_1) - \hat{H}(dM_1, m_0) - \hat{K}_N(d_0, dM_1)$$

$$-\hat{K}_N(d_1, dM_0) - \hat{K}_N(p_1) - \hat{K}_N(p_0, m_1) - \hat{K}_N(p_1, m_0) - \sum_{l=0}^{N} p_0 H^{-2} d_1 - \sum_{l=0}^{N} p_1 H^{-2} d_0$$

$$-N\alpha^2 \Delta t^2 e_2' G^{-2} e_2$$

$$a_7^\kappa = \sum_{l=0}^{N} p_0' H^{-2} Y_l - \Delta t e_2' G^{-2} \left(\hat{H}_N^{(1)} e_2 - \left(e_1 e_1' - \Delta t e_1 e_2' + e_2 e_2'\right) \hat{H}_N^{(2)} e_2\right.$$

$$\left. - \left(\mu - \frac{1}{2}\sigma_1^2\right) \Delta t e_1 \left(\hat{L}_N^{(1)}\right)' e_2\right)$$

$$+\hat{K}_N(dM_0) - \hat{H}(dM_0) - \hat{H}(dM_{-1}, m_1) - \hat{H}(dM_0, m_0)$$

$$-\hat{H}(dM_1, m_{-1}) - \hat{K}_N(d_{-1}, dM_1) - \hat{K}_N(d_0, dM_0)$$

$$-\hat{K}_N(d_1, dM_{-1}) - \hat{K}_N(p_0) - \hat{K}_N(p_{-1}, m_1) - \hat{K}_N(p_0, m_0)$$

$$-\hat{K}_N(p_1, m_{-1}) - \sum_{l=0}^{N} p_{-1} H^{-2} d_1 - \sum_{l=0}^{N} p_0 H^{-2} d_0 - \sum_{l=0}^{N} p_1 H^{-2} d_{-1}$$

$$+\alpha \Delta t e_2' G^{-2} \left(\hat{L}_N^{(2)} - Q_l \hat{L}_N^{(1)} - N\left(\mu - \frac{1}{2}\sigma_1^2\right) \Delta t e_1\right)$$

$$a_6^\kappa = \sum_{l=0}^{N} p_{-1}' H^{-2} Y_l + \hat{K}_N^\kappa(dM_{-1}) - \hat{H}(dM_{-1}) - \hat{H}(dM_{-2}, m_1)$$

$$-\hat{H}(dM_{-1}, m_0) - \hat{H}(dM_0, m_{-1}) - \hat{K}_N(d_{-2}, dM_1) - \hat{K}_N^\kappa(d_{-1}, dM_0)$$

$$-\hat{K}_N^\kappa(d_0, dM_{-1}) - \hat{K}_N^\kappa(d_1, dM_{-2}) - \hat{K}_N(p_{-1}) - \hat{K}_N(p_{-2}, m_1)$$

$$-\hat{K}_N(p_{-1}, m_0) - \hat{K}_N(p_0, m_{-1}) - \sum_{n=-2}^{1} \left(\sum_{l=0}^{N} p_{-n-1} H^{-2} d_n\right)$$

$$a_5^\kappa = \sum_{l=0}^{N} p'_{-2} H^{-2} Y_l + \hat{K}_N^\kappa (dM_{-2}) - \hat{H}(dM_{-2}) - \hat{H}(dM_{-2}, m_0)$$

$$- \hat{H}(dM_{-1}, m_{-1}) - \hat{K}_N (d_{-3}, dM_1) - \hat{K}_N (d_{-2}, dM_0) - \hat{K}_N (d_{-1}, dM_{-1})$$

$$- \hat{K}_N (d_0, dM_{-2}) - \hat{K}_N (p_{-2}) - \hat{K}_N (p_{-3}, m_1) - \hat{K}_N (p_{-2}, m_0) - \hat{K}_N (p_{-1}, m_{-1})$$

$$- \sum_{n=-3}^{1} \left(\sum_{l=0}^{N} p_{-2-n} H^{-2} d_n \right)$$

$$a_4^\kappa = \sum_{l=0}^{N} p'_{-3} H^{-2} Y_l - \hat{H}(dM_{-2}, m_{-1}) - \hat{K}_N (d_{-3}, dM_0) - \hat{K}_N (d_{-2}, dM_{-1})$$

$$- \hat{K}_N (d_{-1}, dM_{-2}) - \hat{K}_N (p_{-3}) - \hat{K}_N (p_{-4}, m_1) - \hat{K}_N (p_{-3}, m_0) - \hat{K}_N (p_{-2}, m_{-1})$$

$$- \sum_{n=-3}^{1} \left(\sum_{l=0}^{N} p_{-3-n} H^{-2} d_n \right)$$

$$a_3^\kappa = \sum_{l=0}^{N} p'_{-4} H^{-2} Y_l - \hat{K}_N (d_{-3}, dM_{-1}) - \hat{K}_N (d_{-2}, dM_{-2}) - \hat{K}_N (p_{-4}) - \hat{K}_N (p_{-4}, m_0)$$

$$- \hat{K}_N (p_{-3}, m_{-1}) - \sum_{n=0}^{3} \left(\sum_{l=0}^{N} p_{n-4} H^{-2} d_{-n} \right)$$

$$a_2^\kappa = -\hat{K}_N (d_{-3}, dM_{-2}) - \hat{K}_N (p_{-4}, m_{-1}) - \sum_{l=0}^{N} p_{-4} H^{-2} d_{-1} - \sum_{l=0}^{N} p_{-3} H^{-2} d_{-2} - \sum_{l=0}^{N} p_{-2} H^{-2} d_{-3}$$

$$a_1^\kappa = -\sum_{l=0}^{N} p_{-4} H^{-2} d_{-2} - \sum_{l=0}^{N} p_{-3} H^{-2} d_{-3}$$

$$a_0^\kappa = -\sum_{l=0}^{N} p_{-4} H^{-2} d_{-3}$$

From all above, we proved the EM update for each parameter.

Appendix B

Proof of Finite-dimensional Filters

In this section, we give a outline for deriving the filters given in Theorem 3.1-3.8. The proof is based on the work of Elliott and Krishnamurthy (1999) and Elliott and Hyndman (2007). We calculate the filters with the reference probability measure introduced in section (2.3) and apply Bayes' theorem. In the E-step of the EM algorithm, the $(k+1)$-th update of the parameter set $\hat{\theta}_{k+1}$ is based on the k-th update $\hat{\theta}_k$. Therefore, the Radon-Nikodym derivative Λ_k in the E-step is associated with $\hat{\theta}_k$ and used to construct the measure $P_{\hat{\theta}_k}$ from the reference probability $\bar{P}$. Let α_k, $\gamma_k^{ij(m)}$, q_k, and $\beta_k^{ij(n)}$ be the unnormalized (measured valued) densities

$$
\begin{aligned}
\alpha_k(x) &= \bar{E}\left[\Lambda_k I_{\{x_k \in dx\}} | \mathcal{Y}_k\right] \\
\gamma_k^{ij(n)} &= \bar{E}\left[\Lambda_k J_k^{ij(n)} I_{\{x_k \in dx\}} | \mathcal{Y}_k\right], m = 1, 2 \\
q_k(x) &= \bar{E}\left[\Lambda_k K_k I_{\{x_k \in dx\}} | \mathcal{Y}_k\right] \\
\beta_k^{ij(n)}(x) &= \bar{E}\left[\Lambda_k U_k^{ij(n)} I_{\{x \in dx\}} | \mathcal{Y}_k\right], n = 1, \ldots, 4 \\
\zeta_k^{i(n)}(x) &= \bar{E}\left[\Lambda_k L_k^{i(n)} I_{\{x \in dx\}} | \mathcal{Y}_k\right], n = 1, \ldots, 2, i = 1, \ldots, m \\
p_k^{ij(n)}(x) &= \bar{E}\left[\Lambda_k H_k^{ij(n)} I_{\{x \in dx\}} | \mathcal{Y}_k\right], n = 0, \ldots, 3.
\end{aligned}
$$

Then for any measurable test function $g : \mathcal{R}^m \to \mathcal{R}$

$$\bar{E}\left[\Lambda_k g(X_k)|Y_k\right] = \int_{\mathcal{R}^m} \alpha_k(x)\, g(x)\, dx \tag{B.1}$$

$$\bar{E}\left[\Lambda_k J_k^{ij(m)} g(X_k)|\mathcal{Y}_k\right] = \int_{\mathcal{R}^m} \gamma_k^{ij(m)}(x)\, g(x)\, dx, m = 1,2 \tag{B.2}$$

$$\bar{E}\left[\Lambda_k K_k g(X_k)|\mathcal{Y}_k\right] = \int_{\mathcal{R}^m} q_k(x)\, g(x)\, dx \tag{B.3}$$

$$\bar{E}\left[\Lambda_k U_k^{ij(n)} g(X_k)|\mathcal{Y}_k\right] = \int_{\mathcal{R}^m} \beta_k^{ij(n)}(x) g(x)\, dx, n = 1,\ldots,4 \tag{B.4}$$

$$\bar{E}\left[\Lambda_k L_k^{i(n)} g(X_k)|\mathcal{Y}_k\right] = \int_{\mathcal{R}^m} \zeta_k^{i(n)}(x) g(x)\, dx, n = 1,\ldots,2, \tag{B.5}$$

$$\bar{E}\left[\Lambda_k H_k^{ij(n)} g(X_k)|\mathcal{Y}_k\right] = \int_{\mathcal{R}^m} p_k^{ij(n)}(x) g(x)\, dx, n = 0,\ldots,3. \tag{B.6}$$

Similar to Theorem 4.1 of Elliott and Krishnamurthy (1999), the following recursions can be derived:

$$\beta_k^{ij(1)}(x) = \frac{\phi\left(H^{-1}(Y_k - M_k x - d_k)\right)}{|G||H|\phi(Y_k)} \int_{\mathcal{R}^m} \varphi\left(G^{-1}(x - Q_k z - c_k)\right) \beta_{k-1}^{ij(1)}(z)\, dz$$
$$+ < X, e_1 M_k^{(i,1)} >< X, e_1 M_k^{(j,1)} > \alpha_k(x) \tag{B.7}$$

$$\beta_k^{ij(2)}(x) = \frac{\phi\left(H^{-1}(Y_k - M_k x - d_k)\right)}{|G||H|\phi(Y_k)} \int_{\mathcal{R}^m} \varphi\left(G^{-1}(x - Q_k z - c_k)\right) \beta_{k-1}^{ij(2)}(z)\, dz$$
$$+ < X, e_1 M_k^{(i,1)} >< X, e_2 M_k^{(j,2)} > \alpha_k(x) \tag{B.8}$$

$$\beta_k^{ij(3)}(x) = \frac{\phi\left(H^{-1}(Y_k - M_k x - d_k)\right)}{|G||H|\phi(Y_k)} \int_{\mathcal{R}^m} \varphi\left(G^{-1}(x - Q_k z - c_k)\right) \beta_{k-1}^{ij(3)}(z)\, dz$$
$$+ < X, e_2 M_k^{(i,2)} >< X, e_1 M_k^{(j,1)} > \alpha_k(x) \tag{B.9}$$

$$\beta_k^{ij(4)}(x) = \frac{\phi\left(H^{-1}(Y_k - M_k x - d_k)\right)}{|G||H|\phi(Y_k)} \int_{\mathcal{R}^m} \varphi\left(G^{-1}(x - Q_k z - c_k)\right) \beta_{k-1}^{ij(4)}(z)\, dz$$
$$+ < X, e_2 M_k^{(i,2)} >< X, e_2 M_k^{(j,2)} > \alpha_k(x) \tag{B.10}$$

$$\gamma_k^{ij(1)}(x) = \frac{\phi\left(H^{-1}(Y_k - M_k x - d_k)\right)}{|G||H|\phi(Y_k)} \int_{\mathcal{R}^m} \varphi\left(G^{-1}(x - Q_k z - c_k)\right) \gamma_{k-1}^{ij(1)}(z)\, dz$$
$$+ < x, e_1 M_k^{(i,1)} \left(Y_k^{(j)} - d_k^{(j)}\right) > \alpha_k(x) \tag{B.11}$$

$$\gamma_k^{ij(2)}(x) = \frac{\phi\left(H^{-1}\left(Y_k - M_k x - d_k\right)\right)}{|G||H|\phi\left(Y_k\right)} \int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Q_k z - c_k\right)\right) \gamma_{k-1}^{ij(2)}(z)\, dz$$
$$+ < X, e_2 M_k^{(i,2)}\left(Y_k^{(j)} - d_k^{(j)}\right) > \alpha_k(x) \tag{B.12}$$

$$\alpha_k(x) = \frac{\phi\left(H^{-1}\left(Y_k - M_k x - d_k\right)\right)}{|G||H|\phi\left(Y_k\right)} \int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Q_k z - c_k\right)\right) \alpha_{k-1}(z)\, dz \tag{B.13}$$

$$q_k(x) = \frac{\phi\left(H^{-1}\left(Y_k - M_k x - d_k\right)\right)}{|G||H|\phi\left(Y_k\right)} \int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Q_k z - c_k\right)\right) q_{k-1}(z)\, dz + f_k^{\prime} H^{-2} M_k x \alpha_k(x) \tag{B.14}$$

$$\zeta_k^{i(1)}(x) = \frac{\phi\left(H^{-1}\left(Y_k - M_k x - d_k\right)\right)}{|G||H|\phi\left(Y_k\right)} \Bigg[\int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Q_k z - c_k\right)\right) \beta_{k-1}^{i(1)}(z)\, dz$$
$$\int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Qz - c\right)\right) < z, e_i > \alpha_{k-1}(z)\, dz\Bigg] \tag{B.15}$$

$$\zeta_k^{i(2)}(x) = \frac{\phi\left(H^{-1}\left(Y_k - M_k x - d_k\right)\right)}{|G||H|\phi\left(Y_k\right)} \int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Q_k z - c_k\right)\right) \zeta_{k-1}^{i(2)}(z)\, dz$$
$$+ < x, e_i > \alpha_k(x) \tag{B.16}$$

$$p_k^{ij(0)}(x) = \frac{\phi\left(H^{-1}\left(Y_k - M_k x - d_k\right)\right)}{|G||H|\phi\left(Y_k\right)} \int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Q_k z - c_k\right)\right) p_{k-1}^{ij(0)}(z)\, dz$$
$$+ \overline{f}_l^{\prime} H^{-2} \overline{g}_l < x, e_i >< x, e_j > \alpha_k(x) \tag{B.17}$$

$$p_k^{ij(1)}(x) = \frac{\phi\left(H^{-1}\left(Y_k - M_k x - d_k\right)\right)}{|G||H|\phi\left(Y_k\right)} \Bigg[\int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Q_k z - c_k\right)\right) p_{k-1}^{ij(1)}(z)\, dz$$
$$+ < x, e_i > \int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Qz - c\right)\right) < z, e_j > \alpha_{k-1}(z)\, dz\Bigg] \tag{B.18}$$

$$p_k^{ij(2)}(x) = \frac{\phi\left(H^{-1}\left(Y_k - M_k x - d_k\right)\right)}{|G||H|\phi\left(Y_k\right)} \Bigg[\int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Q_k z - c_k\right)\right) p_{k-1}^{ij(2)}(z)\, dz$$
$$+ \int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Qz - c\right)\right) < z, e_i >< z, e_j > \alpha_{k-1}(z)\, dz\Bigg] \tag{B.19}$$

$$p_k^{ij(3)}(x) = \frac{\phi\left(H^{-1}\left(Y_k - M_k x - d_k\right)\right)}{|G||H|\phi\left(Y_k\right)} \int_{\mathcal{R}^m} \varphi\left(G^{-1}\left(x - Q_k z - c_k\right)\right) p_{k-1}^{ij(3)}(z)\, dz$$
$$+ < x, e_i >< x, e_j > \alpha_k(x) \tag{B.20}$$

For brevity, we only prove (B.7) as others have similar procedure. we can write

$$U_k^{ij(1)} = U_{k-1}^{ij(1)} + < X_k, e_1 M_k^{(i,1)} >< X_k, e_1 M_k^{(j,1)} >$$

Then using the definition of Λ_k in equation (2.14), we can write equation (B.4) as

$$\begin{aligned}
&\bar{E}\left[\Lambda_k U_k^{ij(1)} g(X_k) | \mathcal{Y}_k\right] \\
=&\bar{E}\left[\Lambda_{k-1} \frac{\phi\left(H^{-1}(Y_k - M_k X_k - d_k)\right)}{|H|\phi(Y_k)} \frac{\varphi\left(G^{-1}(X_k - Q_k X_{k-1} - c_k)\right)}{|G|\varphi(X_k)} \times U_{k-1}^{ij(1)} g(X_k) | \mathcal{Y}_k\right] \\
&+\bar{E}\left[\Lambda_{k-1} \frac{\phi\left(H^{-1}(Y_k - M_k X_k - d_k)\right)}{|H|\phi(Y_k)} \frac{\varphi\left(G^{-1}(X_k - Q_k X_{k-1} - c_k)\right)}{|G|\varphi(X_k)}\right. \\
&\left.\times < X_k, e_1 M_k^{(i,1)} >< X_k, e_1 M_k^{(j,1)} > g(X_k) | \mathcal{Y}_k\right] \\
=&\frac{1}{|H||G|\phi(Y_k)}\left[\bar{E}\left[\Lambda_{k-1} U_{k-1}^{ij(1)} \int_{\mathcal{R}^m} \phi\left(H^{-1}(Y_k - M_k x - d_k)\right) \varphi\left(G^{-1}(X_k - Q_k X_{k-1} - c_k)\right)\right.\right. \\
&g(x)\,dx|\mathcal{Y}_k] + \bar{E}\left[\Lambda_{k-1} \int_{\mathcal{R}^m} \phi\left(H^{-1}(Y_k - M_k x - d_k)\right) \varphi\left(G^{-1}(X_k - Q_k X_{k-1} - c_k)\right)\right. \\
&\left.\left. < x, e_1 M_k^{(i,1)} >< x, e_1 M_k^{(j,1)} > g(x)\,dx|\mathcal{Y}_k\right]\right] \\
=&\frac{1}{|H||G|\phi(Y_k)}\left[\int_{\mathcal{R}^m}\int_{\mathcal{R}^m} \beta_{k-1}^{ij(1)}(z) \phi\left(H^{-1}(Y_k - M_k x - d_k)\right) \varphi\left(G^{-1}(x - Q_k z - c_k)\right) g(x)\,dxdz\right. \\
&+\int_{\mathcal{R}^m}\int_{\mathcal{R}^m} \alpha_{k-1}(z) \phi\left(H^{-1}(Y_k - M_k x - d_k)\right) \varphi\left(G^{-1}(x - Q_k z - c_k)\right) \\
&\left. < x, e_1 M_k^{(i,1)} >< x, e_1 M_k^{(j,1)} > g(x)\,dxdz\right]
\end{aligned}$$

From the definition of $\beta_k^{ij(1)}(x)$ in equation (B.4), it is easy to conclude that

$$\begin{aligned}
\beta_k^{ij(1)}(x) =& \frac{\phi\left(H^{-1}(Y_k - M_k x - d_k)\right)}{|H||G|\phi(Y_k)} \int_{\mathcal{R}^m} \beta_{k-1}^{ij(1)}(z) \varphi\left(G^{-1}(x - Q_k z - c_k)\right) dz \\
&+ \frac{\phi\left(H^{-1}(Y_k - M_k x - d_k)\right)}{|H||G|\phi(Y_k)} \int_{\mathcal{R}^m} \alpha_{k-1}(z) \varphi\left(G^{-1}(x - Q_k z - c_k)\right) \\
&\times < X_k, e_1 M_k^{(i,1)} >< X_k, e_1 M_k^{(j,1)} > dz \\
=& \frac{\phi\left(H^{-1}(Y_k - M_k x - d_k)\right)}{|H||G|\phi(Y_k)} \int_{R^m_{\mathcal{R}^m}} \beta_{k-1}^{ij(1)}(z) \varphi\left(G^{-1}(x - Q_k z - c_k)\right) dz \\
&+ < X_k, e_1 M_k^{(i,1)} >< X_k, e_1 M_k^{(j,1)} > \alpha_k(x)
\end{aligned}$$

Similarly, we can prove the other recursions.

Note that when k=0, we have the initial conditions as follows:

$$\alpha_0(x) = \frac{\phi\left(H^{-1}(Y_0 - M_0x - d_0)\right)}{|H|\phi(Y_0)}\varphi(x)$$

$$\beta_0^{ij(1)}(x) =< x, e_1M_0^{(i,1)} >< x, e_1M_0^{(j,1)} > \alpha_0(x) \tag{B.21}$$

$$\beta_0^{ij(2)}(x) =< x, e_1M_0^{(i,1)} >< x, e_2M_0^{(j,2)} > \alpha_0(x)$$

$$\beta_0^{ij(3)}(x) =< x, e_2M_0^{(i,2)} >< x, e_1M_0^{(j,1)} > \alpha_0(x)$$

$$\beta_0^{ij(4)}(x) =< x, e_2M_0^{(i,2)} >< x, e_2M_0^{(j,2)} > \alpha_0(x)$$

$$\gamma_0^{ij(1)}(x) =< x, e_1M_0^{(i,1)}\left(Y_0^{(j)} - d_0^{(j)}\right) > \alpha_0(x)$$

$$\gamma_0^{ij(2)}(x) =< x, e_2M_0^{(i,2)}\left(Y_0^{(j)} - d_0^{(j)}\right) > \alpha_0(x)$$

$$q_0(x) = f_0'H^{-2}M_0x\alpha_0(x)$$

$$\zeta_0^{i(1)} = 0$$

$$\zeta_0^{i(2)} = 0$$

$$p_0^{ij(0)} = \vec{f}_l'H^{-2}\overline{g}_l < x, e_i >< x, e_j > \alpha_0(x)$$

$$p_0^{ij(1)} = 0$$

$$p_0^{ij(2)} = 0$$

$$p_0^{ij(3)} = 0$$

(B.22)

Using these recursions, we are able to give finite-dimensional sufficient statistics for the densities (B.7) to (B.14). Define

$$\sigma_k = P_{k-1}^{-1} + Q_l' G^{-2} Q_l$$

$$\Sigma_k = G^{-2} Q_l \sigma_k^{-1}$$

$$S_k = \sigma_{k+1}^{-1} \left(P_k^{-1} \mu_k - Q_l' G^{-2} c_l \right)$$

Similar to Theorem 5.2 of Elliott and Krishnamurthy (1999), we can prove

$$\beta_k^{ij(n)}(x) = \left[a_k^{ij(n)} + \left(b_k^{ij(n)} \right)' x + x' d_k^{ij(n)} x \right] \alpha_k(x), n = 1, \ldots, 4 \tag{B.23}$$

where $a_{k+1}^{ij(n)} \in \mathcal{R}$, $b_{k+1}^{ij(n)} \in \mathcal{R}^m$ and $d_{k+1}^{ij(n)} \in \mathcal{R}^{m\times m}$, a symmetric matrix satisfy the following recursions

$$a_{k+1}^{ij(1)} = a_k^{ij(1)} + \left(b_k^{ij(1)}\right)' S_k + Tr\left[d_k^{ij(1)} \sigma_{k+1}^{-1}\right] + S_k' d_k^{ij(1)} S_k, a_0^{ij(1)} = 0$$

$$b_{k+1}^{ij(1)} = \Sigma_{k+1}\left(b_k^{ij(1)} + 2d_k^{ij(1)} S_k\right), b_0^{ij(1)} = 0$$

$$d_{k+1}^{ij(1)} = \Sigma_{k+1} d_k^{ij(1)} \Sigma_{k+1}' + \frac{1}{2}\left(M_{k+1}^{(i,1)} e_1 e_1' M_{k+1}^{(j,1)} + M_{k+1}^{(j,1)} e_1 e_1' M_{k+1}^{(i,1)}\right)$$

$$d_0^{ij(1)} = \frac{1}{2}\left(M_0^{(i,1)} e_1 e_1' M_0^{(j,1)} + M_0^{(j,1)} e_1 e_1' M_0^{(i,1)}\right)$$

$$a_{k+1}^{ij(2)} = a_k^{ij(2)} + \left(b_k^{ij(2)}\right)' S_k + Tr\left[d_k^{ij(2)} \sigma_{k+1}^{-1}\right] + S_k' d_k^{ij(2)} S_k, a_0^{ij(2)} = 0$$

$$b_{k+1}^{ij(2)} = \Sigma_{k+1}\left(b_k^{ij(2)} + 2d_k^{ij(2)} S_k\right), b_0^{ij(2)} = 0$$

$$d_{k+1}^{ij(2)} = \Sigma_{k+1} d_k^{ij(2)} \Sigma_{k+1}' + \frac{1}{2}\left(M_{k+1}^{(i,1)} e_1 e_2' M_{k+1}^{(j,2)} + M_{k+1}^{(j,2)} e_2 e_1' M_{k+1}^{(i,1)}\right)$$

$$d_0^{ij(2)} = \frac{1}{2}\left(M_0^{(i,1)} e_1 e_2' M_0^{(j,2)} + M_0^{(j,2)} e_2 e_1' M_0^{(i,1)}\right)$$

$$a_{k+1}^{ij(3)} = a_k^{ij(3)} + \left(b_k^{ij(3)}\right)' S_k + Tr\left[d_k^{ij(3)} \sigma_{k+1}^{-1}\right] + S_k' d_k^{ij(3)} S_k, a_0^{ij(3)} = 0$$

$$b_{k+1}^{ij(3)} = \Sigma_{k+1}\left(b_k^{ij(3)} + 2d_k^{ij(3)} S_k\right), b_0^{ij(3)} = 0$$

$$d_{k+1}^{ij(3)} = \Sigma_{k+1} d_k^{ij(3)} \Sigma_{k+1}' + \frac{1}{2}\left(M_{k+1}^{(i,2)} e_2 e_1' M_{k+1}^{(j,1)} + M_{k+1}^{(j,1)} e_1 e_2' M_{k+1}^{(i,2)}\right)$$

$$d_0^{ij(3)} = \frac{1}{2}\left(M_0^{(i,2)} e_2 e_1' M_0^{(j,1)} + M_0^{(j,1)} e_1 e_2' M_0^{(i,2)}\right)$$

$$a_{k+1}^{ij(4)} = a_k^{ij(4)} + \left(b_k^{ij(4)}\right)' S_k + Tr\left[d_k^{ij(4)} \sigma_{k+1}^{-1}\right] + S_k' d_k^{ij(4)} S_k, a_0^{ij(4)} = 0$$

$$b_{k+1}^{ij(4)} = \Sigma_{k+1}\left(b_k^{ij(4)} + 2d_k^{ij(1)} S_k\right), b_0^{ij(4)} = 0$$

$$d_{k+1}^{ij(4)} = \Sigma_{k+1} d_k^{ij(4)} \Sigma_{k+1}' + \frac{1}{2}\left(M_{k+1}^{(i,2)} e_2 e_2' M_{k+1}^{(j,2)} + M_{k+1}^{(j,2)} e_2 e_2' M_{k+1}^{(i,2)}\right)$$

$$d_0^{ij(4)} = \frac{1}{2}\left(M_0^{(i,2)} e_2 e_2' M_0^{(j,2)} + M_0^{(j,2)} e_2 e_2' M_0^{(i,2)}\right)$$

Also, we can prove

$$\gamma_{k+1}^{ij(n)}(x) = \left[\bar{a}_k^{ij(n)} + \left(\bar{b}_k^{ij(n)}\right)' x\right] \alpha_k(x), n = 1, 2$$

where $\bar{a}_{k+1}^{ij(n)}$ and $\bar{b}_{k+1}^{ij(n)}$ satisfy the following recursions

$$\bar{a}_{k+1}^{ij(1)} = \bar{a}_k^{ij(1)} + \left(\bar{b}_k^{ij(1)}\right)' S_k, \bar{a}_0^{ij(1)} = 0$$

$$\bar{b}_{k+1}^{ij(1)} = \Sigma_{k+1}\bar{b}_k^{ij(1)} + e_1 M_{k+1}^{(i,1)} \left(Y_{k+1}^{(j)} - d_{k+1}^{ij(j)}\right), \bar{b}_0^{ij(1)} = e_1 M_0^{(i,1)} \left(Y_0^{(j)} - d_0^{(j)}\right) \in \mathcal{R}^m$$

$$\bar{a}_{k+1}^{ij(2)} = \bar{a}_k^{ij(2)} + \left(\bar{b}_k^{ij(2)}\right)' S_k, \bar{a}_0^{ij(2)} = 0$$

$$\bar{b}_{k+1}^{ij(2)} = \Sigma_{k+1}\bar{b}_k^{ij(2)} + e_2 M_{k+1}^{(i,2)} \left(Y_{k+1}^{(j)} - d_{k+1}^{ij(j)}\right), \bar{b}_0^{ij(1)} = e_2 M_0^{(i,2)} \left(Y_0^{(j)} - d_0^{(j)}\right) \in \mathcal{R}^m$$

Next we have that

$$q_k(x) = \left[\bar{r}_k + (\bar{s}_k)' x\right] \alpha_k(x)$$

where $\bar{r}_{k+1}$ and $\bar{s}_{k+1}$ satisfy the recursions

$$\bar{r}_{k+1} = \bar{r}_k + (\bar{s}_k)' S_k, \bar{s}_0 = 0$$

$$\bar{s}_{k+1} = \Sigma_{k+1}\bar{s}_k + M_{k+1}' H^{-2} f_{k+1}, \bar{s}_0 = M_0' H^{-2} f_0 \in \mathcal{R}^m$$

Similarly we can show that

$$\zeta_k^{i(n)} = \left[u_k^{i(n)} + \left(v_k^{i(n)}\right)' x\right] \alpha_k(x) \tag{B.24}$$

(B.25)

where $u_{k+1}^{i(n)}$ and $v_{k+1}^{i(n)}$ satisfy the recursions

$$u_{k+1}^{i(1)} = u_k^{i(1)} + \left(v_k^{i(1)} + e_i\right)' S_k, \quad u_0^{i(1)} = 0, \tag{B.26}$$

$$v_{k+1}^{i(1)} = \Sigma_k \left(v_k^{i(1)} + e_i\right), \quad v_0^{i(1)} = \mathbf{0} \in \mathcal{R}^m, \tag{B.27}$$

$$u_{k+1}^{i(2)} = u_k^{i(2)} + \left(v_k^{i(2)}\right)' S_k, \quad u_0^{i(2)} = 0, \tag{B.28}$$

$$v_{k+1}^{i(2)} = \Sigma_k v_k^{i(2)} + e_i, \quad v_0^{i(2)} = \mathbf{0} \in \mathcal{R}^m, \tag{B.29}$$

(B.30)

Finally, we can show that

$$p_k^{ij(n)}(x) = \left[\bar{u}_k^{ij(n)} + \left(\bar{v}_k^{ij(n)}\right)' x + x' \bar{d}_k^{ij(n)} x\right] \alpha_k(x), n = 1, \ldots, 4 \tag{B.31}$$

where $\bar{u}_k^{ij(n)}$, $\bar{v}_k^{ij(n)}$ and $\bar{d}_k^{ij(n)}$ satisfy the recursions

$$\bar{u}_{k+1}^{ij(0)} = \bar{u}_k^{ij(0)} + \left(\bar{v}_k^{ij(0)}\right)' S_k + Tr\left[\bar{d}_k^{ij(0)} \sigma_{k+1}^{-1}\right] + S_k' \bar{d}_k^{ij(n)} S_k, \bar{u}_0^{ij(0)} = 0, \tag{B.32}$$

$$\bar{v}_{k+1}^{ij(0)} = \Sigma_{k+1}\left(\bar{v}_k^{ij(0)} + 2\bar{d}_k^{ij(0)} S_k\right), \bar{v}_0^{ij(0)} = 0 \in \mathcal{R}^m \tag{B.33}$$

$$\bar{d}_{k+1}^{ij(0)} = \Sigma_{k+1} \bar{d}_k^{ij(0)} (\Sigma_{k+1})' + \frac{1}{2} \bar{f}_l H^{-2} \bar{g}_l \left(e_i e_j' + e_j e_i'\right), \tag{B.34}$$

$$\bar{u}_{k+1}^{ij(1)} = \bar{u}_k^{ij(1)} + \left(\bar{v}_k^{ij(1)}\right)' S_k + Tr\left[\bar{d}_k^{ij(1)} \sigma_{k+1}^{-1}\right] + S_k' \bar{d}_k^{ij(n)} S_k, \bar{u}_0^{ij(1)} = 0, \tag{B.35}$$

$$\bar{v}_{k+1}^{ij(1)} = \Sigma_{k+1}\left(\bar{v}_k^{ij(1)} + 2\bar{d}_k^{ij(1)} S_k\right) + e_i e_j' S_k, \bar{v}_0^{ij(1)} = \mathbf{0} \in \mathcal{R}^m, \tag{B.36}$$

$$\bar{d}_{k+1}^{ij(1)} = \Sigma_{k+1} \bar{d}_k^{ij(1)} \Sigma_k' + \frac{1}{2}\left(e_i e_j \Sigma_k' + \Sigma_k e_j e_i'\right), \bar{d}_0^{ij(1)} = \mathbf{0} \in \mathcal{R}^{m \times m} \tag{B.37}$$

$$\bar{u}_{k+1}^{ij(2)} = \bar{u}_k^{ij(2)} + \left(\bar{v}_k^{ij(2)}\right)' S_k + Tr\left[\bar{d}_k^{ij(2)} \sigma_{k+1}^{-1}\right] + S_k' \left(\bar{d}_k^{ij(2)} + e_i e_j'\right) S_k + Tr\left[e_i e_j' \sigma_{k+1}^{-1}\right], \tag{B.38}$$

$$\bar{u}_0^{ij(2)} = 0, \tag{B.39}$$

$$\bar{v}_{k+1}^{ij(2)} = \Sigma_{k+1}\left(\bar{v}_k^{ij(2)} + \left(2\bar{d}_k^{ij(2)} + e_i e_j' + e_j e_i'\right) S_k\right), \bar{v}_0^{ij(2)} = \mathbf{0} \in \mathcal{R}^m, \tag{B.40}$$

$$\bar{d}_{k+1}^{ij(2)} = \Sigma_{k+1}\left(\bar{d}_k^{ij(2)} + \frac{1}{2}\left[e_i e_j' + e_j e_i'\right]\right) (\Sigma_{k+1})', \bar{d}_0^{ij(2)} = \mathbf{0} \in \mathcal{R}^{m \times m} \tag{B.41}$$

$$\bar{u}_{k+1}^{ij(3)} = \bar{u}_k^{ij(3)} + \left(\bar{v}_k^{ij(3)}\right)' S_k + Tr\left[\bar{d}_k^{ij(3)} \sigma_{k+1}^{-1}\right] + S_k' \bar{d}_k^{ij(n)} S_k, \bar{u}_0^{ij(3)} = 0, \tag{B.42}$$

$$\bar{v}_{k+1}^{ij(3)} = \Sigma_{k+1}\left(\bar{v}_k^{ij(3)} + 2\bar{d}_k^{ij(3)} S_k\right), b_0^{ij(3)} = \mathbf{0} \in \mathcal{R}^m \tag{B.43}$$

$$\bar{d}_{k+1}^{ij(3)} = \Sigma_{k+1} \bar{d}_k^{ij(3)} (\Sigma_{k+1})' + \frac{1}{2}\left(e_i e_j' + e_j' e_i\right), \bar{d}_0^{ij(3)} = \mathbf{0} \in \mathcal{R}^{m \times m} \tag{B.44}$$

Proof. We only prove equation (B.23) for $n = 1$. For convenience , we drop the superscript i and j. From equation (B.21), we have

$$\beta_0^{(1)}(x) = < x, e_1 M_0^{(i,1)} > < x, e_1 M_0^{(j,1)} > \alpha_0(x)$$

Therefore, we have the initial condition for a_0, b_0 and d_0 given by equation (B.45)-(B.47).

$$a_0 = 0 \in \mathcal{R} \tag{B.45}$$

$$b_0 = \mathbf{0} \in \mathcal{R}^m \tag{B.46}$$

$$d_0 = \frac{1}{2}\left(e_1 M_0^{(i,1)} M_0^{(j,1)'} e_1' + e_1 M_0^{(j,1)} M_0^{(i,1)} e_1'\right) \in \mathcal{R}^{m\times m} \tag{B.47}$$

Now assume equation (B.23) holds for $n = 1$ at time k. Then at time $k+1$, using equation (B.23) for $n = 1$ and recursion (B.7), we get,

$$\begin{aligned}\beta_{k+1}^{(1)}(x) =& \frac{\phi\left(H^{-1}\left(Y_{k+1} - M_{k+1}x - d_{k+1}\right)\right)}{|G||H|\phi\left(Y_{k+1}\right)} \times \\ & \int_{\mathcal{R}^2} \varphi\left(G^{-1}\left(x - Q_{k+1}z - c_{k+1}\right)\right)\left(a_k + b_k' z + z' d_k z\right)\alpha_k(z)\,dz \\ & + < x, e_1 M_{k+1}^{(i,j)} >< x, e_1 M_{k+1}^{(j,1)} > \alpha_{k+1}(x)\end{aligned} \tag{B.48}$$

Let us first concentrate on the first term of equation (B.48).

Denote

$$\begin{aligned}I_1 =& \frac{\phi\left(H^{-1}\left(Y_{k+1} - M_{k+1}x - d_{k+1}\right)\right)}{|G||H|\phi\left(Y_{k+1}\right)} \int_{\mathcal{R}^2} \varphi\left(G^{-1}\left(x - Q_{k+1}z - c_{k+1}\right)\right)\left(a_k + b_k' z + z' d_k z\right) \\ & \times \alpha_k(z)\,dz\end{aligned} \tag{B.49}$$

From the definition of φ and $\alpha_k(x)$, we can rewrite equation (B.49) as

$$I_1 = T(x)\int_{\mathcal{R}^m} exp\left[-\frac{1}{2}(x - Qz - c)' G^{-2}(x - Qz - c) - \frac{1}{2}(z - \mu_k)' P_k^{-1}(z - \mu_k)\right] \tag{B.50}$$

$$\times\left(a_k + b_k' z + z' d_k z\right) dz \tag{B.51}$$

where

$$T(x) = \frac{\phi\left(H^{-1}\left(Y_{k+1} - M_{k+1}x - dk + 1\right)\right)}{|G||H|\phi\left(Y_{k+1}\right)}(2\pi)^{-m}|G|^{-1}|P_k|^{-\frac{1}{2}}\bar{\alpha}_k. \tag{B.52}$$

Therefore, I_1 is expanded as

$$I_1 = T(x)\int_{\mathcal{R}^m} exp\left[-\frac{1}{2}\left(x^{'}G^{-2}x - x^{'}G^{-2}Qz - x^{'}G^{-2}c - z^{'}Q^{'}G^{-2}x + z^{'}Q^{'}G^{-2}Qz\right.\right.$$
$$\left.\left.+z^{'}Q^{'}G^{-2}c - c^{'}G^{-2}x + c^{'}G^{-2}Qz + c^{'}G^{-2}c + z^{'}P_k^{-1}z - z^{'}P_k^{-1}\mu_k - \mu^{'}P_k^{-1}z + \mu_k^{'}P_k^{-1}\mu_k\right)\right] \tag{B.53}$$

$$\times\left(a_k + b_k^{'}z + z^{'}d_kz\right)dz$$
$$= T(x)\int_{\mathcal{R}^m} exp\left[-\frac{1}{2}\left[\left(x^{\prime}G^{-2}x + \mu_k^{'}P_k^{-1}\mu_k + c^{\prime}G^{-2}c - 2x^{\prime}G^{-2}c\right) + z^{\prime}\left(Q^{\prime}G^{-2}Q + P_k^{-1}\right)z\right.\right.$$
$$\left.\left.-2\left(x^{\prime}G^{-2}Q - c^{\prime}G^{-2}Q + \mu_kP_k^{-1}\right)z\right]\right] \times \left(a_k + b_k^{'}z + z^{'}d_kz\right)dz \tag{B.54}$$

Denote

$$\delta_{k+1} = 2\left(x^{\prime}G^{-2}Q - c^{\prime}G^{-2}Q + \mu_k^{'}P_k^{-1}\right)^{'} \tag{B.55}$$

$$T_1(x) = T(x)\,exp\left[-\frac{1}{2}\left(x^{\prime}G^{-2}x + \mu_k^{'}P_k^{-1}\mu_k + c^{\prime}G^{-2}c - 2x^{\prime}G^{-2}c\right)\right] \tag{B.56}$$

Then, I_1 can be rewritten as

$$I_1 = T_1(x)\int_{\mathcal{R}^m} exp\left[-\frac{1}{2}\left(z^{\prime}\sigma_{k+1}z - \delta_{k+1}z\right)\right] \times \left(a_k + b_k^{'}z + z^{\prime}d_kz\right)dz \tag{B.57}$$

Completing the square in equation (B.57) yields

$$I_1 = T_1(x)\,exp\left[-\frac{1}{8}\delta_{k+1}^{'}\sigma_{k+1}^{'}\delta_{k+1}\right]\int_{\mathcal{R}^m} exp\left[-\frac{1}{2}\left(z - \frac{\sigma_{k+1}^{-1}\delta_{k+1}}{2}\right)^{\prime}\sigma_{k+1}\left(z - \frac{\sigma_{k+1}^{-1}\delta_{k+1}}{2}\right)\right] \tag{B.58}$$

$$\times\left(a_k + b_k^{'}z + z^{\prime}d_kz\right)dz. \tag{B.59}$$

Now let us focus on the integral in equation (B.59).

Note that $(2\pi)^{-\frac{m}{2}}\,|\sigma_{k+1}|^{\frac{1}{2}}exp\left[\left(z - \frac{\sigma_{k+1}^{-1}\delta_{k+1}}{2}\right)^{\prime}\sigma_{k+1}\left(z - \frac{\sigma_{k+1}^{-1}\delta_{k+1}}{2}\right)\right]$ is the unnormalized

Gaussian density in z with $E(z) = \frac{\sigma_{k+1}^{-1}\delta_{k+1}}{2}$ Therefore

$$\int_{\mathcal{R}^m} \left(a_k + b_k' z + z' d_k z\right) exp\left[-\frac{1}{2}\left(z - \frac{\sigma_{k+1}^{-1}\delta_{k+1}}{2}\right)' \sigma_{k+1}\left(z - \frac{\sigma_{k+1}^{-1}\delta_{k+1}}{2}\right)\right] dz \quad \text{(B.60)}$$

$$= (2\pi)^{\frac{m}{2}} |\sigma_{k+1}|^{-\frac{1}{2}} \left(a_k + b_k' E(z) + E\left(z' d_k z\right)\right) \quad \text{(B.61)}$$

So

$$E\left(z' d_k z\right) = E\left[(z - E(z))' d_k (z - E(z))\right] + E\left(z'\right) d_k E(z)$$

$$= Tr\left[d_k \sigma_{k+1}^{-1}\right] + \frac{1}{4}\left(\sigma_{k+1}^{-1}\delta_{k+1}\right)' d_k \left(\sigma_{k+1}^{-1}\delta_{k+1}\right) \quad \text{(B.62)}$$

From equation (B.49) to equation (B.62), we obtain

$$\beta_{k+1}^{(1)}(x) = T_1(x) exp\left[\frac{1}{8}\delta_{k+1}' \sigma_{k+1}\delta_{k+1}\right](2\pi)^{\frac{m}{2}} |\sigma_{k+1}|^{-\frac{1}{2}} \quad \text{(B.63)}$$

$$\left(a_k + b_k' \frac{\sigma_{k+1}^{-1}\delta_{k+1}}{2} + Tr\left[d_k \sigma_{k+1}^{-1}\right] + \frac{1}{4}\left(\sigma_{k+1}^{-1}\delta_{k+1}\right)' d_k \left(\sigma_{k+1}^{-1}\delta_{k+1}\right)\right) \quad \text{(B.64)}$$

$$+ < x, e_1 M_{k+1}^{(i,1)} >< x, e_1 M_{k+1}^{(j,1)} > \alpha_{k+1}(x). \quad \text{(B.65)}$$

It can be shown that using the similar procedure above, that

$$\alpha_{k+1}(x) = T_1(x) exp\left[\frac{1}{8}\delta_{k+1}' \sigma_{k+1}\delta_{k+1}\right](2\pi)^{\frac{m}{2}} |\sigma_{k+1}|^{-\frac{1}{2}} \quad \text{(B.66)}$$

Thus

$$\beta_{k+1}^{(1)}(x) = \alpha_{k+1}(x)\left(a_k^{(1)} + b_k^{(1)'} \frac{\sigma_{k+1}^{-1}\delta_{k+1}}{2} + Tr\left[d_k^{(1)}\sigma_{k+1}^{-1}\right] + \frac{1}{4}\delta_{k+1}' \sigma_{k+1}^{-1} d_k \sigma_{k+1}^{-1}\delta_{k+1}\right.$$

$$\left. + x' M_{k+1}^{(i,1)} e_1 e_1' M_{k+1}^{(j,1)} x\right)$$

Substituting for δ_{k+1}, we can obtain

$$\beta_{k+1}^{(1)}(x) = \alpha_{k+1}(x)\left[a_{k+1}^{(1)} + b_{k+1}^{(1)'} x + x' d_{k+1}^{(1)} x\right]$$

where

$$
\begin{aligned}
a_{k+1}^{(1)} &= a_k^{(1)} + \left(b_k^{(1)}\right)' S_k + Tr\left[d_k^{(1)} \sigma_{k+1}^{-1}\right] + S_k' d_k^{(1)} S_k, \\
b_{k+1}^{(1)} &= \Sigma_{k+1}\left(b_k^{(1)} + 2 d_k^{(1)} S_k\right) \\
d_{k+1}^{(1)} &= \Sigma_{k+1} d_k^{(1)} \Sigma_{k+1}' + \frac{1}{2}\left(M_{k+1}^{(i,1)} e_1 e_1' M_{k+1}^{(j,1)} + M_{k+1}^{(j,1)} e_1 e_1' M_{k+1}^{(i,1)}\right)
\end{aligned}
$$

Similarly, we can prove for $\beta_{k+1}^{ij(n)}(x)$ n=2,3,4 , $q_k(x)$ and $\gamma_{k+1}^{ij(m)}(x)$, m=1,2. The proofs of recursion (B.23) for $n = 2,3,4$, (B.2), (B.1) and (B.3) are very similar with the above one. Hence we omitted them. Once we get the finite sufficient statistics for densities $\beta_k^{ij(n)}(x), n = 1,2,3,4, \gamma_k(x), q_k(x)$, we can finally obtain finite-dimensional filters for $J_k^{ij(n)}$, K_k and $U_k^{ij(n)}, n = 1,2,3,4$. Applying the general version of Bayes' theorem, we complete the the proof of Theorem 3.9. For example, using Bayes' theorem, we can write

$$
E_{\hat{\theta}_j}\left[U_k^{ij(n)} | \mathcal{Y}_k\right] = \frac{\bar{E}\left[\Lambda_k U_k^{ij(n)} | \mathcal{Y}_k\right]}{E\left[\Lambda_k | \mathcal{Y}_k\right]} = \frac{\int_{\mathcal{R}^m} \beta_k^{ij(n)}(x)\, dx}{\bar{\alpha}_k} \tag{B.67}
$$

where $\bar{\alpha}_k = \int_{\mathcal{R}^m} \alpha_k(x)\, dx$. Since

$$
\int_{\mathcal{R}^m} \beta_k^{ij(n)}(x)\, dx \tag{B.68}
$$

$$
= \bar{\alpha}_k E\left[a_k^{ij(n)} + \left(b_k^{ij(n)}\right)' x + x' d_k^{ij(n)} x\right] \tag{B.69}
$$

$$
= \bar{\alpha}_k E\left[a_k^{ij(n)} + \left(b_k^{ij(n)}\right)' x + x' d_k^{ij(n)} x\right] \tag{B.70}
$$

$$
= \bar{\alpha}_k \left[a_k^{ij(n)} + \left(b_k^{ij(n)}\right)' \mu_k + Tr\left[d_k^{ij(n)} P_k\right] + \mu_k' d_k^{ij(n)} \mu\right]. \tag{B.71}
$$

We can get equation (3.38) by simply substituting in equation (B.67). The proofs of equation (3.37) and (3.36) are similar. □

Therefore, we proved the finite-dimensional filters in M-step of the EM algorithm, which enables us to implement the filter-based EM algorithm.

Bibliography

B. D. Anderson and J. B. Moore. *Optimal filtering*. Prentice-Hall,Inc. Englewood cliffs, New Jersey, 1979. 3

P. Bjerksund. Contingent claims evalutation when the convenience yeild is stochastic: analytical results, working paper, norwegian school of economics and business administration. 1995. 10

R. J. Elliott and C. B. Hyndman. Parameter estimation in commodity markets-a filtering approach. *Journal of Economic Dynamics and Control*, 31(7):2350–2373, 2007. 1, 3, 8, 9, 21, 34, 61

R. J. Elliott and V. Krishnamurthy. New finite-dimensional filters for parameter estimation of discrete-time linear gaussian models. *IEEE Transactions on Automatic Control*, 44(5): 938–951, 1999. 1, 3, 34, 61, 62, 66

C. Fontana and W. J. Runggaldier. Credit risk and incomplete information:filtering and em parameter estimation. *Int. J. Theor. Appl. Fin.*, 13(5):638–715, 2010. 1

A. C. Harvey. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge, U.K., 1989. 3

C. Hyndman. Gaussian factor models - futures and forward prices. *IMA Journal Journal of Management Mathematics*, 18(4):353–369, 2007. 11

F. Jamishidian and M. Fein. Closed-form solutions for oil futures and european options in the gibson-schwartz model: A note, working paper, merrill lynch captial markets. 1990. 10

I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus, second ed.* Springer, New York, 1991. 8

E. S. Schwartz. The stochastic behaviour of commodity prices: Implications for valuation and hedging. *Journal of Finance*, LII(3):923–973, 1997. 1, 2, 9, 11, 28, 34

C. Wu. On convergence properties of the em algorithm. *Annals of Statistics*, 11:95–103, 1983. 6, 30