On Nonlinear Filtering Problem:
Structure Theorem and A New Suboptimal Filter

BY
Yang Jiao
B.S. (University of Science and Technology of China) 2006
M.S. (University of Illinois at Chicago) 2008
THESIS
Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Chicago, 2012

Chicago, Illinois

Defense Committee:
Stephen Yau, Chair and Advisor
David Nicholls
Jie Yang
Jan Verschelde
Lixing Jia, Chicago State University
Copyright by

Yang Jiao

2012
To my parents:

Xuchang Jiao and Ling Wang
ACKNOWLEDGMENTS

I am heartily thankful to my advisor, Stephen Yau, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. Professor Yau has supported me Research Assistant Scholarship for two years. He helped me in studying and living during my whole study career in UIC. His sharp mind and enthusiasm in mathematics are qualities for me to work toward.

I appreciate Professor Charles Knessl, Professor Jie Yang, Professor David Nicholls, Professor Jan Verschelde, and Professor Jing Wang for their great assistance. Their teaching has had an important influence on graduate studies. I appreciate Professor Lixing Jia for being my committee member. I also want to give appreciation to my friends and colleagues Dr. Mo Deng, Mr. Hexi ye, Mr. Zheng Fang, and so on.

Moreover, I would like to show my gratitude to my parents, Xuchang Jiao and Ling Wang for their tremendous supports, both emotionally and financially, which made this trip to U.S. possible. I also would like to thank my girlfriend Jing Wang for her support during past years.

Lastly, I offer my regards and blessings to all of those who supported me in any respect during the completion of this thesis.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
</tr>
<tr>
<td>2</td>
<td>MITTER CONJECTURE AND STRUCTURE THEOREM FOR SIX-DIMENSIONAL ESTIMATION ALGEBRAS</td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>2.2</td>
<td>Filtering Model And Basic Theorems</td>
</tr>
<tr>
<td>2.3</td>
<td>Six-dimensional Mitter Conjecture and Estimation Algebras</td>
</tr>
<tr>
<td>2.4</td>
<td>Structure theorem of Six-dimensional estimation algebras</td>
</tr>
<tr>
<td>2.5</td>
<td>Examples</td>
</tr>
<tr>
<td>2.6</td>
<td>Conclusions</td>
</tr>
<tr>
<td>3</td>
<td>NEW SUBOPTIMAL FILTER FOR POLYNOMIAL FILTERING PROBLEMS</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>3.2</td>
<td>Filtering Model And Basic Theorems</td>
</tr>
<tr>
<td>3.3</td>
<td>Some widely used suboptimal filters and previous results</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Unscented Kalman filter</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Ensemble Kalman filter</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Particle filter</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Some previous results for nonlinear filtering problems</td>
</tr>
<tr>
<td>3.4</td>
<td>New suboptimal filter for polynomial filtering problems</td>
</tr>
<tr>
<td>3.5</td>
<td>Examples, numerical results and discussions</td>
</tr>
<tr>
<td>3.6</td>
<td>CONCLUSIONS</td>
</tr>
<tr>
<td>CITED LITERATURE</td>
<td>89</td>
</tr>
<tr>
<td>VITA</td>
<td>94</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5.1</td>
<td>NSF compared with EKF for cubic sensor problem, with $R = 1$ and initial condition $x_0 \sim N(2, 0.01)$.</td>
</tr>
<tr>
<td>3.5.2</td>
<td>NSF compared with UKF for cubic sensor problem, with $R = 1$ and initial condition $x_0 \sim N(2, 0.01)$.</td>
</tr>
<tr>
<td>3.5.3</td>
<td>NSF compared with EnKF for cubic sensor problem, with $R = 1$ and initial condition $x_0 \sim N(2, 0.01)$, 20 ensemble for EnKF.</td>
</tr>
<tr>
<td>3.5.4</td>
<td>NSF compared with PF for cubic sensor problem, with $R = 1$ and initial condition $x_0 \sim N(2, 0.01)$, 50 particles for PF.</td>
</tr>
<tr>
<td>3.5.5</td>
<td>NSF compared with EKF for cubic sensor problem, with $R = 0.1$ and initial condition $x_0 \sim N(0, 0.001)$.</td>
</tr>
<tr>
<td>3.5.6</td>
<td>NSF compared with UKF for cubic sensor problem, with $R = 0.1$ and initial condition $x_0 \sim N(0, 0.001)$.</td>
</tr>
<tr>
<td>3.5.7</td>
<td>NSF compared with EnKF for cubic sensor problem, with $R = 0.1$ and initial condition $x_0 \sim N(0, 0.001)$, 20 ensembles for EnKF.</td>
</tr>
<tr>
<td>3.5.8</td>
<td>NSF compared with PF for cubic sensor problem, with $R = 0.1$ and initial condition $x_0 \sim N(0, 0.001)$, 50 particles for PF.</td>
</tr>
<tr>
<td>3.5.9</td>
<td>NSF compared with EKF for example 3.</td>
</tr>
<tr>
<td>3.5.10</td>
<td>NSF compared with UKF for example 3.</td>
</tr>
<tr>
<td>3.5.11</td>
<td>NSF compared with EnKF for example 3, 20 ensembles for EnKF.</td>
</tr>
<tr>
<td>3.5.12</td>
<td>NSF compared with PF for example 3, 50 particles for PF.</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>KF</td>
<td>Kalman filter</td>
</tr>
<tr>
<td>EKF</td>
<td>Extended Kalman filter</td>
</tr>
<tr>
<td>UKF</td>
<td>Unscented Kalman filter</td>
</tr>
<tr>
<td>EnKF</td>
<td>Ensemble Kalman filter</td>
</tr>
<tr>
<td>PF</td>
<td>Particle filter</td>
</tr>
<tr>
<td>NSF</td>
<td>New suboptimal filter</td>
</tr>
<tr>
<td>DMZ equation</td>
<td>Duncan-Morten-Zakai equation</td>
</tr>
<tr>
<td>SIR</td>
<td>Sampling Importance Resampling</td>
</tr>
</tbody>
</table>
SUMMARY

The nonlinear filtering problem involves the estimation of a stochastic process (called the signal or state process) that cannot be observed directly. Information containing the state is obtained from observations of a related process, i.e., the observation process. The main goal of nonlinear filtering is to determine the conditional expectations, or perhaps even computation of the entire conditional density of the state given the observation history. For an excellent introduction to nonlinear filtering theory, we refer the readers to the article by Marcus [1] and the book by Jazwinski [2].

Kalman [3], and Kalman-Bucy [4] published two historically important mathematical papers in ASME’s Journal of Basic Engineering in 1960 and 1961, respectively, which are highly influential in modern industry. Since then, the resulting Kalman-Bucy filter has been used in many areas such as navigational and guidance systems, radar tracking, solar mapping, satellite and airplane orbit determination, and forecast in weather, econometrics, and finance, see [5]. Despite the usefulness of Kalman-Bucy filter, it is not perfect. One of its weaknesses is that it is restricted to linear dynamical systems. Another weakness is the Gaussian assumption of initial value.

In this thesis, we introduce two methods to solve the nonlinear filtering problem. In chapter 2, we extend Yau and his coauthors’ work of Mitter conjecture for low dimensional algebras in nonlinear filtering problem. We prove the Mitter conjecture when the dimension $n = 6$. Using this result, we give the structure theorem of six-dimensional estimation algebra. We
shall show the structure of six-dimensional estimation algebra is not unique while when $n \leq 5$, the structure of estimation algebra is unique.

It is hard to solve nonlinear filtering problem when we want to find the optimal filter. In Chapter 3, we first introduce several widely used suboptimal filters, Extended Kalman filter, Unscented Kalman filter, Ensemble Kalman filter, Particle filter, and splitting up method. Then, we introduce a new suboptimal filter for polynomial filtering problems. With our special assumption, we can construct a closed form for conditional mean and conditional moments for the state process. Numerical results show that our new suboptimal filter works perfectly.
CHAPTER 1

INTRODUCTION

In early 1960s, Kalman [3], Kalman and Bucy [4] published two historically important papers on linear filtering that are highly influential in modern industry. Since then, the resulting Kalman-Bucy filter has been proved useful in science and engineering, for example in navigational and guidance systems, radar tracking, sonar ranging, satellite and airplane orbit determination, and forecasting in weather, econometrics and finance. However, the Kalman-Bucy filter has limited applicability because of the linearity assumptions of drift term and observation terms as well as the Gaussian assumption of the initial value.

The successes of Kalman filtering for the linear Gaussian estimation problems stimulated many researchers to attempt to generalize the Kalman’s results to systems with nonlinear dynamics. However, this nonlinear filtering problem is an essentially more difficult problem since the resulting optimal filter is, in general, infinite-dimensional, i.e., the conditional density depends on all the moments. In the mid-sixties, Kushner [6], [7] and Stratonovich [8] independently obtained the so-called Kushner-Stratonovich equation, which the conditional density should satisfy. Since the Kushner-Stratonovich equation is a nonlinear stochastic partial integral-differential equation, it is hard to solve even by using numerical methods.

Since then, solving the Kushner-Stratonovich equation has been one of the central problems of nonlinear filtering. One way of solving it is to reduce Kushner-Stratonovich equation to a linear Duncan-Mortensen-Zakai (DMZ) equation [9], [10], [11], which the unnormalized density
should satisfy. There are many ways to solve DMZ equation, for example, using Lie algebraic methods to construct finite dimensional recursive filter \cite{1}, \cite{12}, \cite{13}, \cite{14}, \cite{15}, \cite{16}, using the direct method \cite{17}, \cite{18}, \cite{19}, using the spitting up method \cite{20}, \cite{21}, \cite{22}, or using the approximate method \cite{23}, \cite{24}.

In the late 1970s, a basic approach to nonlinear filtering theory was independently proposed by Brockett and Clark \cite{12}, Brockett \cite{13}, and Mitter \cite{14}. They suggest that the construction of the filter can be divided into two parts: (i), a universal filter which is the evolution equation describing the unnormalized conditional density, the Duncan-Morten-Zakai equation, and (ii) a state-output map, which depends on the statistics being computed, where the state of the filter is the unnormalized conditional density. The reason for focusing on the Duncan-Morten-Zakai equation is that it is a linear equation and is a much simpler object than the nonlinear conditional density equation and can be treated using geometric idea. In 1983, Brockett formally proposed the problem of classifying all finite-dimensional estimation algebras in his lecture at the International Congress of Mathematicians. Recent work on estimation algebras has given us a deeper understanding of the Duncan-Morten-Zakai equation which was essential for progress in nonlinear filtering as well as in stochastic control. The advantage of the Brockett-Mitter approach of using estimation algebra method to solve the Duncan-Mortensen-Zakai equation is the following. As long as the the estimation algebra is finite dimensional, we will get a finite dimensional recursive filter and there is not a need to make any assumption on the initial distribution. Moreover, the approach applies well to nonlinear dynamical systems. Wong \cite{25} introduced a fundamental notion of Wong matrix which plays an important role in the
classification of finite-dimensional estimation algebras. He, see [26], gave a structure theorem of estimation algebra in case the drift term \( f(x) \) is real analytic and its first, second and third derivatives of \( f(x) \) are bounded functions. Under these assumptions, Wong also proved that all finite-dimensional estimation algebras are solvable and observation \( h(x) \) is polynomial of degree one. During 1991-1997, Yau et al., in a series of papers (see [15, 17, 27–32]), introduced new concepts and classified new finite-dimensional estimation algebras. The self-contained proof of the classification of finite-dimensional estimation algebras with maximal rank can be found in, see [33].

Despite the success of the classification of finite dimensional estimation algebras with maximal rank, the problem of classification of non-maximal rank finite dimensional estimation algebras is still wide open except for the case of state space dimension 2 which was finished by Wu and Yau, see [34] and some construction of non-maximal rank finite-dimensional algebras by Rasoulian and Yau, see [32]. Due to the difficulty of the problem, Brockett suggested that one should understand the low-dimensional estimation algebras first. In [35], Yau and Rasoulian have classified estimation algebras of dimension at most four. In [36], Choiu, Chiueh and Yau have classified estimation algebras of dimension five. We will prove the Mitter conjecture in Chapter 2, and give the structure theorem of six-dimensional estimation algebra.

Another approach to solve the nonlinear filtering problem is by using the approximate method to construct a suboptimal filter. The basic approximate approach is to assume the filtering densities belong to a certain class of functions such as Gaussian [2], exponential families [37], or some combinations. The most widely used approximate filter for the nonlinear filtering
problem is EKF, which is basically the Kalman-Bucy filter applied to a linearized system. In fact, EKF, which is the simplest Gaussian assumed-density filter, performs poorly when the dynamic system is significantly nonlinear. For other approximate methods, Unscented Kalman filter (UKF), Ensemble Kalman filter (EnKF), Particle filter (PF), and splitting up method, see [38], [39], [40], and [41]. All of these methods have their own weakness. Unscented Kalman filter and Ensemble Kalman filter assume that the probability density of the state vector is Gaussian. One of the drawbacks of Particle filter can be inefficient and is sensitive to outliers. Resampling step is applied at every iteration, which results in a rapid loss of diversity in particles. Furthermore, particle filters are more applied at low-dimensional systems, see [42] for the obstacles to high dimensional cases. Splitting up methods requires $g$ and $h$ to be bounded, which is not realistic in real life problem. In Chapter 3, we first present some widely used suboptimal filters. Then we introduce a new suboptimal filter for polynomial filtering problems. With our special assumption, we can write the equations of conditional mean and conditional moments in a closed form. We will also have some numerical results, which show that our new method works perfectly.
CHAPTER 2

MITTER CONJECTURE AND STRUCTURE THEOREM FOR
SIX-DIMENSIONAL ESTIMATION ALGEBRAS

2.1 Introduction

The success of classification of finite dimensional estimation algebras of maximal rank was partly inspired by the Mitter Conjecture which was communicated to the third named author in 1990 by Mitter. The conjecture states that all the functions in any finite dimensional estimation algebra are necessarily degree one polynomials. An affirmative answer to this conjecture will give us a deeper understanding of the structure of finite dimensional estimation algebras without maximal rank. Rasoulian and Yau [32] gave a general method to construct finite dimensional estimation algebras without maximal rank. But all their finite dimensional estimation algebras can be viewed as estimation algebras with maximal rank for certain filtering models. Recently, Wu and Yau [34] were able to classify all finite dimensional estimation algebras with state space dimension two. Subsequently the Mitter Conjecture was proved in the case of state space dimension 2. Chiueh and Yau prove the Mitter Conjecture for estimation algebra of dimension at most 5 with arbitrary state space dimension. We will deeper their results to prove the Mitter Conjecture for estimation algebra of dimension 6 with arbitrary state space dimension.
There are two reasons to tackle the problem of classification of non-maximal rank finite-dimensional estimation algebras. The first reason is that the problem was attempted by many well-known engineers without great success in the eighties of the last century. The second reason is that the solution of this problem may give us many new classes of finite-dimensional estimation algebras and hence many new classes of finite-dimensional filters. Due to the difficulty of the problem, Brockett suggested that one should understand the low-dimensional estimation algebras first. Rasoulian and Yau [35] and Chiou, Chiueh, and Yau [36] have classified estimation algebras of dimension at most five. We give a structure theorem for estimation algebras of dimension six.

2.2 Filtering Model And Basic Theorems

The filtering problem considered here is based on following signal observation model:

\[
\begin{align*}
\dot{x}(t) &= f(x(t))dt + g(x(t))dv(t) \quad x(0) = x_0 \\
\dot{y}(t) &= h(x(t))dt + dw(t) \quad y(0) = 0
\end{align*}
\]

(2.2.1)

in which \(x, v, y, \) and \(w\) are, respectively, \(\mathbb{R}^n, \mathbb{R}^p, \mathbb{R}^m, \) and \(\mathbb{R}^m,\) valued processes, and \(v\) and \(w\) have components that are independent, standard Brownian processes. We further assume \(n = p, f, h\) are \(C^\infty\) smooth, and \(g\) is an orthogonal matrix.

Let \(\rho(t, x)\) denote the conditional density of the state given the observation \(\{y(s) : 0 \leq s \leq t\}\). It is well known that \(\rho(t, x)\) is given by normalizing a function \(\sigma(t, x)\) which satisfies the Duncan-Mortensen-Zakai equation:
\[
\begin{aligned}
\left\{ \begin{array}{l}
d\sigma(t, x) = L_0 \sigma(t, x) + \sum_{i=1}^{m} L_i \sigma(t, x) dy_i(t), \\
\sigma(0, x) = \sigma_0,
\end{array} \right.
\end{aligned}
\]  

(2.2.2)

where

\[ L_0 = \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} - \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i} - \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} - \frac{1}{2} \sum_{i=1}^{m} h_i^2, \]

and for \( i = 1, \ldots, m \), \( L_i \) is the zero degree differential operator of multiplication by \( h_i \). Here \( \sigma_0 \) is the probability density of the initial point \( x_0 \). Let

\[
D_i = \frac{\partial}{\partial x_i} - f_i,
\]

\[
\eta = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 + \sum_{i=1}^{m} h_i^2. \quad (2.2.3)
\]

Then \( L_0 \) can be written as \( L_0 = \frac{1}{2} (\sum_{i=1}^{n} D_i^2 - \eta) \).

**Definition 2.2.1.** The estimation algebra \( E \) of a filtering model (2.2.1) is defined to be the Lie algebra generated by \( L_0, L_1, \ldots, L_m \) or \( E = \langle L_0, L_1, \ldots, L_m \rangle_{L_A} \). If \( x_i \in E \) for every \( 1 \leq i \leq n \), then \( E \) is called an estimation algebra of maximal rank.

**Definition 2.2.2.** Wong matrix \( \Omega = (\omega_{ij}) \) is an \( n \times n \) skew symmetric matrix with \( \omega_{ij} = (\partial f_j/\partial x_i) - (\partial f_i/\partial x_j) \).

**Theorem 2.2.3** ([43]). Any function in a finite-dimensional estimation algebra is polynomial of degree at most two.
The following theorem in [15] is useful in the classification of finite-dimensional estimation algebras.

**Theorem 2.2.4** ([15]). Let \( F(x_1, \ldots, x_n) \) be a polynomial on \( \mathbb{R}^n \). Suppose there exists a polynomial path \( c: \mathbb{R} \to \mathbb{R}^n \) such that \( \lim_{t \to \infty} c(t) = \infty \) and \( \lim_{t \to \infty} F(c(t)) = -\infty \). Then there are no \( C^\infty \) function \( f_1, \ldots, f_n \) on \( \mathbb{R}^n \) such that

\[
\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 = F.
\]

Let \( U_k \) be the vector space of differential operators of order up to and including \( k \). Assume that the coefficients of these differential operators are \( C^\infty \) functions. [25] proved the following theorem.

**Theorem 2.2.5** ([25]). If \( Y = \sum_{i=1}^{n} \gamma_i D_i \mod U_0 \) is an element in a finite-dimensional estimation algebra, then \( \gamma_i \) are polynomials of \( x_1, \ldots, x_n \) for all \( i \).

**Theorem 2.2.6** ([25]). If \( Y = \sum_{1 \leq i \leq j \leq n} \gamma_{ij} D_i D_j \mod U_1 \) is an element in a finite-dimensional estimation algebra, then \( \gamma_{ij} \) are polynomials of \( x_1, \ldots, x_n \) for all \( i, j \).

The following structure theorem (see, [35]) for Euler operator \( E_l = x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} + \cdots + x_l \frac{\partial}{\partial x_l} \) will be used frequently.

**Theorem 2.2.7** ([35]). Suppose \( m \in \mathbb{Z} \) is a constant integer and \( \zeta \) is a \( C^\infty \) function on \( \mathbb{R}^n \) and \( l \leq n \) such that \( E_l(\zeta) + m\zeta \) is a polynomial of degree \( k \), \( k \) a positive integer, in
\[ x_1, x_2, \ldots, x_l \text{ variables with coefficient in } \mathbb{C}^\infty \text{ functions of } x_{l+1}, \ldots, x_n \text{ variables.} \] If \( k + m \geq 0 \), \( \zeta \) is a polynomial of degree \( k \) in \( x_1, x_2, \ldots, x_l \) variables with coefficients in \( \mathbb{C}^\infty \) functions of \( x_{l+1}, \ldots, x_n \). If \( k + m < 0 \), \( \zeta \) is a polynomial of degree at most \( m' = -m \) in \( x_1, \ldots, x_l \) variables with coefficient in \( \mathbb{C}^\infty \) functions of \( x_{l+1}, \ldots, x_n \).

**Theorem 2.2.8** ([44]). Suppose that \( h = \frac{1}{2} \sum_{i=1}^{l} x_i^2 \) for some \( 1 \leq l \leq n \) is in the finite dimensional estimation algebra \( E \). Let \( \alpha_i = \sum_{j=1}^{l} x_j \omega_{ij} \). Then

(i) for \( 1 \leq i, j \leq l \), \( E_l(\omega_{ij}) + 2(\omega_{ij}) = \frac{\partial \alpha_i}{\partial x_j} - \frac{\partial \alpha_j}{\partial x_i} \) and \( \omega_{ij} \) is a polynomial in \( x_1, x_2, \ldots, x_n \).

(ii) for \( i \geq l+1, j \leq l \), \( E_l(\omega_{ij}) + \omega_{ij} = \frac{\partial \alpha_i}{\partial x_j} - \frac{\partial \alpha_j}{\partial x_i} \) and \( \omega_{ij} \) is a polynomial in \( x_1, x_2, \ldots, x_n \).

(iii) for \( l+1 \leq i, j \leq n \), \( E_l(\omega_{ij}) = \frac{\partial \alpha_i}{\partial x_j} - \frac{\partial \alpha_j}{\partial x_i} \) and \( \omega_{ij} \) is a polynomial in \( x_1, x_2, \ldots, x_n \) plus a \( \mathbb{C}^\infty \) function in \( x_{l+1}, \ldots, x_n \).

### 2.3 Six-dimensional Mitter Conjecture and Estimation Algebras

In [35], Yau and Rasoulian have classified all finite dimensional estimation algebras of dimension at most 4. As a corollary of their classification result, they have proved the Mitter Conjecture for estimation algebras of dimension at most 4. Chou, Chiueh, and Yau ([36]), have classified all finite dimensional estimation algebras of dimension five, based on the result (see [36]) about Mitter Conjecture estimation algebras of dimension 5. Therefore, we need to prove the Mitter Conjecture to classify all finite dimensional estimation algebras of dimension six.

The following lemma plays an important role in proving the Mitter conjecture for low dimensional estimation algebras.
Lemma 2.3.1 ([36]). For any $1 \leq l \leq n$, if $\gamma_i$, $i = 1, \ldots, l$, are polynomials in $x_1, \ldots, x_l$ with coefficients in $C^\infty$ functions of $x_{l+1}, \ldots, x_n$ satisfying

$$\frac{\partial \gamma_j}{\partial x_i} + \frac{\partial \gamma_i}{\partial x_j} = 0 \text{ for all } 1 \leq i, j \leq l,$$

then each $\gamma_i$ is necessary of the form

$$\gamma_i = \sum_{1 \leq j \leq l} c^j_i(x_{l+1}, \ldots, x_n)x_j + d_i(x_{l+1}, \ldots, x_n)$$

where $c^j_i(x_{l+1}, \ldots, x_n)$ and $d_i(x_{l+1}, \ldots, x_n)$ are $C^\infty$ functions and $c^j_i = -c^i_j$.

Observe

$$[L_0, x_1] = D_1, [D_1, x_1] = 1, [L_0, \sum_{j=2}^{n} a_j x_j + c] = \sum_{j=2}^{n} a_j D_j. \quad (2.3.1)$$

Proposition 2.3.2. If $\dim E = 6$, then

(i) $E$ cannot contain a degree one polynomial and a degree two polynomial.

(ii) $E$ cannot contain two linearly independent degree two polynomials.

Proof. Let

$$h = \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{j=1}^{n} b_j x_j + c_1,$$

$$\tilde{h} = \sum_{i,j=1}^{n} d_{ij} x_i x_j + \sum_{j=1}^{n} e_j x_j + c_2$$
be two linearly independent degree two polynomials. Observe

\[
Y_1 = [L_0, h] = \sum_{i,j=1}^{n} a_{ij} (x_i D_j + x_j D_i) + \sum_{j=1}^{n} b_j D_j \mod U_0 \tag{2.3.2}
\]

\[
\tilde{Y}_1 = [L_0, \tilde{h}] = \sum_{i,j=1}^{n} d_{ij} (x_i D_j + x_j D_i) + \sum_{j=1}^{n} e_j D_j \mod U_0 \tag{2.3.3}
\]

\[
Y_2 = [L_0, Y_1] = \sum_{i,j=1}^{n} a_{ij} D_i D_j \mod U_1 \tag{2.3.4}
\]

\[
\tilde{Y}_2 = [L_0, \tilde{Y}_1] = \sum_{i,j=1}^{n} d_{ij} D_i D_j \mod U_1 \tag{2.3.5}
\]

where \(U_i\) is the space of differential operators with order at most \(i\).

(i) If \(E\) contains a degree one polynomial and degree two polynomial, we may assume that \(x_1\), and degree two polynomial \(h = \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{j=1}^{n} b_j x_j + c_1\), where \(a_{ij} \neq 0\) for some \(i, j\), are in \(E\). From (2.3.1) and (2.3.2), we see that \(E\) contains six linearly independent elements \(L_0, x_1, D_1, 1, h\) and \(Y_1\). Notice that

\[
[D_1, h] = \sum_{i=1}^{n} (a_{1i} + a_{i1}) x_i + b_1
\]

need to be a linear combination of \(x_1\), and 1, which implied that \(a_{1i} + a_{i1} = 0, i \neq 1\). Now again, we consider \([L_0, D_1]\), which must be linear combination of \(x_1, 1, h\), and \(Y_1\):

\[
[L_0, D_1] = \sum_{j=2}^{n} \omega_{1j} D_j + \frac{1}{2} \sum_{j=2}^{n} \frac{\partial \omega_{1j}}{\partial x_j} + \frac{1}{2} \frac{\partial \eta}{\partial x_1}
\]

\[
= \lambda_1 x_1 + \lambda_2 + \lambda_3 h + \lambda_4 Y_1 \tag{2.3.6}
\]
If $a_{ij} + a_{ji} \neq 0$ for some $i \neq j$, then from (2.3.4), we see that $E$ contains seven linearly independent elements $L_0, x_1, D_1, h, Y_1$ and $Y_2$.

If $a_{ij} + a_{ji} = 0$ for all $i \neq j$, then $h = \sum_{i=1}^{n} (2a_{ii}x_i^2 + \sum_{j=1}^{n} b_j x_j + c_1)$, we have $\lambda_4(2a_{11}x_1 + b_1) = 0$ (D_1 term) from (2.3.6). If $\lambda_4 \neq 0$, then we know $a_{11} = 0$. Again, $E$ contains seven linearly independent elements $L_0, x_1, D_1, h, Y_1$ and $Y_2$. If $\lambda_4 = 0$, then $\omega_{1i} = 0$ for $1 \leq i \leq n$ from (2.3.6). Then we have

$$\frac{\partial \eta}{\partial x_1} = \lambda_1 x_1 + \lambda_2 + \lambda_3 h$$

$$\Rightarrow \eta = \lambda_1 x_1^2 + \frac{2}{3} \lambda_3 x_1^3 + 2x_1 f_1(x_2, \ldots, x_n),$$

where $f_1$ is a function of $x_2, \ldots, x_n$ only. Thus, $\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_1} + \sum_{i=1}^{n} f_i^2 = \eta - h^2 - x_1^2$. Then by Theorem 2.2.4, there is no $C^\infty$ solution $f = (f_1, \ldots, f_n)$ for equation (2.2.3), a contradiction.

(ii) Suppose $E$ contains two linearly independent degree two polynomial, $h$ and $\tilde{h}$. Denote $h^{(i)}$ be the degree $i$ monomial part of $h$. If $h^{(2)} = c \tilde{h}^{(2)}$ for some constant $c$, then $E$ contains a degree one polynomial $h^{(1)} + c_1$ and a degree two polynomial $h^{(2)} + c_2$. In view of the proof of (i), we know it is impossible. One the other hand, if $h^{(2)} \neq c \tilde{h}^{(2)}$ for any constant $c$, then we may assume that $\sum a_{ij} x_i x_j \neq \sum x_i^2$. From (2.3.2)–(2.3.5), we see that $E$ contains the seven linearly independent elements $L_0, h, \tilde{h}, Y_1, \tilde{Y}_1, Y_2, \tilde{Y}_2$.

The following result is provided by [36] with dim $E = 5$. However, without any change of the proof, we will find the result holds for dim $E = 6$ too.
Theorem 2.3.3. Let $E$ be the estimation algebra of a filtering model (2.2.1) with state space dimension $n$. Suppose $\dim E = 6$. Then

(i) $n$ is at least two,

(ii) If $E$ contains a degree two polynomial, then $E$ contains a polynomial of the form $h_1 = \sum_{i=1}^{l} x_i^2$ for some $1 \leq l \leq n$.

In view of the above discussion, in order to prove the Mitter conjecture of estimation algebras of dimension six, we only need to consider two possible cases: case 1, $h_1(x) = \frac{1}{2} \sum_{i=1}^{l} x_i^2$, $1 \leq l < n$ and case 2, $h_1(x) = \frac{1}{2} \sum_{i=1}^{n} x_i^2$. In both cases, $h_j = c_j h_1$ for some constants $c_j$, $j = 2, \ldots, m$. We may assume that $n \geq 2$. Notice that $\sum_{j=1}^{m} h_j^2 = c^2 h_1^2$ where $c^2 = 1 + c_2^2 + \cdots + c_m^2$.

Theorem 2.3.4. Let $E$ be a finite dimensional estimation algebra associate to filtering model (2.2.1) with arbitrary state space dimension. Then any function in $E$ is a polynomial of degree at most one if $\dim E = 6$.

Proof. Denote $E_i = x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} + \cdots + x_i \frac{\partial}{\partial x_i}$.

Case 1: $h_1(x) = \frac{1}{2} \sum_{i=1}^{l} x_i^2$, $l \leq n - 1$. Observe that

$$Y_1 : = [L_0, h_1] = \sum_{i=1}^{l} x_i D_i + \frac{l}{2}$$

$$Y_2 : = [L_0, Y_1] = \sum_{i=1}^{l} D_i^2 - \sum_{i=1}^{n} \alpha_i D_i - \frac{1}{2} \beta_n + \frac{1}{2} E_l(\eta)$$

where $\alpha_i = \sum_{j=1}^{l} x_j \omega_{ij}$, and $\beta_n = \sum_{i=1}^{n} \frac{\partial \alpha_i}{\partial \eta} = \sum_{i=1}^{n} \sum_{j=1}^{l} x_j \frac{\partial \omega_{ij}}{\partial \eta}$. 


\begin{equation}
Z_{21} := [Y_2, Y_1] = 2 \sum_{i=1}^{l} D_i^2 + \sum_{i=1}^{n} E_l(\alpha_i) D_i - 3 \sum_{i=1}^{l} \alpha_i D_i \\
- \beta_i + \sum_{i=1}^{n} \alpha_i^2 + \frac{1}{2} E_l(\beta_n) - \frac{1}{2} E_l(\xi),
\end{equation}

where \( \beta_i = \sum_{i=1}^{n} \frac{\partial \alpha_i}{\partial x_i} = \sum_{i=1}^{l} \sum_{j=1}^{l} x_j \frac{\partial \omega_{ij}}{\partial x_i} \) and \( \sum_{i=1}^{n} \alpha_i^2 = \sum_{i=1}^{l} \sum_{j,k=1}^{l} x_j x_k \omega_{ij} \omega_{ik} \).

\begin{equation}
\Delta := Z_{21} - 2Y_2 = \sum_{i=1}^{l} (E_l(\alpha_i) - \alpha_i) D_i + \sum_{i=l+1}^{n} (E_l(\alpha_i) - 2\alpha_i) D_i + \sum_{i=1}^{n} \alpha_i^2 - \beta_i + \frac{1}{2} (E_l(\beta_n) + 2\beta_n) - \frac{1}{2} E_l(\xi). 
\end{equation}

\begin{equation}
Y_3 := [L_0, Y_2] = 2 \sum_{j=1}^{n} \sum_{i=1}^{l} \omega_{ji} D_j D_i - \sum_{j=1}^{n} \sum_{i=1}^{l} \sum_{j=1}^{l} \frac{\partial \alpha_i}{\partial x_i} D_j D_i \quad \mod U_1 
\end{equation}

\begin{equation}
= - \sum_{1 \leq i < j \leq l} \left( \frac{\partial \alpha_i}{\partial x_i} + \frac{\partial \alpha_i}{\partial x_j} \right) D_i D_j - \sum_{l+1 \leq i \leq j \leq n} \left( \frac{\partial \alpha_i}{\partial x_i} + \frac{\partial \alpha_i}{\partial x_j} \right) D_i D_j \\
+ \sum_{i=l+1}^{n} \sum_{j=1}^{l} \left( \frac{\partial \omega_{ji}}{\partial x_i} + \frac{\partial \omega_{ji}}{\partial x_j} \right) D_i D_j \quad \mod U_1
\end{equation}

Based on the proof of the main theorem from [45], \( L_0, h_1, Y_1, Y_2, \Delta \) are linearly independent.

We have the following multiplication table:
Since \( \dim E = 6 \), then there is one and only one operator, which is linearly independent of \( L_0, h_1, Y_1, Y_2, \Delta \), among the upper triangle operators in above table. Notice that \([h_1, Y_1], [h_1, \Delta] \in U_0 \) and \([h_1, Y_2] - Y_1 \in U_0 \), so \([h_1, Y_1], [h_1, \Delta] \) and \([h_1, Y_2] - Y_1 \) must be a constant multiple of \( h_1 \). Thus the operator, which is linearly independent of \( L_0, h_1, Y_1, Y_2, \Delta \), must be among \([L_0, Y_2], [L_0, \Delta], [Y_1, \Delta], [Y_2, \Delta] \). Let us assume \( Y_3 = [L_0, Y_2] \) is linearly independent of \( L_0, h_1, Y_1, Y_2, \Delta \). Then we have

\[
Y_3 := [L_0, Y_2] = 2 \sum_{i=1}^{n} \sum_{j=1}^{l} \omega_{ji} D_j D_i - \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial \alpha_i}{\partial x_j} D_i D_j \mod U_1
\]

\[
= -\sum_{1 \leq i < j \leq l} \left( \frac{\partial \alpha_j}{\partial x_i} + \frac{\partial \alpha_i}{\partial x_j} \right) D_i D_j - \sum_{l+1 \leq i < j \leq n} \left( \frac{\partial \alpha_j}{\partial x_i} + \frac{\partial \alpha_i}{\partial x_j} \right) D_i D_j
\]

\[
+ \sum_{i=l+1}^{n} \sum_{j=1}^{l} \left[ 2 \omega_{ji} - \left( \frac{\partial \alpha_j}{\partial x_i} + \frac{\partial \alpha_i}{\partial x_j} \right) \right] D_i D_j \mod U_1
\]

Consider

\[
Y_4 := [L_0, Y_3]
\] (2.3.7)
\[\begin{align*}
&= \sum_{k=1}^{n} \sum_{1 \leq i,j \leq l} \frac{\partial}{\partial x_k} \left( \frac{\partial \alpha_j}{\partial x_i} + \frac{\partial \alpha_i}{\partial x_j} \right) D_k D_i D_j - \sum_{k=1}^{n} \sum_{l+1 \leq i,j \leq n} \frac{\partial}{\partial x_k} \left( \frac{\partial \alpha_j}{\partial x_i} + \frac{\partial \alpha_i}{\partial x_j} \right) D_k D_i D_j \\
&\quad + \sum_{k=1}^{n} \sum_{l+1 \leq i,j \leq n} \sum_{1 \leq l' \leq i,j \leq n} \frac{\partial}{\partial x_k} \left[ 2\omega_{ji} - \left( \frac{\partial \alpha_j}{\partial x_i} + \frac{\partial \alpha_i}{\partial x_j} \right) \right] D_k D_i D_j \mod U_2
\end{align*}\]

Suppose there exist \(\lambda_1, \ldots, \lambda_6\) in \(\mathbb{R}\) such that

\[Y_4 = \lambda_1 L_0 + \lambda_2 h_1 + \lambda_3 Y_1 + \lambda_4 Y_2 + \lambda_5 \Delta + \lambda_6 Y_3 \quad (2.3.8)\]

Then we get:

\[\alpha_k = \sum_{s=1}^{l} d_s^k x_s x_k + c_k x_k + e_k,\]

where \(d_s^k, c_k\) are functions independent of \(x_1, \ldots, x_l, d_s^k = 0,\) and \(e_k\) is independent of \(x_k, \frac{\partial^2 e_k}{\partial x_i^2} = -2d_i^k\) for \(1 \leq i, k \leq l,\) \(2.3.9\)

\[\alpha_k = \sum_{s=l+1}^{n} d_s^k x_s x_k + c_k x_k + e_k,\]

where \(d_s^k, c_k\) are functions independent of \(x_{l+1}, \ldots, x_n, d_s^k = 0,\) and \(e_k\) is independent of \(x_k, \frac{\partial^2 e_k}{\partial x_i^2} = -2d_i^k\) for \(1 + 1 \leq i, k \leq n,\) \(2.3.10\)

\[\frac{\partial}{\partial x_k} \left[ 2\omega_{ji} - \left( \frac{\partial \alpha_j}{\partial x_i} + \frac{\partial \alpha_i}{\partial x_j} \right) \right] = 0 \text{ for } 1 \leq j \leq l, l + 1 \leq i \leq n, 1 \leq k \leq n \quad (2.3.11)\]

The analysis is similar to case 2 in [45]. Notice \(\alpha_k = \sum_{s=1}^{l} d_s^k x_s x_k + c_k x_k + e_k = \sum_{i=1}^{l} x_i \omega_{ki},\) for \(1 \leq k \leq l,\) from (2.3.9), we have \(\sum_{s=1}^{l} d_s^k x_s x_k + c_k x_k = 0\) (simply put \(x_k = 0\)), which implies that \(d_s^k = c_k = 0, 1 \leq s, k \leq l.\) Similarly, \(d_s^k = c_k = 0, l + 1 \leq s, k \leq n\) from (2.3.10).
Then \(E_l(\alpha_i) - \alpha_i = E_l(e_i) - e_i\), and \(E_l(e_i) - e_i\) is independent of \(x_i\) for \(1 \leq i \leq l\). We get

\[
\Delta = \sum_{i=1}^{l} (E_l(e_i) - e_i)D_i + \sum_{i=l+1}^{n} (E_l(\alpha_i) + 2\alpha_i)D_i + \sum_{i=1}^{n} \alpha_i^2 \\
- \beta_l + \frac{1}{2}(E_l(\beta_n) + 2\beta_n) - \frac{1}{2}E_l(E_l(\eta) + 2\eta))
\]

\[\[\Delta, h_1\] = \sum_{i=1}^{l} (E_l(e_i) - e_i)x_i\]

Since \(\dim E = 6\), \(\sum_{i=1}^{l} (E_l(e_i) - e_i)x_i\) must be a constant multiple of \(\sum_{i=1}^{l} x_i^2\). This implies \(E_l(e_i) - e_i = 0\) for \(i = 1, \ldots, l\) (put \(x_j = 0\), \(j \neq i\)). So for any fixed \(l \geq 1\)

\[
\Delta = \sum_{i=l+1}^{n} (E_l(\alpha_i) + 2\alpha_i)D_i + \sum_{i=1}^{n} \alpha_i^2 - \beta_l + \frac{1}{2}(E_l(\beta_n) + 2\beta_n) - \frac{1}{2}E_l(E_l(\eta) + 2\eta)) \tag{2.3.12}
\]

From (2.3.11), we have

\[\[2\omega_{ji} - \left(\frac{\partial \alpha_j}{\partial x_i} + \frac{\partial \alpha_i}{\partial x_j}\right)\] = constant \(\lambda_{ij}\), for \(1 \leq j \leq l, l + 1 \leq i \leq n\) \tag{2.3.13}

From (2.3.12)

\[\[Y_2, \Delta\] = 2 \sum_{j=1}^{l} \sum_{i=l+1}^{n} \frac{\partial}{\partial x_j} (E_l(\alpha_i) + 2\alpha_i)D_jD_i \mod U_1
\]

\[= \tilde{\lambda}_1 L_0 + \tilde{\lambda}_2 h_1 + \tilde{\lambda}_3 Y_1 + \tilde{\lambda}_4 Y_2 + \tilde{\lambda}_5 \Delta + \tilde{\lambda}_6 Y'_3, \quad \text{for some } \tilde{\lambda}_1, \ldots, \tilde{\lambda}_6 \in \mathbb{R} \tag{2.3.14}\]
It follows that (by (2.3.12))

\[
\frac{\partial}{\partial x_j}(E_l(\alpha_i) + 2\alpha_i) = \text{constant, for } 1 \leq j \leq l, l+1 \leq i \leq n \quad (2.3.15)
\]

Observe that \( E_l(\alpha_i) + 2\alpha_i, l+1 \leq i \leq n \) are polynomials of \( x_1, \ldots, x_n \) by Theorem 2.2.5. (2.3.15) implies that \( E_l(\alpha_i) + 2\alpha_i, l+1 \leq i \leq n \) are polynomial of degree one in \( x_1, x_2, \ldots, x_l \) with coefficients in \( \mathbb{C}^\infty \) function of \( x_{l+1}, \ldots, x_n \). In view of Theorem 2.2.7, \( \alpha_i, l+1 \leq i \leq n \) are polynomial of degree one in \( x_1, x_2, \ldots, x_l \) with coefficients in \( \mathbb{C}^\infty \) function of \( x_{l+1}, \ldots, x_n \),

\[
\alpha_i = \sum_{j=1}^l c^j_i (x_l + 1, \ldots, x_n) x_j + d_i (x_{l+1}, \ldots, x_n), \text{ for } l+1 \leq i \leq n
\]

It follows that \( d_i(x_{l+1}, \ldots, x_n) = 0, c^j_i (x_l + 1, \ldots, x_n) = \omega_{ij} \) because \( \alpha_i = \sum_{j=1}^l x_j \omega_{ij} \) where \( \omega_{ij} \) are \( \mathbb{C}^\infty \) smooth.

Then from (2.3.13), we have

\[
2\omega_{ji} - (\omega_{ij} + \sum_{k=1}^l x_k \left( \frac{\partial \omega_{jk}}{\partial x_i} + \frac{\partial \omega_{ik}}{\partial x_j} \right)) = \lambda_{ij}, 1 \leq j \leq l, l+1 \leq i \leq n
\]

\[
\Rightarrow 3\omega_{ji} - \sum_{k=1}^l x_k \frac{\partial \omega_{jk}}{\partial x_i} = \lambda_{ij} (\omega_{ik} \text{ is independent of } x_1, \ldots, x_l)
\]

\[
\Rightarrow 3\omega_{ji} = \lambda_{ij}, \sum_{k=1}^l x_k \frac{\partial \omega_{jk}}{\partial x_i} = 0 \text{ (by setting } x_1 = \ldots = x_l = 0) \]

Then \( \alpha_i = \frac{1}{n} \sum_{j=1}^l \lambda_{ji} x_j \). We have \( \alpha_i = \tilde{\lambda}_{ij} \) from (2.3.14), which implies that \( \tilde{\lambda}_{ij} = \lambda_{ji} = 0 \), further, \( \alpha_i = 0, l+1 \leq i \leq n \).
From $E_l(e_i) - e_i = 0$, $1 \leq i \leq l$, by Theorem 2.2.7, we know $e_i$ is a polynomial of degree at most 1 in $x_1, \ldots, x_l$ with coefficients in $C^\infty$ functions of $x_{l+1}, \ldots, x_n$. Thus, $\alpha_i$ is a polynomial of degree at most 1 in $x_1, \ldots, x_l$ with coefficients in $C^\infty$ functions of $x_{l+1}, \ldots, x_n$. Now $\Delta$ is of form

$$\Delta = \sum_{i=1}^{l} \alpha_i^2 - \beta_l + \frac{1}{2}(E_l(\beta_n) + 2\beta_n) - \frac{1}{2}E_l(E_l(\eta) + 2\eta)$$

$$= e \sum_{i=1}^{l} x_i^2,$$

for some constant $e$.

The last equality comes from the fact that $\dim E = 6$. Notice that $\beta_l = \sum_{i=1}^{l} \frac{\partial \alpha_i}{\partial x_i}$, and $\beta_n = \sum_{i=1}^{n} \frac{\partial \alpha_i}{\partial x_i}$, $\beta_l$ and $\beta_n$ are polynomials of degree at most one by (2.3.9), (2.3.10). By Theorem 2.2.7, $\eta$ is a polynomial of degree at most two in $x_1, \ldots, x_l$. We get a contradiction since there is no $C^\infty$ smooth solution for $\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 = \sum_{j=1}^{m} h_j^2 + \eta = -c^2(\sum_{i=1}^{l} x_i^2)^2 + \eta$ in view of Theorem 2.2.4. We have shown that six dimensional estimation algebra cannot contain an element of the form $\frac{1}{7} \sum_{i=1}^{l} x_i^2$ with $1 \leq l < n$ if $L_0, h_1, Y_1, Y_2, \Delta, Y_3$ are linearly independent.

If $Y_3' := [Y_1, \Delta]$ is linearly independent of $L_0, h_1, Y_1, Y_2, \Delta$, then we can still use a similar proof since we are focusing on the coefficients of $D_i D_j D_k$ terms and $[Y_1, \Delta], [Y_2, \Delta] \in U_2$. We only need to change equation (2.3.8) to $Y_4 = \lambda_1 L_0 + \lambda_2 h_1 + \lambda_3 Y_1 + \lambda_4 Y_2 + \lambda_5 \Delta + \lambda_6 Y_3'$ and equation (2.3.14) to

$$[Y_2, \Delta] = 2 \sum_{j=1}^{l} \sum_{i=l+1}^{n} \frac{\partial}{\partial x_j}(E_l(\alpha_i) + 2\alpha_i)D_j D_i \mod U_1$$

$$= \tilde{\lambda}_1 L_0 + \tilde{\lambda}_2 h_1 + \tilde{\lambda}_3 Y_1 + \tilde{\lambda}_4 Y_2 + \tilde{\lambda}_5 \Delta + \tilde{\lambda}_6 Y_3',$$

for some $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_6 \in \mathbb{R}$.
If $Y'_3 := [Y_2, \Delta]$ or $Y'_3 := [L_0, \Delta]$ is linearly independent of $L_0, h_1, Y_1, Y_2, \Delta$, we do not even need to change equation (2.3.8). We only need to change equation (2.3.14) to

\[
[L_0, Y_2] = -\sum_{1 \leq i \leq j \leq l} \left( \frac{\partial \alpha_i}{\partial x_i} + \frac{\partial \alpha_j}{\partial x_j} \right) D_i D_j \sum_{l+1 \leq i \leq j \leq n} \left( \frac{\partial \alpha_i}{\partial x_i} + \frac{\partial \alpha_j}{\partial x_j} \right) D_i D_j \\
+ \sum_{i=l+1}^{n} \sum_{j=1}^{l} \left[ 2 \omega_{ji} - (\frac{\partial \alpha_i}{\partial x_i} + \frac{\partial \alpha_j}{\partial x_j}) \right] D_i D_j \pmod{U_1} \\
= \tilde{\lambda}_1 L_0 + \tilde{\lambda}_2 h_1 + \tilde{\lambda}_3 Y_1 + \tilde{\lambda}_4 Y_2 + \tilde{\lambda}_5 \Delta + \tilde{\lambda}_6 Y'_3, \quad \text{for some } \tilde{\lambda}_1, \ldots, \tilde{\lambda}_6 \in \mathbb{R}
\]

**Case 2:** $h_1(x) = \frac{1}{2} \sum_{i=1}^{n} x_i^2$. In this case, $\omega_{ij}, 1 \leq i, j \leq n$ are polynomial in $x_1, \ldots, x_n$ by Theorem 2.2.8. Observe that

\[
Y_1 := [L_0, h_1] = \sum_{i=1}^{n} x_i D_i + \frac{n}{2} Y_1, h_1] = \sum_{i=1}^{n} x_i^2 = 2 h_1,
\]

\[
Y_2 := 2 L_0 - [L_0, Y_1] = \sum_{i=1}^{n} \alpha_i D_i + g_1,
\]

\[
\Delta := [Y_1, Y_2] = \sum_{i=1}^{n} [E_n(\alpha_i) - \alpha_i] D_i + \sum_{i=1}^{n} \alpha_i^2 + E_n\left[ \frac{1}{2} \beta_n - \frac{1}{2} E_n(\eta) - \eta \right],
\]

\[
Y_3 := [L_0, Y_2] = \sum_{1 \leq i \leq j \leq n} \left( \frac{\partial \alpha_j}{\partial x_i} + \frac{\partial \alpha_i}{\partial x_j} \right) D_i D_j \pmod{U_1},
\]

\[
[L_0, Y_3] = \sum_{i,j,k=1}^{n} \frac{\partial^2 \alpha_j}{\partial x_k \partial x_i} D_k D_i D_j + \sum_{i,j,k=1}^{n} \frac{\partial \alpha_j}{\partial x_i} (\omega_{jk} D_i D_k + \omega_{ik} D_k D_j) \pmod{U_1},
\]

where $\alpha_i = \sum_{j=1}^{n} x_j \omega_{ij}$, $\beta_n = \sum_{j=1}^{n} \frac{\partial \alpha_i}{\partial x_j}$ and $g_1 = \frac{1}{2} \beta_n - \frac{1}{2} [E_n(\eta) + 2 \eta]$. Based on the analysis from [45], $L_0, h_1, Y_1, Y_2, Y_3$ are linearly independent. If $L_0, h_1, Y_1, Y_2, Y_3, \Delta$ are linearly dependent, then $[L_0, Y_3]$ is a linear combination of $L_0, h_1, Y_1, Y_2, Y_3, \Delta$. Using the analysis from [45], case
2 of their main theorem, we will obtain a contradiction. So $\Delta$ is a linear combination of
$L_0, h_1, Y_1, Y_2, Y_3$: $\Delta = \lambda_1 L_0 + \lambda_2 h_1 + \lambda_3 Y_1 + \lambda_4 Y_2 + \lambda_5 Y_3$. If $\lambda_5 \neq 0$, then we could use the
same argument in [45] to discuss the coefficient of $D_i D_j$ terms to obtain a contradiction. So
$\Delta = \lambda_1 L_0 + \lambda_2 h_1 + \lambda_3 Y_1 + \lambda_4 Y_2$, and

$$\lambda_1 = 0$$

$$E_n(\alpha_i) = \lambda_3 x_i + (\lambda_4 + 1) \alpha_i, \quad 1 \leq i \leq n$$  \tag{2.3.16}

$$E_n(g_1) - \lambda_4 g_1 = -\sum_{i=1}^{n} \alpha_i^2 + \frac{\lambda_2}{2} \sum_{i=1}^{n} x_i^2 + \frac{n\lambda_3}{2}$$  \tag{2.3.17}

from (2.3.16), we have $\sum_{j=1}^{n} x_j (E_n(\omega_{1j}) - \lambda_4 \omega_{1j}) = \lambda_3 x_1$, which implies $\lambda_3 = 0$. Let $p_i = \deg(\alpha_i)$. [45] shows that $(\alpha_1, \ldots, \alpha_n) \neq C(x_1, \ldots, x_n)$ for any constant $C$ and $p_i \neq 0$. For any polynomial
$\psi$, we shall denote the homogeneous part of degree $s$ of $\psi$ by $\psi^{(s)}$. Equation (2.3.16) implies

$$\sum_{s=1}^{p_i} s \alpha_i^{(s)} = E_n(\alpha_i) = (\lambda_4 + 1) \sum_{s=1}^{p_i} \alpha_i^{(s)}.$$  (2.3.18)

Therefore $\alpha_i = \alpha_i^{(p_i)}$, where $p_i = \lambda_4 + 1$ if $p_i > 0$. Without loss of generality we shall assume
deg $\alpha_1 = \deg \alpha_2 = \cdots = \deg \alpha_\mu = \lambda_4 + 1 := p$, deg $\alpha_\mu+1 = \cdots = \deg \alpha_n = 0$ for some
$0 < \mu \leq n$.

We now claim that

$$\alpha_i = \sum a_{j_1 \ldots j_p}^i x_{j_1} \ldots x_{j_p}, \quad 1 \leq i \leq \mu$$  \tag{2.3.18}
where the summation is over distinct indices $j_1 \neq i$, $j_2 \neq i$, ..., $j_p \neq i$ and $a_{j_1 \ldots j_p}^i$ are real constants.

For $L_0, h_1, Y_1, Y_2, Y_3$, we have the following multiplication table:

<table>
<thead>
<tr>
<th></th>
<th>$L_0$</th>
<th>$h_1$</th>
<th>$Y_1$</th>
<th>$Y_2$</th>
<th>$Y_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_0$</td>
<td>0</td>
<td>$Y_1 = [L_0, h_1]$</td>
<td>$2L_0 - Y_2 = [L_0, Y_1]$</td>
<td>$Y_3 = [L_0, Y_2]$</td>
<td>$[L_0, Y_3]$</td>
</tr>
<tr>
<td>$h_1$</td>
<td>$-[L_0, h_1]$</td>
<td>0</td>
<td>$[h_1, Y_1]$</td>
<td>$[h_1, Y_2]$</td>
<td>$[h_1, Y_3]$</td>
</tr>
<tr>
<td>$Y_1$</td>
<td>$-[L_0, Y_1]$</td>
<td>$-[h_1, Y_1]$</td>
<td>0</td>
<td>$[Y_1, Y_2]$</td>
<td>$[Y_1, Y_3]$</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>$-[L_0, Y_2]$</td>
<td>$-[h_1, Y_2]$</td>
<td>$-[Y_1, Y_2]$</td>
<td>0</td>
<td>$[Y_2, Y_3]$</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>$-[L_0, Y_3]$</td>
<td>$-[h_1, Y_3]$</td>
<td>$-[Y_1, Y_3]$</td>
<td>$-[Y_2, Y_3]$</td>
<td>0</td>
</tr>
</tbody>
</table>

Since dim $E = 6$, then there is one and only one operator, which is linearly independent of $L_0, h_1, Y_1, Y_2, Y_3$, among the upper triangle operators in above table. Notice $[h_1, Y_1], [h_1, Y_2] \in U_0$, so $[h_1, Y_1], [h_1, Y_2]$ must be a constant multiple of $h_1$. We already know $\Delta = [Y_1, Y_2]$ is a linear combination of $L_0, h_1, Y_1, Y_2, Y_3$. Thus the operator, which is linearly independent of $L_0, h_1, Y_1, Y_2, Y_3$, among $[L_0, Y_3] \in U_3, [Y_1, Y_3], [Y_2, Y_3] \in U_2, [h_1, Y_3] \in U_1$. Let us assume $Y_4 := [L_0, Y_3]$ (if $Y_4 := [Y_1, Y_3], [Y_2, Y_3], \text{ or } [h_1, Y_3]$, the proof is similar and easier since they do not contain $D_i D_j D_k$ terms).

We next consider

$$Y_5 := [L_0, Y_4] = \sum_{i,j,k,l=1}^n \frac{\partial^3 \alpha_{lj}}{\partial x_k \partial x_l \partial x_i} D_l D_k D_i D_j +$$
\[ \sum_{i,j,k,l=1}^{n} \left( \frac{\partial^2 \alpha_l}{\partial x_i \partial x_k} \omega_{j} + \frac{\partial^2 \alpha_j}{\partial x_l \partial x_k} \omega_{l} \right) D_k D_i D_j \mod U_2 \]

since \( Y_5 \) has to be a linear combination of \( L_0, h_1, Y_1, Y_2, Y_3, Y_4 \), we have

\[ Y_5 = \sum_{i,j,k,l=1}^{n} \left( \sum_{l=1}^{n} \left( \frac{\partial^2 \alpha_l}{\partial x_i \partial x_k} \omega_{j} + \frac{\partial^2 \alpha_j}{\partial x_l \partial x_k} \omega_{l} \right) \right) D_k D_i D_j \mod U_2 \]  

(2.3.19)

\[ \frac{\partial^3 \alpha_i}{\partial x_i^3} = 0, \ 1 \leq i \leq n \ (D_i^4 \text{ terms}) \]  

(2.3.20)

\[ \frac{\partial^3 \alpha_i}{\partial x_j^3} + 3 \frac{\partial^3 \alpha_j}{\partial x_i \partial x_j^2} = 0, \ 1 \leq i, j \leq n \ (D_i^3 D_j \text{ terms}) \]  

(2.3.21)

\[ \frac{\partial^3 \alpha_i}{\partial x_j \partial x_k^2} + \frac{\partial^3 \alpha_j}{\partial x_i \partial x_k^2} + 2 \frac{\partial^3 \alpha_k}{\partial x_i \partial x_j \partial x_k} = 0, \ 1 \leq i, j, k \leq n \ (D_i^2 D_j D_k \text{ terms}) \]  

(2.3.22)

\[ \frac{\partial^3 \alpha_i}{\partial x_j \partial x_k \partial x_l} + \frac{\partial^3 \alpha_j}{\partial x_i \partial x_l \partial x_k} + \frac{\partial^3 \alpha_k}{\partial x_l \partial x_i \partial x_j} + \frac{\partial^3 \alpha_l}{\partial x_i \partial x_j \partial x_k} = 0, \ 1 \leq i, j, k, l \leq n \ (D_i D_j D_k D_l \text{ terms}) \]  

(2.3.23)

Differentiating (2.3.23) with respect to \( x_k \), together with (2.3.22), we get

\[ \frac{\partial^4 \alpha_k}{\partial x_i \partial x_j \partial x_k \partial x_l} = 0, \]

which implies

\[ \frac{\partial \alpha_k}{\partial x_k} = \sum_{s,r=1}^{n} d_{s}^{r} x_{s} x_{r} + \sum_{s=1}^{n} d_{s}^{s} x_{s} + c_{k}, \]  

(2.3.24)
where \( d_{s,r}^k, d_s^k, c_k \) are constants, \( d_{i,i}^j + d_{j,j}^i = 0 \), \( d_{i,j}^j + d_{j,i}^i = d_{j,i}^i + d_{i,j}^j \), and \( d_{k,k}^k = 0 \) from (2.3.20)–(2.3.23).

It follows that
\[
\alpha_k = \sum_{s,r=1}^{n} d_{s,r}^k x_s x_r x_k + \sum_{s=1}^{n} d_s^k x_s x_k + c_k x_k + e_k,
\]
where \( e_k \) is independent of \( x_k \) variable. (2.3.20) and (2.3.23) imply
\[
\frac{\partial^3 e_i}{\partial x_j^3} = -3(d_{i,j}^j + d_{j,i}^j). \tag{2.3.25}
\]

(a) For the case \( p > 3 \), in view of (2.3.20), (2.3.23), we have
\[
c_k = 0, \quad \text{for } 1 \leq k \leq \mu,
\]
\[
d_i^k, d_j^k = 0, \quad \text{for } 1 \leq i,j,k \leq n,
\]
\[
\frac{\partial^3 e_i}{\partial x_j^3} = 0, \quad \text{for } 1 \leq i,j \leq n.
\]

From (2.3.22) we have \( \frac{\partial}{\partial x_j} \left( \frac{\partial^2 \alpha_i}{\partial x_k^2} \right) + \frac{\partial}{\partial x_i} \left( \frac{\partial^2 \alpha_j}{\partial x_k^2} \right) = 0 \). From Lemma 2.3.1, we have
\[
\frac{\partial^2 e_j}{\partial x_k^2} = \sum_{s=1}^{n} c_s^j x_s + d_j, \quad \text{for all } 1 \leq j,k \leq n
\]
where \( c_s^j, d_j \) are constants and \( c_s^j = c_j^s \). Notice \( \deg(\alpha_j) > 3 \), we have \( c_s^j = 0, d_j = 0 \). Therefore \( \alpha_j \) does not contain \( x_k^k \) terms, \( k > 1 \). We have established (2.3.18) for \( p > 3 \).
(b) For the case \( p = 3 \), from (2.3.24) and (2.3.25) and the fact that \( e_j \) is independent of \( x_j \), we have

\[
\alpha_j = \sum_{r,s=1}^{n} d^j_{rs} x_r x_s x_j - \frac{1}{2} \sum_{s=1}^{n} (d^s_{ij} + d^s_{js}) x_s^2 + \sum_{r,s=1}^{n} b^j_{rs} x_r x_s + \sum_{r,s,k=1}^{n} c^j_{rsk} x_r x_s x_k
\]  

(2.3.26)

for some constants \( d^j_{rs}, b^j_{rs}, c^j_{rsk} \) with \( d^j_{jj} = 0, b^j_{jr} = b^j_{rj} = b^j_{rr} = 0, c^j_{rsk} = c^j_{rjk} = c^j_{rsj} = c^j_{rrr} = c^j_{rss} = c^j_{ttr} = 0, \forall \ 1 \leq r, s, k \leq n, 1 \leq j \leq \mu \), \( d^s_{rs} = d^s_{js} = b^s_{rs} = c^s_{rsk} = 0, \forall \ 1 \leq r, s, k \leq n, \mu + 1 \leq j \leq n \). In view of \( \alpha_j = \sum_{s=1}^{n} x_s (\sum_{r=1}^{n} d^j_{rs} x_r x_j - \frac{1}{2}(d^s_{ij} + d^s_{js}) x_s^2 + \sum_{r=1}^{n} b^j_{rs} x_r x_s + \sum_{r,k=1}^{n} c^j_{rsk} x_r x_s x_k) \) = \( x_s \omega_j \). Recall that \( \omega_{ij}, 1 \leq i, j \leq n \) are polynomials. Then \( \omega_{js}^{(2)} \), homogeneous polynomial of degree 2 part of \( \omega_{js} \) must be of form

\[
\omega_{js}^{(2)} = \sum_{r=1}^{n} d^j_{rs} x_r x_j - \frac{1}{2} (d^s_{ij} + d^s_{js}) x_s^2 + \sum_{r=1}^{n} b^j_{rs} x_r x_s + \sum_{r,k=1}^{n} c^j_{rsk} x_r x_s x_k, 1 \leq j, s \leq n,
\]  

(2.3.27)

for some constant \( c^j_{rsk} \) with \( c^j_{rk} = c^j_{kj} = c^j_{sk} = c^j_{rsk} = 0 \). Next we shall show that \( d^j_{js} + d^j_{sj} = d^j_{rr} = b^j_{rs} = 0 \). Compute the coefficient \( \sum_{i=1}^{n} \left( \frac{\partial^2 \alpha_i}{\partial x_j \partial x_i} \right) \omega_{li} \) of the term \( D^3_i \), \( 1 \leq i \leq n \) in equation (2.3.19) which is a polynomial. Denote this polynomial by \( u_i \), we have

\[
u_i = \sum_{l=1}^{n} \sum_{l 
eq i}^{n} [2d^j_{il} x_l - 3(d^s_{il} + d^s_{li}) x_i + \sum_{r=1}^{n} 2b^j_{ir} x_r + \sum_{r=1}^{n} (d^i_{il} + d^i_{li}) x_r] \cdot \]

\[
\sum_{r=1}^{n} d^j_{rs} x_r x_j - \frac{1}{2} (d^s_{ij} + d^s_{js}) x_s^2 + \sum_{r=1}^{n} b^j_{rs} x_r x_s + \sum_{r,k=1}^{n} c^j_{rsk} x_r x_s x_k + (\omega_{li} - \omega_{li}^{(2)}) \]  

(2.3.28)
Since $Y_5$ is a linear combination of $L_0, h_1, Y_1, Y_2, Y_3, Y_4$, and notice that the coefficient $\frac{\partial^2 \alpha_i}{\partial x_i}$ of $D_3$ term in $Y_4$ is $2 \sum_{s=1}^{n} d^i_{ls} x_s + \text{constant}$, which is a degree 1 polynomial, we have that the coefficients $\frac{1}{2}(d^i_{l} + d^i_{l_i})^2$ of $x_3^2$, $(d^i_{l})^2$ of $x_i x_i^2$ and $2(b^l_{r_i})^2$ of $x_i x_i^2$ are necessary zero. It follows that $d^i_{l} + d^i_{l_i} = d^i_{l_i} = b^l_{r_i} = 0$ for $i \neq l$. Therefore $\alpha_j = \sum_{r,s,k=1}^{n} c^j_{rsk} x_r x_s x_k$. Therefore, $\alpha_j$, $1 \leq j \leq \mu$ are of the form (2.3.18) with $p = 3$.

(c) For the case $p = 2$, from (2.3.24),(2.3.25) and fact that $e_j$ is independent of $x_j$, we have

$$\alpha_j = \sum_{s=1}^{n} d^j_s x_s x_j - \sum_{s=1}^{n} b^j_s x^2_s + \sum_{1 \leq s, r \leq n} c^j_{rs} x_r x_s, \ 1 \leq j \leq n,$$

(2.3.29)

With the similarity to case $p = 3$, we will have $b^j_s = 0$. Therefore $\alpha_j$, $1 \leq j \leq \mu$ are of the form (2.3.18) with $p = 2$.

Now, we have shown equation (2.3.18), hence

$$\frac{\partial \alpha_k}{\partial x_k} = 0, \ 1 \leq k \leq n.$$  

(2.3.30)

(2.3.17) becomes

$$E_n(g_1) - (p - 1)g_1 = -\sum_{i=1}^{\mu} (\sum_{j_1 \neq i, \ldots, j_p \neq i} a^i_{j_1 \ldots j_p} x_{j_1} \cdots x_{j_p})^2$$

$$+ \lambda_2 \left( \frac{1}{2} \sum_{i=1}^{n} x_i^2 \right) \quad \text{if } \mu > 0$$

(2.3.31)
\[ E_n(g_1) = \lambda_2 \left( \frac{1}{2} \sum_{i=1}^{n} x_i^2 \right) \quad \text{if } \mu = 0 \]  

(2.3.32)

By Theorem 2.2.7, \( g_1 \) is a polynomial of degree 2p (if \( \mu > 0 \)) or 2 (if \( \mu = 0 \)). By expressing \( g_1 \) as a polynomial in general form and by comparing the coefficient in both sides of (2.3.31) and (2.3.32), we get that \( g_1 \) is a polynomial of degree 2 in \( x_r \) for some \( x_r \) variable. Therefore \( g_1 \) can be written in the form

\[ u_2(x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n)x_r^2 + u_1(x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n)x_r + u_0(x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n) \],

where \( u_i \) is a polynomial of \( x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n \). Recall \( g_1 = \beta_n - \frac{1}{2}(E_n(\eta) + 2\eta) \), where \( \beta_n = \sum_{i=1}^{n} \frac{\partial \alpha}{\partial x_i} = 0 \) by (2.3.30). By Theorem 2.2.7, \( \eta \) is a polynomial of degree 2p, \( p \geq P1 \) in \( x_1, \ldots, x_n \). Therefore, we can write \( \eta \) in the form

\[ \theta_2(x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n)x_r^2 + \theta_1(x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n)x_r + \theta_0(x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n), \]

where \( u_i \) is a polynomial of \( x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n \). It follows that

\[
\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 = \sum_{j=1}^{m} h_j^2 + \eta = -c^2 \left( \sum_{i=1}^{l} x_i^2 \right)^2 + \eta \\
= -c^2 x_r^4 + \text{lower degree in } x_r
\]

By Theorem 2.2.4, we get a contradiction by taking a polynomial path \( x_r = t, x_i = 0 \) for \( i \neq r \). Thus we have finished proof Mitter conjecture for \( \dim E = 6 \). \( \square \)

2.4 **Structure theorem of Six-dimensional estimation algebras**

In this section, we shall prove the structure of six-dimensional estimation algebra. Since \( \dim E = 6 \), we only need to consider the following two cases by theorem 2.3.4. Otherwise, if \( E \) contains more than 2 degree one polynomial, we will find \( \dim E > 6 \).
Case 1 $E$ does not contain 2 degree one polynomials. By proposition 2.3.2 and theorem 2.3.4, we have $m = 1$ (hence $\sum_{i=1}^{m} h_i^2 = h_1^2$) and we may assume that $h_1 = x_1$. Observe that

$$[L_0, x_1] = D_1, [D_1, x_1] = 1 \quad (2.4.1)$$

we have $\{1, x_1, D_x, L_0\} \subseteq E$. Consider

$$Y_1 = [L_0, D_1] = \sum_{i=1}^{n} \omega_{1i} D_i + \frac{1}{2} \sum_{i=1}^{n} \frac{\partial \omega_{1i}}{\partial x_i} + \frac{1}{2} \frac{\partial \eta}{\partial x_1} \quad (2.4.2)$$

From [36], we know that $\omega_{1i} \neq 0$ for some $2 \leq i \leq n$. And $1, x_1, D_1, L_0, Y_1$ are linearly independent.

$$Y_2 := [L_0, Y_1]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial \omega_{1i}}{\partial x_i} D_i D_j + \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{1j} \omega_{ji} D_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 \omega_{1j}}{\partial x_i^2} D_j + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial \omega_{1j}}{\partial x_i} + \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^3 \omega_{1j}}{\partial x_i \partial x_j^2} + \frac{1}{4} \sum_{i=1}^{n} \frac{\partial \eta}{\partial x_1} + \frac{1}{2} \sum_{i=1}^{n} \omega_{1i} \frac{\partial \eta}{\partial x_i} \quad (2.4.3)$$
If \( \frac{\partial \omega_{1j}}{\partial x_i} + \frac{\partial \omega_{1i}}{\partial x_j} = 0, \) \( \sum_{j=1}^{n} \omega_{1j} \omega_{ji} + \frac{1}{2} \frac{\partial^2 \eta}{\partial x_i \partial x_1} = 0, \) in view of (2.4.2), proposition 2.3.2 and theorem 2.3.4, \( E \) is a five-dimensional estimation algebra. Based on the analysis of [36], we know \( 1, x_1, D_1, L_0, Y_1, \) and \( Y_2 \) are linearly independent. Consider

\[
Y_3 := [L_0, Y_2] = \sum_{i,j,k=1}^{n} \frac{\partial^2 \omega_{1j}}{\partial x_i \partial x_k} D_k D_i D_j \mod U_2 \quad (2.4.4)
\]

Since \( Y_3 \) has to be a linear combination of \( 1, x_1, D_1, L_0, Y_1, \) and \( Y_2, \) we can write

\[
Y_3 = 2\lambda L_0 + C_0 x_1 + C_1 Y_1 + C_2 + C_3 D_1 + C_4 Y_2 \quad (2.4.5)
\]

From the above equation, we have

\[
\frac{\partial^2 \omega_{1i}}{\partial x_i^2} = 0, \ 1 \leq i \leq n \ (D_i^3 \text{ terms}) \quad (2.4.6)
\]

\[
\frac{\partial^2 \omega_{1j}}{\partial x_i^2} + 2 \frac{\partial^2 \omega_{1i}}{\partial x_i \partial x_j} = 0, \ 1 \leq i, j \leq n \ (D_i^2 D_j \text{ terms}) \quad (2.4.7)
\]

\[
\frac{\partial^2 \omega_{1k}}{\partial x_i \partial x_j} + \frac{\partial^2 \omega_{1j}}{\partial x_i \partial x_k} + \frac{\partial^2 \omega_{1i}}{\partial x_j \partial x_k} = 0, \ 1 \leq i, j, k \leq (D_i D_j D_k \text{ terms}) \quad (2.4.8)
\]

It follows

\[
\omega_{1i} = \sum_{s=1}^{n} d^i_s x_s x_i + c_i x_i + e_i, \ 1 \leq i \leq n \quad (2.4.9)
\]

\[
\frac{\partial^2 e_j}{\partial x_i^2} = -2d^j_i \quad (2.4.10)
\]
where \( d_i^j, c_i \) are constants and \( d_i^j = 0 \), \( c_i \) is independent of \( x_i \).

It follows that

\[
Y_2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial \omega_{ij}}{\partial x_i} D_i D_j + \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij} \omega_{ji} D_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i^2} D_j \\
+ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij} \frac{\partial \omega_{ij}}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i \partial x_j} D_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 \eta}{\partial x_i \partial x_j} D_i \\
+ \frac{1}{4} \sum_{i=1}^{n} \frac{\partial^2 \eta}{\partial x_i \partial x_i} + \frac{1}{2} \sum_{i=1}^{n} \omega_{ij} \frac{\partial \eta}{\partial x_j} \tag{2.4.11}
\]

\[
Y_3 = \sum_{i,j,k=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i \partial x_k} D_k D_i D_j + \sum_{i,j,k=1}^{n} \frac{\partial \omega_{ij}}{\partial x_i} \omega_{jk} D_k D_i + \sum_{i,j,k=1}^{n} \frac{\partial \omega_{ij}}{\partial x_i} \omega_{jk} D_k D_i + \frac{1}{2} \sum_{i,j,k=1}^{n} \omega_{ij} \omega_{jk} D_k D_i \\
+ \frac{1}{2} \sum_{i,j,k=1}^{n} \frac{\partial \omega_{ij}}{\partial x_i} \frac{\partial \omega_{jk}}{\partial x_k} D_k + \frac{1}{2} \sum_{i,j,k=1}^{n} \omega_{ij} \frac{\partial \omega_{jk}}{\partial x_k} D_k + \frac{1}{2} \sum_{i,j,k=1}^{n} \omega_{ij} \omega_{jk} \omega_{ki} D_k \\
+ \frac{1}{2} \sum_{i,j,k=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i^2} D_k + \frac{1}{2} \sum_{i,j,k=1}^{n} \omega_{ij} \frac{\partial \omega_{ki}}{\partial x_k} D_k - \frac{1}{2} \sum_{i,j,k=1}^{n} \omega_{ij} \omega_{kj} D_k \\
+ \frac{1}{4} \sum_{i,j,k=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i \partial x_k} \omega_{kj} D_k - \frac{1}{4} \sum_{i,j,k=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i \partial x_k} \omega_{ki} D_k - \frac{1}{4} \sum_{i,j,k=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i \partial x_j} \omega_{ki} D_k \\
+ \frac{1}{2} \sum_{i,j,k=1}^{n} \frac{\partial^2 \eta}{\partial x_i \partial x_k} D_k D_i - \frac{1}{2} \sum_{i,j,k=1}^{n} \frac{\partial^2 \eta}{\partial x_i \partial x_k} \omega_{ki} D_k + \frac{1}{4} \sum_{i,j,k=1}^{n} \frac{\partial^4 \eta}{\partial x_i \partial x_k \partial x_i \partial x_k} D_i \\
- \frac{1}{4} \sum_{i,j,k=1}^{n} \frac{\partial^2 \eta}{\partial x_i \partial x_k} \omega_{ki} + \frac{1}{2} \sum_{i,j,k=1}^{n} \frac{\partial \omega_{ij}}{\partial x_k} \omega_{ji} D_k + \frac{1}{4} \sum_{i,j,k=1}^{n} \frac{\partial^2 \eta}{\partial x_k^2} \omega_{ji} \frac{\partial \omega_{ji}}{\partial x_i} + \frac{1}{4} \sum_{i,j,k=1}^{n} \frac{\partial^2 \eta}{\partial x_k^2} \omega_{ji} \frac{\partial \omega_{ji}}{\partial x_i} \\
+ \frac{1}{4} \sum_{i,j,k=1}^{n} \frac{\partial^2 \eta}{\partial x_k \partial x_i} \frac{\partial \omega_{ij}}{\partial x_j} + \frac{1}{2} \sum_{i,j,k=1}^{n} \frac{\partial \omega_{ij}}{\partial x_k} \frac{\partial \omega_{ij}}{\partial x_i} + \frac{1}{2} \sum_{i,j,k=1}^{n} \frac{\partial \omega_{ij}}{\partial x_k} \frac{\partial \omega_{ij}}{\partial x_j} + \frac{1}{4} \sum_{i,j,k=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i \partial x_j} \frac{\partial \omega_{ij}}{\partial x_i} \\
+ \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial \omega_{ij}}{\partial x_i} \frac{\partial \omega_{ij}}{\partial x_j} + \frac{1}{2} \sum_{i,j=1}^{n} \omega_{ij} \frac{\partial \omega_{ji}}{\partial x_i} + \frac{1}{4} \sum_{i,j,k=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i^2} \frac{\partial \omega_{ij}}{\partial x_i} + \frac{1}{4} \sum_{i,j,k=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i \partial x_j} \frac{\partial \omega_{ij}}{\partial x_j}
\begin{equation}
\frac{1}{4} \sum_{i=1}^{n} \frac{\partial^2 \eta}{\partial x_i \partial x_1} \frac{\partial \eta}{\partial x_i} \tag{2.4.12}
\end{equation}

By comparing the coefficients of \( D_1^2 \) term, \( D_i^2 \), \( 2 \leq i \leq n \) terms, and \( D_i D_j \), \( 1 \leq i \neq j \leq n \) terms from (2.4.5) and (2.4.12), we have

\begin{equation}
- \frac{3}{2} \sum_{i=1}^{n} \frac{\partial \omega_{ij}^2}{\partial x_1} + \frac{1}{2} \frac{\partial^3 \eta}{\partial x_1^3} = \lambda \tag{2.4.13}
\end{equation}

\begin{equation}
\sum_{j=1}^{n} \frac{\partial \omega_{ij}^2}{\partial x_j} + \sum_{j=1}^{n} \frac{\partial \omega_{1j}^2}{\partial x_j} = \frac{1}{2} \frac{\partial \omega_{1j}^2}{\partial x_1} = \lambda + C_4 \frac{\partial \omega_{1j}}{\partial x_1}, \quad 2 \leq i \leq n \tag{2.4.14}
\end{equation}

\begin{equation}
\sum_{j=1}^{n} \frac{\partial \omega_{1j} \omega_{ij}}{\partial x_k} + \frac{\partial \omega_{ij} \omega_{ij}}{\partial x_k} + \frac{\partial \omega_{ii} \omega_{ij}}{\partial x_k} + \frac{\partial \omega_{ij} \omega_{jk}}{\partial x_k} = C_4 \left( \frac{\partial \omega_{1i}}{\partial x_k} + \frac{\partial \omega_{1k}}{\partial x_i} \right), \quad 1 \leq i < k \leq n \tag{2.4.15}
\end{equation}

By comparing the coefficients of \( D_1 \) term, \( D_k \), \( 2 \leq k \leq n \) terms from (2.4.5) and (2.4.12), we have

\begin{equation}
- \sum_{i,j=1}^{n} \left( \frac{\partial \omega_{ij}^2}{\partial x_i} \right)^2 + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial \omega_{ij}^2}{\partial x_1} \frac{\partial \omega_{ij}}{\partial x_i} - \sum_{i,j=1}^{n} \omega_{1j} \omega_{ij} \omega_{1i} - \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 (\omega_{ij}^2)}{\partial x_i \partial x_j} \\
- \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i^2} \omega_{ij} - \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i \partial x_j} \omega_{ij} - \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 \eta}{\partial x_i \partial x_1} \omega_{1i} + \frac{1}{4} \sum_{i=1}^{n} \frac{\partial^4 \eta}{\partial x_1^4} \\
+ \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial \omega_{ij}}{\partial x_1} \frac{\partial \omega_{ij}}{\partial x_i} + \frac{1}{4} \sum_{i=1}^{n} \frac{\partial^4 \eta}{\partial x_i^4} + \frac{1}{2} \sum_{i=1}^{n} \frac{\partial \omega_{1i}}{\partial x_i} \frac{\partial \eta}{\partial x_1} \\
+ \frac{1}{2} \sum_{i=1}^{n} \frac{\partial \omega_{1i}}{\partial x_1} \frac{\partial \eta}{\partial x_i} = C_3 + C_4 \left( \sum_{i=1}^{n} \omega_{1i}^2 \right)^2 + \frac{1}{2} \sum_{j=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i \partial x_j} + \frac{1}{2} \frac{\partial^2 \eta}{\partial x_1^2}. \tag{2.4.16}
\end{equation}
by comparing the zero order differential operators in both sides of (2.4.5), we get
\[
\sum_{i,j=1}^{n} \frac{\partial \omega_{ij}}{\partial x_i} \frac{\partial \omega_{j}}{\partial x_j} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial \omega_{ij}}{\partial x_k} \frac{\partial \omega_{ji}}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial \omega_{ij}}{\partial x_i} \frac{\partial \omega_{ji}}{\partial x_j} - \sum_{i,j=1}^{n} \omega_{ij} \omega_{ji} \omega_{ki}
\] 
\[+ \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i^2} - \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i^2} \omega_{ki} - \frac{1}{4} \sum_{i,j=1}^{n} \frac{\partial^2 \omega_{ij}}{\partial x_i \partial x_j} \omega_{ki} - \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 \eta}{\partial x_i \partial x_1} \omega_{ki}
\] 
\[+ \frac{1}{4} \sum_{i=1}^{n} \frac{\partial^4 \eta}{\partial x_k \partial x_i^2 \partial x_1} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial \omega_{ki}}{\partial x_i} \frac{\partial \eta}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial \omega_{ki}}{\partial x_i} \frac{\partial \eta}{\partial x_j} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial \omega_{ki}}{\partial x_j} \frac{\partial \eta}{\partial x_1} \frac{\partial \omega_{ij}}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial \omega_{ji}}{\partial x_1} \frac{\partial \omega_{ij}}{\partial x_i} \frac{\partial \omega_{ji}}{\partial x_j} \frac{\partial \eta}{\partial x_1} \frac{\partial \omega_{ij}}{\partial x_i} \frac{\partial \omega_{ji}}{\partial x_j}
\] 
\[= C_1 \omega_{1k} + C_4 \left( \sum_{i=1}^{n} \omega_{1i} \omega_{ik} + \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 \omega_{1k}}{\partial x_i^4} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \omega_{1j}}{\partial x_i \partial x_j} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 \eta}{\partial x_i \partial x_1} \right), \quad 2 \leq k \leq n. \quad (2.4.17)
\]
\[
\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij} \frac{\partial \omega_{ji}}{\partial x_i} + \frac{1}{4} \sum_{i=1}^{n} \frac{\partial^3 \eta}{\partial x_1 \partial x_i^2} + \frac{1}{2} \sum_{i=1}^{n} \omega_{ii} \frac{\partial \eta}{\partial x_i} + \sum_{1 \leq i < j \leq n} \frac{\partial \omega_{ij}}{\partial x_i} \omega_{ji}. \tag{2.4.18}
\]

**Case 2** \(E\) contains two degree one polynomials. Then \(E = \{L_0, x_1, D_1, 1, \sum_{j=2}^{n} a_j x_j, \sum_{j=2}^{n} a_j D_j\}\).

Consider \(Y_1 := [L_0, D_1]\) and \(\tilde{Y}_1 := [L_0, \sum_{j=2}^{n} a_j D_j]\), which must be linear combination of the six elements:

\[
Y_1 := [L_0, D_1] = \sum_{j=2}^{n} \omega_{1j} D_j + \frac{1}{2} \sum_{j=2}^{n} \frac{\partial \omega_{1j}}{\partial x_j} + \frac{1}{2} \frac{\partial \eta}{\partial x_1}
= \lambda L_0 + C_0 x_1 + C_1 D_1 + C_2 + C_3 \sum_{j=2}^{n} a_j x_j + C_4 \sum_{j=2}^{n} a_j D_j \tag{2.4.19}
\]

\[
\tilde{Y}_1 := [L_0, \sum_{j=2}^{n} a_j D_j] = \sum_{i=1}^{n} \sum_{j=2}^{n} \omega_{ji} D_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=2}^{n} \frac{\partial \omega_{ji}}{\partial x_i} + \frac{1}{2} \sum_{j=2}^{n} \frac{\partial \eta}{\partial x_j}
= \tilde{\lambda} L_0 + \tilde{C}_0 x_1 + \tilde{C}_1 D_1 + \tilde{C}_2 + \tilde{C}_3 \sum_{j=2}^{n} a_j x_j + \tilde{C}_4 \sum_{j=2}^{n} a_j D_j \tag{2.4.20}
\]

Then we have \(\lambda = \tilde{\lambda} = C_1 = 0, \tilde{C}_1 = \sum_{j=2}^{n} \omega_{1j}, \omega_{1j} = C_4 a_j, 2 \leq j \leq n, \sum_{j=2}^{n} \omega_{ji} = \tilde{C}_4 a_i, 2 \leq i \leq n.\)

Then we have \(\frac{1}{2} \sum_{j=2}^{n} \frac{\partial \omega_{ij}}{\partial x_j} = 0, \frac{1}{2} \sum_{i=1}^{n} \sum_{j=2}^{n} \frac{\partial \omega_{ji}}{\partial x_i} = 0.\) So \(\frac{1}{2} \frac{\partial \eta}{\partial x_1}\) is a linear combination of \(x_1, 1\) and \(\sum_{j=2}^{n} a_j x_j:\)

\[
\frac{1}{2} \frac{\partial \eta}{\partial x_1} = C_0 x_1 + C_2 \sum_{j=2}^{n} a_j x_j + C_3. \tag{2.4.21}
\]

And \(\frac{1}{2} \sum_{j=2}^{n} \frac{\partial \eta}{\partial x_j}\) is also a linear combination of \(x_1, 1\) and \(\sum_{j=2}^{n} a_j x_j:\)

\[
\frac{1}{2} \sum_{j=2}^{n} \frac{\partial \eta}{\partial x_j} = \tilde{C}_0 x_1 + \tilde{C}_2 \sum_{j=2}^{n} a_j x_j + \tilde{C}_3. \tag{2.4.22}
\]
Summarizing what we have shown above, we have the following structure theorem for six-dimensional estimation algebras.

**Theorem 2.4.1.** Suppose that the state space of the filtering model (2.2.1) is of dimension at least two. Then the six-dimensional estimation algebra is isomorphic to a Lie algebra generated by:

(i) $L_0$ and $x_1$, with a basis given by $1, x_1, D_1 = \frac{\partial}{\partial x_1} - f_1, Y_1 = [L_0, D_1], Y_2 = [L_0, Y_1]$ and $L_0$, and:

(i.1) $\omega_{1i} \neq 0$ for some $2 \leq i \leq n$ and each $\omega_{1i}$ is of the form

$$\omega_{1i} = \sum_{s=1}^{n} d^i_s x_s x_i + c_i x_i + e_i, \quad 1 \leq i \leq n$$

(2.4.23)

$$\frac{\partial^2 e_j}{\partial x^2_i} = -2d^i_j$$

(2.4.24)

where $d^i_s, c_i$ are constants and $d^i_i = 0$, $e_i$ are independent of $x_i$.

(ii) There exist constants $\lambda, C_0, C_1, C_2, C_3, C_4$ such that equations (2.4.13)-(2.4.18) are satisfied.

Or,

(ii) $L_0, x_1$ and $\sum_{j=2}^{n} a_j x_j$, with a basis given by $L_0, x_1, D_1, 1, \sum_{j=2}^{n} a_j x_j, \sum_{j=2}^{n} a_j D_j$, and:

(ii.1) $\omega_{1j} = C_4 a_j, \quad 2 \leq j \leq n, \quad \sum_{j=2}^{n} \omega_{ji} = \tilde{C_4} a_i, \quad 2 \leq i \leq n, \quad \frac{1}{2} \sum_{j=2}^{n} \frac{\partial \omega_{ji}}{\partial x_j} = 0$, where $C_4$ and $\tilde{C_4}$ are constants. And $\frac{1}{2} \sum_{i=1}^{n} \sum_{j=2}^{n} \frac{\partial \omega_{ji}}{\partial x_i} = 0$. 


(ii.2) There exist constants $C_0, C_2, C_3, \tilde{C}_0, \tilde{C}_2, \text{ and } \tilde{C}_3$ such that

\[
\frac{1}{2} \frac{\partial \eta}{\partial x_1} = C_0 x_1 + C_2 \sum_{j=2}^{n} a_j x_j + C_3. \tag{2.4.25}
\]

\[
\frac{1}{2} \sum_{j=2}^{n} \frac{\partial \eta}{\partial x_j} = \tilde{C}_0 x_1 + \tilde{C}_2 \sum_{j=2}^{n} a_j x_j + \tilde{C}_3. \tag{2.4.26}
\]

2.5 Examples

Example 1 Consider the filtering model

\[
\begin{cases}
    dx(t) = f(x(t))dt + g(x(t))dv(t) \\
    dy(t) = h(x(t))dt + dw(t)
\end{cases}
\]

where

\[
\begin{align*}
    f_1 &= a - \frac{3}{\sqrt{2}} x_1 \\
    f_2 &= b + x_1 + 2x_3 \\
    f_3 &= c - x_2 \\
    f_i &= g_i(x_4, \ldots, x_n), \quad 4 \leq i \leq n \\
    h(x) &= x_1
\end{align*}
\]

and $a, b, c$ are constants. Then

\[
\omega_{12} = 1, \omega_{13} = 0, \omega_{23} = -3, \omega_{1j} = 0, \quad 4 \leq i \leq n
\]
\[ \sum_{i=1}^{n} f_i^2 + \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} = (a - \frac{3}{\sqrt{2}} x_1)^2 + (b + x_1 + 2x_3)^2 + (c - x_2)^2 - \frac{3}{\sqrt{2}} \]

It is easy to check that \( \lambda = 0, C_0 = -\frac{33}{2}, C_1 = -\frac{7}{2}, C_2 = C_3 = C_4 = 0 \). The estimation algebra \( E \) is six-dimensional with basis \( \{1, x_1, D_1, Y_1 = D_2 - \sqrt{3} a + b + \frac{13}{2} x_1 + 2x_3, Y_2 = -D_3 - c + x_2, L_0\} \).

**Example 2** Consider the filtering model

\[
\begin{aligned}
    dx(t) &= f(x(t))dt + g(x(t))dv(t) \\
    dy(t) &= h(x(t))dt + dw(t)
\end{aligned}
\]

where

\[
\begin{aligned}
    f_1 &= x_1, \\
    f_2 &= ax_2 - dx_2^2 + bx_3 + 2dx_2x_3 - \frac{1}{2}(1 + 2d)x_3^2 + cx_4 + x_3x_4 - \frac{1}{2}x_4^2, \\
    f_3 &= bx_2 + dx_2^2 + (a + c - e)x_3 - (1 + 2d)x_2x_3 + \frac{1}{2}(1 + 2d - 2k)x_3^2 \\
    &\quad + ex_4 + x_2x_4 + 2kx_3x_4 + (\frac{1}{2} + 2d - 4d^2 - k)x_4^2, \\
    f_4 &= cx_2 + ex_3 + x_2x_3 + kx_3^2 + (a + b - e)x_4 - x_2x_4 \\
    &\quad + (1 + 4d - 8d^2 - 2k)x_3x_4 + (1 + k)x_4^2, \\
    h_1(x) &= x_1, h_2(x) = x_2 + x_3 + x_4,
\end{aligned}
\]

and \( a, b, c, d, e, k \) are constants. Then

\[ \omega_{ij} = 0, 1 \leq i \leq 4, 1 \leq j \leq 4, \]
\[
\frac{1}{2} \frac{\partial \eta}{\partial x_1} = 2x_1,
\]
\[
\frac{1}{2} \sum_{j=2}^{n} \frac{\partial \eta}{\partial x_j} = (a + b + c)^2(x_2 + x_3 + x_4)
\]

It is easy to check that \( \lambda = C_1 = C_2 = C_3 = C_4 = \tilde{\lambda} = \tilde{C}_0 = \tilde{C}_1 = \tilde{C}_2 = \tilde{C}_4 = 0, C_0 = 2, \tilde{C}_3 = (a + b + c)^2 \). The estimation algebra \( E \) is six-dimensional with basis \( \{1, x_1, D_1, x_2 + x_3 + x_4, D_2 + D_3 + D_4, L_0\} \).

### 2.6 Conclusions

Despite the success of the classification of finite dimensional estimation algebras with maximal rank, the problem of classification of non-maximal rank finite dimensional estimation algebras is still open except for the case of state space dimension 2 which is finished by Wu and Yau, see [34]. Due the difficulty of the problem, Brockett suggested that one should understand the low-dimensional estimation algebras. In [35] and [36], Yau and Rasoulian, Chiou, Chuieh, and Yau have classified estimation algebras at most five. In this paper, we give a structure theorem for estimation algebras of dimension six. Notice that there are two structures for estimation algebras. We also provide two family of six-dimensional estimation algebras. In the first example, the filtering model is the same as those filtering models of Rasoulian and Yau. In the second example, the filtering model is of Yau type, filtering model with nonlinear observations and drift terms equal to gradient vector field plus affine vector field. In the future work, we shall investigate whether there are more new classes of finite-dimensional estimation algebras, and new method to discuss the structure for low dimensional estimation algebras.
CHAPTER 3

NEW SUBOPTIMAL FILTER FOR POLYNOMIAL FILTERING PROBLEMS

3.1 Introduction

Although Kalman filter and Kalman-Bucy filter are widely used in many areas such as navigational and guidance systems, radar tracking, solar mapping, satellite and airplane orbit determination, and forecast in weather, econometrics, and finance, see [5], the use is still limited because of the linearity assumption and the Gaussian assumption of the initial value. However, the nonlinear filtering problem is an essentially more difficult problem since the resulting optimal filter is, in general, infinite-dimensional. A lot of work have been done to find certain suboptimal filters. Most widely used suboptimal filters include Extended Kalman filter (EKF), Unscented Kalman filter (UKF), Ensemble Kalman filter (EnKF), Particle filter (PF), and splitting up method, see [38], [39], [46], and [41]. Those suboptimal filters all have their own drawbacks, separately. We will provide some previous results and a brief introduction of these suboptimal filters as prerequisite. Then we introduce our new suboptimal filter with our special assumptions for polynomial filtering problems. We will also show numerical results which show that our suboptimal filter works perfectly.
3.2 Filtering Model And Basic Theorems

The model we consider here is:

\[
\begin{align*}
\text{dx}_t &= f(x_t)dt + g(x_t)dv_t \\
\text{dy}_t &= h(x_t)dt + dw_t,
\end{align*}
\]

in which \(x_t, v_t, y_t,\) and \(w_t\) are, respectively, \(\mathbb{R}^n, \mathbb{R}^p, \mathbb{R}^m,\) and \(\mathbb{R}^m,\) valued processes, and \(v_t\) and \(w_t\) have components that are independent, standard Brownian processes. We assume \(f_i, 1 \leq i \leq n,\)
\(g_{ij}, 1 \leq i \leq n, 1 \leq j \leq m,\) and \(h_j, 1 \leq j \leq n\) are polynomial functions of \(x_t.\) \(\{v_t, t \leq 0\}\) and \(\{w_t, t \leq 0\}\) are Brownian motion processes with \(\text{Var}[dv_t] = Q(t)dt\) and \(\text{Var}[dw_t] = R(t)dt,\)
separately. \(y_0 = 0.\) We further assume that \(\{v_t, t \leq 0\}, \{w_t, t \leq 0\}\) and \(x_0\) are independent and \(m = n.\) For simplicity, we could assume \(Q(t)\) is diagonal matrixes, \(Q(t) = \text{diag}(q_1^2, \ldots, q_n^2).\)

Actually, if \(Q(t)\) is not diagonal matrix, since we have spectral decompositi on of \(Q(t) = \Lambda P P',\)
where \(P P' = I, \Lambda\) is diagonal matrix, then we could let \(g^* = gP, dv^* = P' dv.\) Thus \(\text{Var}[dv^*] = \Lambda dt.\) We could further assume that \(Q(t) = I,\) since the \(Q(t)\) could be absorbed in the function
of \(g\) (replace \(g\) by \(gQ^{1/2}).\)

Researchers also could discretize the model using certain scheme, such as Euler scheme to obtain a discretized model.

For simplicity, we shall omit the dependent variables, \(x, x_t\) and \(t\) if no confusion occurs, i.e., \(x_i = x_{it}, f = f(x, t), f_t = f(x_t, t),\) et. al.. Let \(p \equiv p(x, t \mid Y_t)\) be the conditional probability
density function of the state $x_t$, given the observation history $Y_t \equiv \{y_s, 0 \leq s \leq t\}$, then the conditional expectation for some function of $x_t$ is defined as

$$\hat{(\circ)} \equiv E^t[(\circ)] \equiv E[(\circ) \mid Y_t].$$

(3.2.2)

The conditional mean of $x_i$, $1 \leq i \leq n$ is defined as

$$\hat{x}_i \equiv E^t[x_i] \equiv E[x_i \mid Y_t].$$

(3.2.3)

We denote $P_{k_1...k_n}$ as

$$P_{k_1...k_n} \equiv E^t[(x_1 - \hat{x}_1)^{k_1} \cdots (x_n - \hat{x}_n)^{k_n}] \equiv E[(x_1 - \hat{x}_1)^{k_1} \cdots (x_n - \hat{x}_n)^{k_n} \mid Y_t].$$

Notice that $P_{k_1...k_n} = 1$ when $k_i = 0$, $1 \leq i \leq n$; $P_{k_1...k_n} = 0$ when $n_i = 1$, $n_j = 0$, $1 \leq i \leq n$, $j \neq i$. $P_{k_1...k_n}$, $k_i = 2$, $k_j = 0$, $1 \leq i \leq n$, $j \neq i$ is the conditional variance of $x_i$, and $P_{k_1...k_n}$, $k_i = 1$, $k_j = 1$, $k_l = 0$, $1 \leq i, j \leq n$, $j \neq i, l \neq i, j$ is the conditional covariance of $x_i$ and $x_j$.

For convenience, we denote

$$P^\alpha_a = (x_a - \hat{x}_a)^\alpha$$

(3.2.4)
3.3 Some widely used suboptimal filters and previous results

3.3.1 Unscented Kalman filter

In subsection 3.3.1 to 3.3.3, we consider the corresponding discrete model of (3.2.1) using a certain scheme.

Unscented Kalman filter improves the extended Kalman filter by using a deterministic sampling approach [38]. The unscented Kalman filter approach is presented as following:

1. Initialization

\[ \hat{x}_0 = E[x_0] \] (3.3.1)
\[ P_0 = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T] \] (3.3.2)
\[ \hat{x}_0^a = E[x_0^a] = [\hat{x}_0^T 0 0]^T \] (3.3.3)
\[ P_0^a = E[(x_0^a - \hat{x}_0^a)(x_0^a - \hat{x}_0^a)^T] = \begin{bmatrix}
    P_0 & 0 & 0 \\
    0 & P_v & 0 \\
    0 & 0 & P_w
\end{bmatrix} \] (3.3.4)

2. Calculate sigma points

\[ X^a_k = [\hat{x}_k^a \hat{x}_k^a + (\sqrt{(L + \lambda)P_k^a})_{1:L} \hat{x}_k^a - (\sqrt{(L + \lambda)P_k^a})_{1:L}] \] (3.3.5)

3. Time update

\[ X^x_{k+1|k} = f(X^x_k) + X^a_k \] (3.3.6)
\[ \hat{x}_{k+1} = \sum_{i=0}^{2L} W_i^{(m)} \chi_{i,k+1|k}^x \]  
\[ P_{k+1} = \sum_{i=0}^{2L} W_i^{(c)} [\chi_{i,k+1|k}^x - \hat{x}_{k+1}] [\chi_{i,k+1|k}^x - \hat{x}_{k+1}]^T \]  
\[ \mathcal{Y}_{k+1|k} = h[\chi_{k+1|k}^x] + \chi_{k+1}^w \]  
\[ \hat{y}_{k+1} = \sum_{i=0}^{2L} W_i^{(m)} \mathcal{Y}_{i,k+1|k} \]  
\[ W_0^{(m)} = \frac{\lambda}{L + \lambda} \]  
\[ W_0^{(c)} = \frac{\lambda}{L + \lambda} + (1 - \alpha^2 + \beta) \]  
\[ W_i^{(m)} = W_i^{(c)} = \frac{1}{2(L + \lambda)}, \quad i = 1, \ldots, 2L \]  

(4) Measurement update:

\[ P_{\hat{y}_{k+1}\hat{y}_{k+1}} = \sum_{i=0}^{2L} W_i^{(c)} [\mathcal{Y}_{i,k+1|k}^x - \hat{y}_{k+1}] [\mathcal{Y}_{i,k+1|k}^x - \hat{y}_{k+1}]^T \]  
\[ P_{x_{k+1}y_{k+1}} = \sum_{i=0}^{2L} W_i^{(c)} [\chi_{i,k+1|k}^x - \hat{x}_{k+1}] [\chi_{i,k+1|k}^x - \hat{x}_{k+1}]^T \]  
\[ K_{k+1} = P_{x_{k+1}y_{k+1}} P_{y_{k+1}|y_{k+1}}^{-1} \]  
\[ \hat{x}_{k+1} = \hat{x}_{k+1} + K_{k+1} (y_{k+1} - \hat{y}_{k+1}) \]  
\[ P_{k+1} = P_{k+1} - K_{k+1} P_{y_{k+1}|y_{k+1}} K_{k+1}^T \]  
\[ P_{k+1}^a = \begin{bmatrix} P_{k+1} & 0 & 0 \\ 0 & P_v & 0 \\ 0 & 0 & P_w \end{bmatrix} \]
where \( x^a = [x^T v^T w^T]^T \), \( X^a = [(X^x)^T (X^v)^T (X^w)^T]^T \), \( \lambda = \alpha^2 (L + \kappa) - L \) is composite scaling parameter, \( \alpha \) determines the spread of the sigma points around \( \hat{x} \) and is usually set to a small positive value, \( \kappa \) is a secondary scaling parameter which is usually set to 0, \( \beta = 2 \) is optimal for Gaussian distribution, \( L \) is the dimension of augmented state, \( P_v \) is the initial state noise covariance matrix, \( P_w \) is the initial measurement noise covariance matrix. \( (\sqrt{(L + \lambda)P_a^k})_i \) is the \( i \)th row of the matrix square root.

The unscented Kalman filter is generally much better than EKF although the computational complexity are roughly the same. However, unscented Kalman filter is not always better than EKF. Therefore we need other methods.

### 3.3.2 Ensemble Kalman filter

The Ensemble Kalman filter is a suboptimal filter for nonlinear filter problem. It is presented in the following three steps:

1. **Ensemble step**

Assume at time \( k \), we have an ensemble of \( q \) forecasted state estimates with random sample errors. We denote this ensemble as \( X^f_k = (x^f_1, \cdots, x^f_q) \in \mathbb{R}^{n \times q} \). The superscript \( f_i \) refers the \( i \)-th forecast ensemble member. Then the ensemble mean is given by:

\[
\overline{x}_k = \frac{1}{q} \sum_{i=1}^{q} x^f_k \in \mathbb{R}^n.
\]
Now define the ensemble error matrix $E_{f_{k}}^{f} \in \mathbb{R}^{n \times q}$ by

$$E_{f_{k}}^{f} = \begin{bmatrix} x_{f_{1}}^{f_{k}} - \bar{x}_{f_{k}}, \ldots, x_{f_{q}}^{f_{k}} - \bar{x}_{f_{k}} \end{bmatrix}$$  \hspace{1cm} (3.3.20)$$

and the ensemble output error $E_{y_{k}}^{f} \in \mathbb{R}^{p \times q}$ by

$$E_{y_{k}}^{f} = \begin{bmatrix} y_{f_{1}}^{f_{k}} - \bar{y}_{f_{k}}, \ldots, y_{f_{q}}^{f_{k}} - \bar{y}_{f_{k}} \end{bmatrix},$$  \hspace{1cm} (3.3.21)$$

where $y_{f_{k}}^{f_{i}} = h(x_{f_{k}}^{f_{i}})$, and $\bar{y}_{f_{k}}^{f_{i}} = \frac{1}{q} \sum_{i=1}^{q} y_{f_{k}}^{f_{i}}$. We then approximate $P_{f_{k}}^{f_{i}}$ by $\hat{P}_{f_{k}}^{f_{i}}$, $P_{x_{y_{k}}}^{f_{i}}$ by $\hat{P}_{x_{y_{k}}}^{f_{i}}$, and $P_{y_{y_{k}}}^{f_{i}}$ by $\hat{P}_{y_{y_{k}}}^{f_{i}}$, respectively, where

$$\hat{P}_{f_{k}}^{f_{i}} = \frac{1}{q-1} E_{f_{k}}^{f_{i}} (E_{f_{k}}^{f_{i}})^{T}$$  \hspace{1cm} (3.3.22)$

$$\hat{P}_{x_{y_{k}}}^{f_{i}} = \frac{1}{q-1} E_{x_{y_{k}}}^{f_{i}} (E_{x_{y_{k}}}^{f_{i}})^{T}$$  \hspace{1cm} (3.3.23)$

$$\hat{P}_{y_{y_{k}}}^{f_{i}} = \frac{1}{q-1} E_{y_{y_{k}}}^{f_{i}} (E_{y_{y_{k}}}^{f_{i}})^{T}$$  \hspace{1cm} (3.3.24)$

(2) Analysis step

Let

$$x_{k}^{a_{i}} = x_{k}^{f_{i}} + \hat{K}_{k}(y_{k}^{f_{i}} - h(x_{k}^{f_{i}})),$$  \hspace{1cm} (3.3.25)$
where $y^i_k = y_k + w^i_k$, is the perturbed observation. $w^i_k$'s follow a normal distribution $N(0, R_k)$.

We use $\hat{P}^a_k$ to approximate $P^a_k$, where

$$
\hat{P}^a_k = \frac{1}{q-1} E^a_k (E^a_k)^T, \quad (3.3.26)
$$

where

$$
E^a_k = \begin{bmatrix} x^{a_1}_k - \overline{x}_k, \cdots, x^{a_q}_k - \overline{x}_k \end{bmatrix}, \quad (3.3.27)
$$

$$
\overline{x}_k = \frac{1}{q} \sum_{i=1}^q x^{a_i}_k, \quad (3.3.28)
$$

which is a mean of analysis estimate ensemble members. We use $\hat{K}_k$ to express the filter gain which is approximated by the classic Kalman gain:

$$
\hat{K}_k = \hat{P}^f_{yk} (\hat{P}^f_{yy})^{-1}. \quad (3.3.29)
$$

(3) Forecast step

The forecast estimate ensemble members are given by:

$$
x^{f_i}_{k+1} = f(x^{a_i}_k) + v^i_k, \quad (3.3.30)
$$
where $v^i_k$’s follow a normal distribution $N(0, Q_k)$. Now the EnKF could be summarized as following:

\[
\hat{K}_k = \hat{P}^f_{xy_k} (\hat{P}^f_{yy_k})^{-1},
\]

\[
x^a_k = x^f_k + \hat{K}_k (y_k + v^i_k - h(x^f_k)),
\]

\[
\bar{x}_k = \frac{1}{q} \sum_{i=1}^{q} x^a_i,
\]

\[
x^f_{k+1} = f(x^a_k) + v^i_k,
\]

\[
\bar{x}^f_{k+1} = \frac{1}{q} \sum_{i=1}^{q} x^f_{k+1}
\]

where $E^f_k, E^a_k, \hat{P}^f_{xy_k}, \hat{P}^f_{yy_k}$ are given by (3.3.20), (3.3.21), (3.3.23), and (3.3.24). Notice that this algorithm does not involve the Jacobian matrix, which means it does not involve linearity approximation.

EnKF requires more ensembles if we want a better approximation, which will result a longer computation time.

### 3.3.3 Particle filter

Ensemble Kalman filter and Particle filter are both easy to implement. One can code a particle filter algorithm in 2 pages with MATLAB. But particle filters are more applied at low-dimensional systems, see [42] for the obstacles to high dimensional cases. We only introduce a widely used particle filter – Sampling Importance Resampling (SIR) filter. For more details, refer [47] and [40].
Algorithm 1 SIR Particle Filter

\[ \{x_k^i, \omega_k^i\}_{i=1}^{N_s} = \text{SIR}\{x_{k-1}^i, \omega_{k-1}^i\}_{i=1}^{N_s}, z_k \]

- FOR \( i = 1 : N_s \)
  - Draw \( x_k^i \sim p(x_k | x_{k-1}^i) \)
  - Calculate \( \omega_k^i = p(z_k | x_k^i) \)
- END FOR

- Calculate total weight: \( t = \text{SUM}\{\omega_k^i\}_{i=1}^{N_s} \)

- FOR \( i = 1 : N_s \)
  - Normalize: \( \omega_k^i = t^{-1} \omega_k^i \)
- END FOR

- Resample using algorithm 2:
  \[ -\{x_k^i, \omega_k^i, -\}_{i=1}^{N_s} = \text{RESAMPLE}\{x_k^i, \omega_k^i\}_{i=1}^{N_s} \]
Algorithm 2 Resampling Algorithm

\[ \{x^*_k, \omega^i_k, i^j\}_{j=1}^{N_s} \] \(=\) \text{RESAMPLE}[\{x^i_k, \omega^i_k\}_{i=1}^{N_s}]

- Initialize the CDF: \( c_1 = \omega^1_k \)
- FOR \( i = 2 : N_s \)
  - Construct CDF \( c_i = c_{i-1} + \omega^i_k \)
- END FOR
- Start at the bottom of the CDF: \( i = 1 \)
- Draw a starting point: \( u_1 \sim U[0, N_s^{-1}] \)
- FOR \( j = 1 : N_s \)
  - Move along the CDF: \( u_j = u_1 + N_s^{-1}(j - 1) \)
  - WHILE \( u_j > c_i \)
    * \( i = i + 1 \)
  - END WHILE
  - Assign sample: \( x^*_k = x^i_k \)
  - Assign weight: \( \omega^j_k = N_s^{-1} \)
  - Assign parent: \( i^j = i \)
- END FOR
Although the algorithm is easy to implement, one of the drawbacks is that a few particles may absorb all the weights. There are many remedy to modify this algorithm, such as sequential importance resampling, and variance reduction. However, we still need to notice that, if we want to obtain more accuracy, we need to have bigger sample size, which may result a longer computation time.

3.3.4 Some previous results for nonlinear filtering problems

**Theorem 3.3.1.** For the continuous filtering problems given by the system (3.2.1), the conditional density \( p(x,t) \) satisfies the following Kushner’s equation \([6] [7]\)

\[
d p(x,t) = L^* p dt + p(x,t) (dy - \hat{h} dt)' R^{-1} (h - \hat{h}) ,
\]

(3.3.32)

where

\[
L^* p(x,t) = - \sum_i \frac{\partial (f_i p(x,t))}{\partial x_i} + \frac{1}{2} \sum_{i,j} \frac{\partial^2 v_{ij}}{\partial x_i \partial x_j},
\]

which is the joint of \( L \),

\[
L = \sum_i f_i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j} v_{ij} \frac{\partial^2}{\partial x_i \partial x_j},
\]

where \( \{v_{ij}\} = g'g = V \).

**Theorem 3.3.2.** For the continuous filtering problems given by the system (3.2.1), the conditional density \( \hat{x}_i \) satisfies the following stochastic differential equation

\[
d\hat{x}_i = \hat{f}_i dt + (dy - \hat{h} dt)' R^{-1} (h \hat{x}_i - \hat{h} \hat{x}_i) ,
\]

(3.3.33)
Theorem 3.3.3. For the continuous filtering problems given by the system (3.2.1), EKF is given by

\[
\begin{align*}
    d\hat{x}(t) &= f(\hat{x}(t))dt + P(t)M^T(t)R^{-1}(t)(dy(t) - h(\hat{x}(t)))dt \\
    \frac{dP(t)}{dt} &= F(t)P(t) + P(t)F^T(t) + g(t)Q(t)g^T(t) - P(t)M^T(t)R^{-1}(t)M(t)P(t),
\end{align*}
\]

where \( F(t) \) and \( M(t) \) are the Jacobian matrix defined by:

\[
F(t) = \left[ \frac{\partial f_i(\hat{x}(t))}{\partial x_j} \right], \quad M(t) = \left[ \frac{\partial h_i(\hat{x}(t))}{\partial x_j} \right].
\]

3.4 New suboptimal filter for polynomial filtering problems

Let \( f_i(x, t)'s, g_{i,j}(x, t)'s, h_i(x, t)'s \) be polynomials of degree \( M_f, M_g, M_h \), respectively defined as follows

\[
\begin{align*}
    f_i(x, t) &= \sum_{m_l \leq M_f} f_{i; m_1, \ldots, m_n} \prod_{a=1}^{n} x_a^{m_a} \\
    g_{i,j}(x, t) &= \sum_{m_l \leq M_g} g_{i,j; m_1, \ldots, m_n} \prod_{a=1}^{n} x_a^{m_a} \\
    h_i(x, t) &= \sum_{m_l \leq M_h} h_{i; m_1, \ldots, m_n} \prod_{a=1}^{n} x_a^{m_a}
\end{align*}
\]

where \( M_f \) is the highest degree among \( f_i \)'s, \( M_g \) is the highest degree among \( g_{i,j} \)'s, \( M_h \) is the highest degree among \( h_i \)'s, \( f_{i; m_1, \ldots, m_n} \)'s, \( g_{i,j; m_1, \ldots, m_n} \)'s, \( h_{i; m_1, \ldots, m_n} \)'s are functions of \( t \). \( \sum_{m_l \leq M_f} \) means we sum every term in the form of \( x_1^{m_1}x_2^{m_2}...x_n^{m_n} \), where \( m_l \leq M_f, 1 \leq l \leq n. \)
Then using Taylor expansion,

\[
\prod_{a=1}^{n} x_a^{m_a} = \prod_{a=1}^{n} (x_a - \hat{x}_a + \hat{x}_a)^{m_a}
\]

\[
= \prod_{a=1}^{n} \sum_{k_a=0}^{m_a} \binom{m_a}{k_a} (x_a - \hat{x}_a)^{k_a}(\hat{x}_a)^{m_a-k_a}
\]

\[
= \sum_{0 \leq k_a \leq m_a} \prod_{a=1}^{n} \binom{m_a}{k_a} (x_a - \hat{x}_a)^{k_a}(\hat{x}_a)^{m_a-k_a}
\]

(3.4.4)

Then

\[
f_i(x, t) = \sum_{m_i \leq M_f} \sum_{0 \leq k_a \leq m_a} f_i;m_1,\ldots,m_n \prod_{a=1}^{n} \binom{m_a}{k_a} (x_a - \hat{x}_a)^{k_a}(\hat{x}_a)^{m_a-k_a}
\]

(3.4.5)

\[
h_i(x, t) = \sum_{m_i \leq M_h} \sum_{0 \leq k_a \leq m_a} h_i;m_1,\ldots,m_n \prod_{a=1}^{n} \binom{m_a}{k_a} (x_a - \hat{x}_a)^{k_a}(\hat{x}_a)^{m_a-k_a}
\]

(3.4.6)

\[
g_{i,j}(x, t) = \sum_{m_i \leq M_g} \sum_{0 \leq k_a \leq m_a} g_{i,j};m_1,\ldots,m_n \prod_{a=1}^{n} \binom{m_a}{k_a} (x_a - \hat{x}_a)^{k_a}(\hat{x}_a)^{m_a-k_a}
\]

(3.4.7)

\[
g_{i_1,j_1}g_{i_2,j_2} = \left( \sum_{m_{i_1} \leq M_g} g_{i_1,j_1};m_{i_1} \prod_{a=1}^{n} x_a^{m_{i_1}a} \right) \left( \sum_{m_{i_2} \leq M_g} g_{i_2,j_2};m_{i_2} \prod_{a=1}^{n} x_a^{m_{i_2}a} \right)
\]

\[
= \sum_{m_{i_1},m_{i_2} \leq M_g} g_{i_1,j_1};m_{i_1},g_{i_2,j_2};m_{i_2} \prod_{a=1}^{n} x_a^{m_{i_1}a+m_{i_2}a}
\]

\[
= \sum_{m_{i_1},m_{i_2} \leq M_g} \sum_{0 \leq k_a \leq m_{i_1}a+m_{i_2}a} g_{i_1,j_1};m_{i_1},g_{i_2,j_2};m_{i_2} \prod_{a=1}^{n} \binom{m_a}{k_a} (x_a - \hat{x}_a)^{k_a}(\hat{x}_a)^{m_a-k_a}
\]

(3.4.8)
The conditional expectation of the above equations are given by

\[
\tilde{f}_i(x, t) = \sum_{m_i \leq M_f} \sum_{0 \leq k_a \leq m_a} f_{i; m_1, \ldots, m_n} \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{k_1, \ldots, k_n}
\]

\[
\tilde{h}_i(x, t) = \sum_{m_i \leq M_h} \sum_{0 \leq k_a \leq m_a} h_{i; m_1, \ldots, m_n} \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{k_1, \ldots, k_n}
\]

\[
\tilde{g}_{i,j}(x, t) = \sum_{m_i \leq M_g} \sum_{0 \leq k_a \leq m_a} g_{i,j; m_1, \ldots, m_n} \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{k_1, \ldots, k_n}
\]

\[
g_{i_1,j_1}g_{i_2,j_2} = \left( \sum_{m_i \leq M_g} \sum_{0 \leq k_a \leq m_a} g_{i_1,j_1; m_1, \ldots, m_n} \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{k_1, \ldots, k_n} \right) \left( \sum_{m_i \leq M_g} \sum_{0 \leq k_a \leq m_a} g_{i_2,j_2; m_1, \ldots, m_n} \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{k_1, \ldots, k_n} \right)
\]

Then we have

\[
\prod_{a=1}^{n} (x_a - \hat{x}_a)^{\alpha_a} f_i = \sum_{m_i \leq M_f} \sum_{0 \leq k_a \leq m_a} f_{i; m_1, \ldots, m_n} \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (x_a - \hat{x}_a)^{\alpha_a+k_a} (\hat{x}_a)^{m_a-k_a} \right)
\]

\[
\prod_{a=1}^{n} (x_a - \hat{x}_a)^{\alpha_a} h_i = \sum_{m_i \leq M_h} \sum_{0 \leq k_a \leq m_a} h_{i; m_1, \ldots, m_n} \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (x_a - \hat{x}_a)^{\alpha_a+k_a} (\hat{x}_a)^{m_a-k_a} \right)
\]

\[
\prod_{a=1}^{n} (x_a - \hat{x}_a)^{\alpha_a} g_{i_1,j_1}g_{i_2,j_2} = \sum_{m_i \leq M_g} \sum_{0 \leq k_a \leq m_a} g_{i_1,j_1; m_1, \ldots, m_n} g_{i_2,j_2; m_1, \ldots, m_n} \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (x_a - \hat{x}_a)^{\alpha_a+k_a} (\hat{x}_a)^{m_a-k_a} \right)
\]
Then we take the conditional expectation of the above equations

\[
E^t \left[ \prod_{a=1}^{n} (x_a - \hat{x}_a)^{\alpha_a} f_i \right] = \sum_{m_i \leq M_k} \sum_{0 \leq k_a \leq m_a} f_i;m_1,\ldots,m_n \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{a_1+k_1,\ldots,a_n+k_n}
\]

(3.4.12)

\[
E^t \left[ \prod_{a=1}^{n} (x_a - \hat{x}_a)^{\alpha_a} h_i \right] = \sum_{m_i \leq M_k} \sum_{0 \leq k_a \leq m_a} h_i;m_1,\ldots,m_n \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{a_1+k_1,\ldots,a_n+k_n}
\]

(3.4.13)

\[
E^t \left[ \prod_{a=1}^{n} (x_a - \hat{x}_a)^{\alpha_a} g_{i_{1,j_1}} g_{i_{2,j_2}} \right] = \sum_{m_i \leq M_k} \sum_{0 \leq k_a \leq m_a} g_{i_{1,j_1}};m_1,\ldots,m_n g_{i_{2,j_2}};m_1,\ldots,m_n \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{a_1+k_1,\ldots,a_n+k_n}
\]

(3.4.14)

Denote \( R^{-1} = (r_{ij})_{n \times n} \). Using these equations (3.4.12)-(3.4.14) we obtain the equation about the conditional mean \( \hat{x}_i \)

\[
d\hat{x}_i = \sum_{m_i \leq M_k} \sum_{0 \leq k_a \leq m_a} f_i;m_1,\ldots,m_n \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{k_1,\ldots,k_n} dt
\]

\[
+ \left[ dy_i - \sum_{m_i \leq M_k} \sum_{0 \leq k_a \leq m_a} h_i;m_1,\ldots,m_n \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{k_1,\ldots,k_n} \right] R^{-1}
\]

\[
= \sum_{m_i \leq M_k} \sum_{0 \leq k_a \leq m_a} f_i;m_1,\ldots,m_n \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{k_1,\ldots,k_n} dt
\]

\[
+ \sum_{i,j} r_{i,j} \left[ dy_i - \sum_{m_i \leq M_k} \sum_{0 \leq k_a \leq m_a} h_i;m_1,\ldots,m_n \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a-k_a} \right) P_{k_1,\ldots,k_n} \right]
\]
To obtain the stochastic differential equation for $P_{\alpha_1,\ldots,\alpha_n}$, we need the following lemma:

**Lemma 3.4.1** (Itô). Let $x_t$ be the unique solution of the vector Itô stochastic differential equation

$$dx_t = f(x_t,t)dt + G(x_t,t)d\beta_t, \quad t \geq t_0$$

where $x$ and $f$ are $n$-vectors, $G$ is $n \times m$, and $\{\beta_t, t \geq t_0\}$ is an $m$-vector Brownian motion process with $E\{d\beta_t d\beta_t^T\} = Q(t)dt$. Let $\phi(x_t,t)$ be a scalar-values real function, continuously in $t$ and having continuous second mixed partial derivatives with respect to the elements of $x$.

Then the (stochastic) differential $d\phi$ of $\phi$ is

$$d\phi = \phi_t dt + \phi_x^T dx_t + \frac{1}{2} tr GQG^T \phi_{xx} dt,$$
Let \( x_{i,t}, i = 1, 2, \ldots, n \) be the unique solution of the Itô equations

\[
dx_{i,t} = f_i(y_{i,t}, t) dt + G_i(x_{i,t}, t) d\beta_t, \quad t \geq t_0
\]

where \( \beta_t, t \geq t_0 \), here and in (3.4.16), is the same Brownian Motion process. Then, by considering the joint process

\[
dz_t = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} dt + \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} d\beta_t, \quad z_t = [x_{1,t}, x_{2,t}]^T,
\]

It is easy to prove the following corollaries of Lemma 3.4.1.

**Corollary 3.4.2.** Let the product \( \phi_1 \phi_2 \) of the scalar function \( \phi_1(x_{1,t}, t) \) and \( \phi_2(x_{2,t}, t) \) satisfy the hypotheses of Lemma 3.4.1. Then

\[
d(\phi_1 \phi_2) = \phi_1 d\phi_2 + \phi_2 d\phi_1 + trG_1 QG_2^T \phi_{x_{2,t}} \phi_{x_{1,t}}^T dt = \phi_1 d\phi_2 + \phi_2 d\phi_1 + d\phi_1 d\phi_2.
\]

Similarly, we have next Corollary.

**Corollary 3.4.3.** Let the product \( \prod_{i=1}^n \phi_i \) of the scalar function \( \phi_i(x_{i,t}, t) \)'s satisfies the hypotheses of Lemma 3.4.1. Then

\[
d(\prod_{i=1}^n \phi_i) = \sum_{i=1}^n (\prod_{j \neq i} \phi_j) d\phi_i + \sum_{i<j} (\prod_{k \neq i,j} \phi_k) d\phi_i d\phi_j.
\]

(3.4.20)
From the definition of $P_{\alpha_1,\ldots,\alpha_n}$, we have

$$P_{\alpha_1,\ldots,\alpha_n} = E^t \prod_{i=1}^n (x_i - \hat{x}_i)^{\alpha_i}$$

$$= \int \prod_{i=1}^n (x_i - \hat{x}_i)^{\alpha_i} p(x,t) dx$$

Then we apply Corollary 3.4.3, we have

$$dP_{\alpha_1,\ldots,\alpha_n} = \sum_{a=1}^n \int (\prod_{b \neq a} P_{b}^{\alpha_b}) dP_{a}^{\alpha_a} p(x,t) dx + \int (\prod_{b=1}^n P_{b}^{\alpha_b}) dp(x,t) dx$$

$$+ \sum_{a<b} \int (\prod_{c \neq a,b} P_{c}^{\alpha_c}) dP_{a}^{\alpha_a} dP_{b}^{\alpha_b} p(x,t) dx + \sum_{a=1}^n \int (\prod_{b \neq a} P_{b}^{\alpha_b}) dP_{a}^{\alpha_a} dp(x,t) dx$$

$$= I + II + III + IV, \quad (3.4.21)$$

where $I, II, III, IV$ represent each term separately. Notice

$$dP_{\alpha_a} = -\alpha_a (x_a - \hat{x}_a)^{\alpha_a-1} d\hat{x}_a + \frac{1}{2} \alpha_a (\alpha_a - 1) (x_a - \hat{x}_a)^{\alpha_a-2} (d\hat{x}_a)^2$$

$$= -\alpha_a (x_a - \hat{x}_a)^{\alpha_a-1} \hat{f}_a dt - \alpha_a (x_a - \hat{x}_a)^{\alpha_a-1} (dy - \hat{h} dt)' R^{-1} (\hat{h} x_a - \hat{h} \hat{x}_a)$$

$$+ \frac{1}{2} \alpha_a (\alpha_a - 1) (x_a - \hat{x}_a)^{\alpha_a-2} (\hat{h} x_a - \hat{h} \hat{x}_a)' R^{-1} (\hat{h} x_a - \hat{h} \hat{x}_a) dt \quad (3.4.22)$$

where we need to recall

$$d\hat{x}_i = \hat{f}_i dt + (dy - \hat{h} dt)' R^{-1} (\hat{h} x_i - \hat{h} \hat{x}_i)$$
From the above equations, we could derive the following expressions using Itô’s lemma.

\[
dp(x, t) = L^*pd\mathbf{t} + p(x, t)(dy - \widehat{h}dt)'R^{-1}(h - \widehat{h})
\]

\[
L = \sum_i f_i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j} v_{ij} \frac{\partial^2}{\partial x_i \partial x_j}
\]

\[
L^*p(x, t) = -\sum_i \frac{\partial(f_i p(x, t))}{\partial x_i} + \frac{1}{2} \sum_{i,j} \frac{\partial^2(v_{ij} p(x, t))}{\partial x_i \partial x_j}
\]

where \(L^*\) is the formal adjoint of \(L\), and for suitable \(q(x)\),

\[
\int q(x)L^*p(x, t)dx = \int (Lq(x))p(x, t)dx,
\]

and \(\{v_{ij}\} = g'g = V\).

From the above equations, we could derive the following expressions using Itô’s lemma.

\[
d\widehat{x}_i \widehat{x}_j = (\widehat{hx}_i - \widehat{h}\widehat{x}_j)'R^{-1}(\widehat{hx}_j - \widehat{h}\widehat{x}_j)dt \tag{3.4.23}
\]

\[
d\widehat{x}_i dp(x, t) = (\widehat{hx}_i - \widehat{h}\widehat{x}_i)'R^{-1}(h - \widehat{h})p(x, t)dt \tag{3.4.24}
\]

Then we have

\[
I = \sum_{a=1}^{n} \left( \int \left( \prod_{b \neq a} P_{ab} \right) \left( -\alpha_a(x_a - \widehat{x}_a)^{\alpha_a-1}\widehat{f}_a dt 
- \frac{1}{2}\alpha_a(\alpha_a - 1)(x_a - \widehat{x}_a)^{\alpha_a-2}(\widehat{hx}_a - \widehat{h}\widehat{x}_a)'R^{-1}(\widehat{hx}_a - \widehat{h}\widehat{x}_a)dt \right) p(x, t)dx \right.

= \sum_{a=1}^{n} \left( -\alpha_a \widehat{f}_a P_{a1,...,a_{a-1},a_{a+1},...,a_n} dt 
- \alpha_a(dy - \widehat{h}dt)'R^{-1}(\widehat{hx}_a - \widehat{h}\widehat{x}_a)P_{a1,...,a_{a-1},a_{a+1},...,a_n} \right)
\]
\[
\begin{align*}
&+ \frac{1}{2} \alpha_a (\alpha_a - 1) (\tilde{h}x_a - \tilde{h}\tilde{x}_a)' R^{-1} (\tilde{h}x_a - \tilde{h}\tilde{x}_a) P_{\alpha_a, \ldots, \alpha_{a-1}, \alpha_a - 2, \alpha_{a+1}, \ldots, \alpha_n} dt \\
= &- \sum_{a=1}^{n} \alpha_a \hat{f}_a P_{\alpha_a, \ldots, \alpha_{a-1}, \alpha_a - 1, \alpha_{a+1}, \ldots, \alpha_n} dt \\
&- (dy - \hat{h}dt)' R^{-1} \sum_{a=1}^{n} \alpha_a (\tilde{h}x_a - \tilde{h}\tilde{x}_a) P_{\alpha_a, \ldots, \alpha_{a-1}, \alpha_a - 1, \alpha_{a+1}, \ldots, \alpha_n} \\
&+ \frac{1}{2} \sum_{a=1}^{n} \alpha_a (\alpha_a - 1) (\tilde{h}x_a - \tilde{h}\tilde{x}_a)' R^{-1} (\tilde{h}x_a - \tilde{h}\tilde{x}_a) P_{\alpha_a, \ldots, \alpha_{a-1}, \alpha_a - 2, \alpha_{a+1}, \ldots, \alpha_n} dt \\
= &- \sum_{a=1}^{n} \alpha_a \sum_{m_l \leq M_1} \sum_{0 \leq k_b \leq m_b} f_{a,m_1, \ldots, m_n} \left( \prod_{b=1}^{n} \left( \frac{m_b}{k_b} \right) (\tilde{x}_b)^{m_b - k_b} \right). \\
\end{align*}
\]

\[P_{k_1, \ldots, k_n} P_{\alpha_1, \ldots, \alpha_{a-1}, \alpha_a - 1, \alpha_{a+1}, \ldots, \alpha_n} dt \]

\[- (dy - \hat{h}dt)' R^{-1} \sum_{a=1}^{n} \alpha_a \cdot \]

\[\left[ \sum_{m_l \leq M_1} \sum_{0 \leq k_b \leq m_b} h_{i;m_1, \ldots, m_n} \left( \prod_{b=1}^{n} \left( \frac{m_b}{k_b} \right) (\tilde{x}_b)^{m_b - k_b} \right) P_{k_1, \ldots, k_{a-1}, k_a + 1, k_{a+1}, \ldots, k_n} \right]_{n \times 1}. \]

\[P_{\alpha_1, \ldots, \alpha_{a-1}, \alpha_a - 1, \alpha_{a+1}, \ldots, \alpha_n} \]

\[+ \frac{1}{2} \sum_{a=1}^{n} \alpha_a (\alpha_a - 1) \left( \sum_{i,j} r_{ij} \left( \sum_{m_l \leq M_1} \sum_{0 \leq k_b \leq m_b} h_{i;m_1, \ldots, m_n} \left( \prod_{b=1}^{n} \left( \frac{m_b}{k_b} \right) (\tilde{x}_b)^{m_b - k_b} \right) \right) \right) \times \]

\[\left( \sum_{m_l \leq M_1} \sum_{0 \leq k_b \leq m_b} h_{j;m_1, \ldots, m_n} \left( \prod_{b=1}^{n} \left( \frac{m_b}{k_b} \right) (\tilde{x}_b)^{m_b - k_b} \right) \right). \]

\[P_{k_1, \ldots, k_{a-1}, k_a + 1, k_{a+1}, \ldots, k_n} \]  \[P_{\alpha_1, \ldots, \alpha_{a-1}, \alpha_a - 2, \alpha_{a+1}, \ldots, \alpha_n} dt \]

(3.4.25)
\[ \text{III} = \sum_{a < b} \left( \int \prod_{c \neq a, b} P_c^{\alpha_c} \alpha_a \alpha_b (x_a - \hat{x}_a)^{\alpha_a-1} (x_b - \hat{x}_b)^{\alpha_b-1}. \right. \\
\left. (\hat{h} x_a - \hat{h} \hat{x}_a)' R^{-1} (\hat{h} x_b - \hat{h} \hat{x}_b) p(x, t) dx \right) dt \\
= \sum_{a < b} \left( \alpha_a \alpha_b P_{a1, \ldots, a_{a-1}, a_{a+1}, \ldots, a_{b-1}, a_{b+1}, \ldots, a_n} (\hat{h} x_a - \hat{h} \hat{x}_a)' R^{-1} (\hat{h} x_b - \hat{h} \hat{x}_b) \right) dt \\
= \sum_{a < b} \left( \alpha_a \alpha_b P_{a1, \ldots, a_{a-1}, a_{a+1}, \ldots, a_{b-1}, a_{b+1}, \ldots, a_n} \right. \\
\left. \left( \sum_{i,j} \sum_{m \leq M_h} \sum_{0 \leq k \leq m_e} h_{i; m_1, \ldots, m_n} \left( \prod_{c=1}^{n} \left( \frac{m_c}{k_c} \right) (\hat{x}_c)^{m_c - k_c} P_{k_1, \ldots, k_{a-1}, k_{a+1}, \ldots, k_{n}} \right) \right) dt \right) \quad (3.4.26) \\
\[ IV = -\sum_{a=1}^{n} \left( \int \prod_{b \neq a} P_b^{\alpha_b} \alpha_a P_a^{\alpha_a-1} \hat{x}_a dp(x, t) dx \right) \\
= -\sum_{a=1}^{n} \left( \int \prod_{b \neq a} P_b^{\alpha_b} \alpha_a P_a^{\alpha_a-1} (\hat{h} x_a - \hat{h} \hat{x}_a)' R^{-1} (h - \hat{h}) p(x, t) dx \right) dt \\
= -\sum_{a=1}^{n} \left( \alpha_a (\hat{h} x_a - \hat{h} \hat{x}_a)' R^{-1} \left( \int \prod_{b \neq a} P_b^{\alpha_b} P_a^{\alpha_a-1} (h - \hat{h}) p(x, t) dx \right) dt \right) \\
= -\sum_{a=1}^{n} \left( \alpha_a (\hat{h} x_a - \hat{h} \hat{x}_a)' R^{-1} \left( \int \prod_{b \neq a} P_b^{\alpha_b} P_a^{\alpha_a-1} \right) \right. \\
\left. \left[ \sum_{m_1 \leq M_h} \sum_{0 \leq k \leq m_e} h_{i; m_1, \ldots, m_n} \left( \prod_{c=1}^{n} \left( \frac{m_c}{k_c} \right) (\hat{x}_c)^{m_c - k_c} P_{k_1, \ldots, k_{n}} \right) \right]_{n \times 1} \right) \cdot \\
\left. p(x, t) dx \right) dt \]
\[
\begin{align*}
&= -\sum_{a=1}^{n} \left( \alpha_a \left[ \sum_{m_1 \leq M_b} \sum_{0 \leq b \leq m_b} h_{i,m_1,...,m_n} \left( \prod_{b=1}^{m_b} \left( m_b - k_b \right) \right) \left( \hat{x}_b \right)^{m_b - k_b} \right] \right)_{n \times 1}.
\end{align*}
\]

\[
R^{-1} \left[ \sum_{m_1 \leq M_b} \sum_{0 \leq b \leq m_b} h_{i,m_1,...,m_n} \left( \prod_{b=1}^{m_b} \left( m_b - k_b \right) \right) \right]_{n \times 1} \right) dt
\]

\[
= -\sum_{a=1}^{n} \left( \alpha_a \sum_{i,j} r^{ij} \left[ \sum_{m_i \leq M_b} \sum_{0 \leq b \leq m_b} h_{i,m_1,...,m_n} \left( \prod_{b=1}^{m_b} \left( m_b - k_b \right) \right) \right] \right)_{n \times 1} \right) dt
\]

\[
\left( \sum_{m_1 \leq M_b} \sum_{0 \leq b \leq m_b} h_{i,m_1,...,m_n} \left( \prod_{b=1}^{m_b} \left( m_b - k_b \right) \right) \right)_{n \times 1} \right) dt
\]

\[
(P_{a_1+k_1,...,a_{n-1}+k_{a-1},a_a+k_a-1,a_{a+1}+k_{a+1},...+a_n+k_n}
\]

\[
- P_{a_1,...,a_{n-1},a_{a+1},...+a_n}(P_{k_1,...,k_n}) \right) dt
\]

(3.4.27)

Now we take a look at II.

\[
\begin{align*}
II &= \int \left( \prod_{b=1}^{n} P_{a_b}^{\alpha_b} \right) L^*(p(x,t)) dx dt + \int \left( \prod_{b=1}^{n} P_{a_b}^{\alpha_b} \right) p(x,t) (dy - \hat{\nu} dt)' R^{-1}(h - \hat{\nu}) dx
d
\end{align*}
\]

\[
= \int L \left( \prod_{b=1}^{n} P_{a_b}^{\alpha_b} \right) p(x,t) dx dt + (dy - \hat{\nu} dt)' R^{-1} \int \left( \prod_{b=1}^{n} P_{a_b}^{\alpha_b} \right) (h - \hat{\nu}) dx
d
\]

\[
= \int \left( \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} v_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} \left( \prod_{b=1}^{n} P_{a_b}^{\alpha_b} \right) \right) p(x,t) dx dt
d
\]

\[
+ (dy - \hat{\nu} dt)' R^{-1} \int \left( \prod_{b=1}^{n} P_{a_b}^{\alpha_b} \right) (h - \hat{\nu}) p(x,t) dx
d
\]

\[
= \int \left( \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i} + \sum_{i<j} v_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} + \frac{1}{2} \sum_{i=1}^{n} v_{i,i} \frac{\partial^2}{\partial x_i^2} \left( \prod_{b=1}^{n} P_{a_b}^{\alpha_b} \right) \right) p(x,t) dx dt
\]
\[ + (dy - \hat{h} dt)' R^{-1} \int \left( \prod_{b=1}^{n} P_{b}^{\alpha_{b}} \right) (h - \hat{h}) p(x, t) dx \]

\[ = A + B + C + D \quad (3.4.28) \]

where

\[ A = \int \left( \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i} \left( \prod_{b=1}^{n} P_{b}^{\alpha_{b}} \right) \right) p(x, t) dx dt \]

\[ = \int \left( \sum_{i=1}^{n} f_i \left( \prod_{b \neq i} P_{b}^{\alpha_{b}} \right) n_i P_{i}^{n_{i}-1} \right) p(x, t) dx dt \]

\[ = \sum_{i=1}^{n} \int \left( \sum_{m_{i} \leq M_{i}} \sum_{0 \leq k_{a} \leq m_{a}} f_{i; m_{i}, \ldots, m_{n}} \left( \prod_{a=1}^{n} \binom{m_{a}}{k_{a}} P_{a}^{k_{a} \left( \hat{x}_{a} \right)^{m_{a} - k_{a}}} \right) \left( \prod_{b \neq i} P_{b}^{\alpha_{b}} \right) n_{i} P_{i}^{n_{i}-1} \right) p(x, t) dx dt \]

\[ = \sum_{i=1}^{n} \int \left( \sum_{m_{i} \leq M_{i}} \sum_{0 \leq k_{a} \leq m_{a}} f_{i; m_{i}, \ldots, m_{n}} \left( \prod_{a=1}^{n} \binom{m_{a}}{k_{a}} \left( \hat{x}_{a} \right)^{m_{a} - k_{a}} \right) \left( \prod_{b \neq i} P_{b}^{\alpha_{b} + k_{a}} \right) n_{i} P_{i}^{n_{i} + k_{i}-1} \right) p(x, t) dx dt \]

\[ = \sum_{i=1}^{n} \sum_{m_{i} \leq M_{i}} \sum_{0 \leq k_{a} \leq m_{a}} n_{i} f_{i; m_{i}, \ldots, m_{n}} \left( \prod_{a=1}^{n} \binom{m_{a}}{k_{a}} \left( \hat{x}_{a} \right)^{m_{a} - k_{a}} \right) \left( P_{\alpha_{1}+k_{1}, \ldots, n_{i}-1+k_{i-1}, i_{i}, n_{i}+1+k_{i+1}, \ldots, \alpha_{n}+k_{n}} \right) dt \quad (3.4.29) \]

\[ B = \int \left( \sum_{i \neq j} v_{i,j} \frac{\partial^2}{\partial x_{i} \partial x_{j}} \left( \prod_{b=1}^{n} P_{b}^{\alpha_{b}} \right) \right) p(x, t) dx dt \]

\[ = \int \left( \sum_{i \neq j} \sum_{k=1}^{n} g_{k,i} g_{k,j} \left( \prod_{b \neq i,j} P_{b}^{\alpha_{b}} \right) n_{i} n_{j} P_{i}^{n_{i}-1} P_{j}^{n_{j}-1} \right) p(x, t) dx dt \]
\[
\begin{align*}
&= \int \left( \sum_{i<j} \sum_{k=1}^{n} \sum_{m_i^1, m_j^2 \leq M} \sum_{0 \leq k_a \leq m_i^1 + m_j^2} g_{k,i;m_i^1,\ldots,m_i^1} g_{k,j;m_j^2,\ldots,m_j^2} \right) \\
&\prod_{a=1}^{n} \left( \frac{m_i^1 + m_j^2}{k_a} \right) \mathcal{P}_{a}^{k_a} \left( \mathcal{a} \cdot \frac{m_i^1 + m_j^2 - k_a}{k^2} \right) \left( \prod_{b \neq i,j}^{n} \mathcal{P}_{b}^{\alpha_{b}} \right) n_{i} n_{j} P_{i}^{n_{i}-1} P_{j}^{n_{j}-1} p(x,t)dxdt \\
&= \int \left( \sum_{i<j} \sum_{k=1}^{n} \sum_{m_i^1, m_j^2 \leq M} \sum_{0 \leq k_a \leq m_i^1 + m_j^2} n_{i} n_{j} g_{k,i;m_i^1,\ldots,m_i^1} g_{k,j;m_j^2,\ldots,m_j^2} \right) \\
&\prod_{a=1}^{n} \left( \frac{m_i^1 + m_j^2}{k_a} \right) \left( \mathcal{P}_{a}^{k_a} \right) \left( \mathcal{a} \cdot \frac{m_i^1 + m_j^2 - k_a}{k^2} \right) \left( \prod_{b \neq i,j}^{n} \mathcal{P}_{b}^{\alpha_{b}} \right) P_{i}^{n_{i}-1} P_{j}^{n_{j}-1} p(x,t)dxdt \\
&= \sum_{i<j} \sum_{k=1}^{n} \sum_{m_i^1, m_j^2 \leq M} \sum_{0 \leq k_a \leq m_i^1 + m_j^2} n_{i} n_{j} g_{k,i;m_i^1,\ldots,m_i^1} g_{k,j;m_j^2,\ldots,m_j^2} \\
&\prod_{a=1}^{n} \left( \frac{m_i^1 + m_j^2}{k_a} \right) \left( \mathcal{a} \cdot \frac{m_i^1 + m_j^2 - k_a}{k^2} \right), \\
&= \sum_{i=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{b \neq i,j}^{n} g_{k,i;m_i^1,\ldots,m_i^1} g_{k,j;m_j^2,\ldots,m_j^2} \\
&\prod_{a=1}^{n} \left( \frac{m_i^1 + m_j^2}{k_a} \right) \left( \mathcal{P}_{a}^{k_a} \right) \left( \mathcal{a} \cdot \frac{m_i^1 + m_j^2 - k_a}{k^2} \right) \left( \prod_{b \neq i,j}^{n} \mathcal{P}_{b}^{\alpha_{b}} \right) (n_{i} - 1) P_{i}^{n_{i}-2} p(x,t)dxdt
\end{align*}
\]

\( P_{\alpha_1+k_1,n_i-1+k_i-1,n_i+k_i-1,m_i^1+k_i+1,n_j-1+k_j-1,n_j+k_j+1,m_j^2+k_j+1,\ldots,\alpha_n+k_a} dt \) (3.4.30)
\[ \frac{1}{2} \int \left( \sum_{i=1}^{n} \sum_{k=1}^{m} \sum_{m_1, m_2 \leq M} n_i (n_i - 1) g_{k,i; m_i, \ldots, m_i, g_{k,i; m_i, \ldots, m_i}^2} \right) \]

\[ \cdot (\prod_{a=1}^{n} \left( m_a + m_a^2 \right) P_a^k (x_a) m_a^2 - k_a) \left( \prod_{b \neq i}^{n} P_b^{k_b} P_i^{n_i - 2} \right) p(x,t) dx dt \]

\[ = \frac{1}{2} \int \left( \sum_{i=1}^{n} \sum_{k=1}^{m} \sum_{m_1, m_2 \leq M} n_i (n_i - 1) g_{k,i; m_i, \ldots, m_i, g_{k,i; m_i, \ldots, m_i}^2} \right) \]

\[ \cdot (\prod_{a=1}^{n} \left( m_a + m_a^2 \right) (x_a) m_a^2 - k_a) \left( \prod_{b \neq i}^{n} P_b^{k_b} P_i^{n_i - 2} \right) p(x,t) dx dt \]

\[ = \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{m} \sum_{m_1, m_2 \leq M} \sum_{0 \leq k_a \leq m_a} n_i (\alpha_i - 1) g_{k,i; m_i, \ldots, m_i, g_{k,i; m_i, \ldots, m_i}} \]

\[ \cdot (\prod_{a=1}^{n} \left( m_a + m_a^2 \right) (x_a) m_a^2 - k_a) P_{\alpha + k_1, \ldots, n_i - 1, k_i - 1, n_i + k_i - 2, \ldots, \alpha_i + k_i, dt} \quad (3.4.31) \]

\[ D = (dy - \hat{h} dt) R^{-1} \int \left( \prod_{b=1}^{n} P_b^{k_b} \right) (h - \hat{h}) p(x,t) dx \]

\[ = (dy - \hat{h} dt) R^{-1} \int \left( \prod_{b=1}^{n} P_b^{k_b} \right) \]

\[ \left[ \sum_{m_1 \leq M_1} \sum_{0 \leq k_a \leq m_a} h_{i1, \ldots, m_a} \left( \prod_{a=1}^{n} \left( m_a \right) P_a^k (x_a) m_a^2 - k_a \right) \right. \]

\[ \left. - \sum_{m_1 \leq M_1} \sum_{0 \leq k_a \leq m_a} h_{i1, \ldots, m_a} \left( \prod_{a=1}^{n} \left( m_a \right) P_a^k (x_a) m_a^2 - k_a \right) P_{\alpha_1 + k_1, \ldots, \alpha_n + k_n} \right]_{n \times 1} p(x,t) dx \]

\[ = (dy - \hat{h} dt) R^{-1} \]

\[ \left[ \sum_{m_1 \leq M_1} \sum_{0 \leq k_a \leq m_a} h_{i1, \ldots, m_a} \left( \prod_{a=1}^{n} \left( m_a \right) (x_a) m_a^2 - k_a \right) P_{\alpha_1 + k_1, \ldots, \alpha_n + k_n} \right. \]

\[ \left. - \sum_{m_1 \leq M_1} \sum_{0 \leq k_a \leq m_a} h_{i1, \ldots, m_a} \left( \prod_{a=1}^{n} \left( m_a \right) (x_a) m_a^2 - k_a \right) P_{\alpha_1, \ldots, \alpha_n, P_{k_1, \ldots, k_n}} \right]_{n \times 1} \]
Substitute these equations into the expression of $dP_{\alpha_1,\ldots,\alpha_n}$, we have:

\[
dP_{\alpha_1,\ldots,\alpha_n} = -\sum_{a=1}^{n} \alpha_a \sum_{m_1 \leq M_f} \sum_{0 \leq k_b \leq m_a} f_{a,m_1,\ldots,m_n} \left( \prod_{b=1}^{n} \left( \hat{m}_b \right) (\hat{x}_b)^{m_b - k_b} \right).
\]

\[
P_{k_1,\ldots,k_n} \cdot P_{\alpha_1,\ldots,\alpha_n} - P_{k_1,\ldots,k_n} \cdot P_{\alpha_1,\ldots,\alpha_n} \ dt
\]

\[
- (dy - \hat{h} dt) \cdot \left( \sum_{a=1}^{n} \alpha_a \right) R^{-1}
\]

\[
\left[ \sum_{m_1 \leq M_f} \sum_{0 \leq k_b \leq m_a} h_{i;j,m_1,\ldots,m_n} \left( \prod_{b=1}^{n} \left( \hat{m}_b \right) (\hat{x}_b)^{m_b - k_b} \right) P_{k_1,\ldots,k_a-1,k_a+1,k_{a+1},\ldots,k_n} \right]_{n \times 1}.
\]

\[
P_{\alpha_1,\ldots,\alpha_n} - P_{\alpha_1,\ldots,\alpha_n} - 1, \ldots, \alpha_{n} - 1\]

\[
+ \frac{1}{2} \sum_{a=1}^{n} \alpha_a (\alpha_a - 1) \left( \sum_{i,j} r_{ij} \sum_{m_1 \leq M_f} \sum_{0 \leq k_b \leq m_a} h_{i;j,m_1,\ldots,m_n} \left( \prod_{b=1}^{n} \left( \hat{m}_b \right) (\hat{x}_b)^{m_b - k_b} \right) \right).
\]

\[
P_{k_1,\ldots,k_a-1,k_a+1,k_{a+1},\ldots,k_n} \cdot P_{\alpha_1,\ldots,\alpha_n} - 2, \alpha_{a+1}, \ldots, \alpha_n \ dt
\]

\[
+ \sum_{i=1}^{n} \sum_{m_1 \leq M_f} \sum_{0 \leq k_b \leq m_a} \alpha_i f_{i;j,m_1,\ldots,m_n} \left( \prod_{a=1}^{n} \left( \hat{m}_a \right) (\hat{x}_a)^{m_a - k_a} \right).
\]

\[
P_{k_1,\ldots,k_a-1,k_a+1,k_{a+1},\ldots,k_n} \cdot P_{\alpha_1,\ldots,\alpha_n} - n \ dt
\]

\[
+ \sum_{i=1}^{n} \sum_{m_1 \leq M_f} \sum_{0 \leq k_b \leq m_a} \alpha_i \alpha_j g_{i;j,m_1,\ldots,m_n} g_{k_1,\ldots,k_n} \ dt.
\]
\[
\prod_{a=1}^{n} \left( \frac{m_1^a + m_2^a}{k_a} \right) (\hat{x}_a)^{m_1^a + m_2^a - k_a}.
\]

\[P_{a_1 + k_1, \ldots, a_n + k_n} = \prod_{a=1}^{n} (\hat{x}_a)^{m_1^a + m_2^a - k_a} \cdot \alpha_i \left( 1 \right) g_{k_i, l_i}^{m_1^i, l_i^i} \cdot \sum_{a=1}^{n} m_a (\hat{x}_a)^{m_a - k_a} \cdot (dy - \hat{h} dt)^{R^{-1}} \sum_{m_l \leq M_l} \sum_{0 \leq k_a \leq m_a} [h_{1; m_1, \ldots, m_n} n \times 1 \prod_{a=1}^{n} (m_a) (\hat{x}_a)^{m_a - k_a}].
\]

\[P_{k_1, \ldots, k_a + 1, k_a + 1, \ldots, k_n}.
\]

\[\left( \sum_{m_l \leq M_l} \sum_{0 \leq k_a \leq m_a} h_{j; m_1, \ldots, m_n} \left( \prod_{c=1}^{n} \left( m_c \right) (\hat{x}_c)^{m_c - k_c} \right) P_{k_1, \ldots, k_a + 1, k_a + 1, \ldots, k_n} \right) \right) dt
\]

\[= \sum_{a=1}^{n} \left( \alpha_i \alpha_j P_{a_1, \ldots, a_{i-1}, a_i - 1, a_{i+1}, \ldots, a_{n-1}, a_n \ldots, a_n} \cdot \alpha_i \left( 1 \right) g_{k_i, l_i}^{m_1^i, l_i^i} \cdot \sum_{a=1}^{n} m_a (\hat{x}_a)^{m_a - k_a} \cdot (dy - \hat{h} dt)^{R^{-1}} \sum_{m_l \leq M_l} \sum_{0 \leq k_a \leq m_a} [h_{1; m_1, \ldots, m_n} n \times 1 \prod_{a=1}^{n} (m_a) (\hat{x}_a)^{m_a - k_a}].
\]

\[P_{k_1, \ldots, k_a + 1, k_a + 1, \ldots, k_n}.
\]
\[ + \frac{1}{2} \sum_{a=1}^{n} \alpha_a (\alpha_a - 1) \left( \sum_{i,j} r_{ij} \left( \sum_{m_i \leq M_i, 0 \leq k_i \leq m_i} h_{i;m_1,\ldots,m_n} \left( \prod_{b=1}^{n} \frac{m_b}{k_b} \right) (\hat{x}_b)^{m_a - k_a} \right) \right). \]

\[ P_{k_1,\ldots,k_{a-1},k_a+1,k_{a+1},\ldots,k_n} \]

\[ \left( \sum_{m_i \leq M_i, 0 \leq k_i \leq m_i} h_{j;m_1,\ldots,m_n} \left( \prod_{b=1}^{n} \frac{m_b}{k_b} \right) (\hat{x}_b)^{m_a - k_a} \right). \]

\[ P_{k_1,\ldots,k_{a-1},k_a+1,k_{a+1},\ldots,k_n} \sum_{i=1}^{n} \sum_{m_i \leq M_i, 0 \leq k_i \leq m_i} \alpha_if_{i;m_1,\ldots,m_n} \left( \prod_{a=1}^{n} \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a - k_a}. \]

\[ P_{a_1+k_1,\ldots,a_i-1+k_i-1,a_i+k_i-1,a_i+1+k_{i+1},\ldots,a_n+k_n} \]

\[ + \sum_{i<j} \sum_{k_1 \leq k_2 \leq M_i, 0 \leq k_i \leq 1} \sum_{m_i \leq M_i} \alpha_i \alpha_j g_{i;j;m_1,\ldots,m_n} \left( \prod_{a=1}^{n} \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a + m_a^2 - k_a}. \]

\[ \left( \prod_{a=1}^{n} \left( \frac{m_a + m_a^2}{k_a} \right) \right) (\hat{x}_a)^{m_a + m_a^2 - k_a}. \]

\[ P_{a_1+k_1,\ldots,a_i-1+k_i-1,a_i+k_i-1,a_i+1+k_{i+1},\ldots,a_n+k_n} \]

\[ + \frac{1}{2} \sum_{i=1}^{n} \sum_{k_1 \leq k_2 \leq M_i, 0 \leq k_i \leq 1} \sum_{m_i \leq M_i} \alpha_i (\alpha_1 - 1) g_{i;j;m_1,\ldots,m_n} \left( \prod_{a=1}^{n} \frac{m_a}{k_a} \right) (\hat{x}_a)^{m_a + m_a^2 - k_a}. \]

\[ \left( \prod_{a=1}^{n} \left( \frac{m_a + m_a^2}{k_a} \right) \right) (\hat{x}_a)^{m_a + m_a^2 - k_a}. \]

\[ + \sum_{a<b} \left( \alpha_a \alpha_b P_{a_1,\ldots,a_i-1,a_i-1,a_i+1,\ldots,a_n-1,a_{b-1},a_{b+1},\ldots,a_n} \right). \]

\[ \left( \sum_{i,j} r_{ij} \left( \sum_{m_i \leq M_i, 0 \leq k_i \leq c_i} h_{i;m_1,\ldots,m_n} \left( \prod_{c=1}^{c_i} \frac{m_c}{k_c} \right) (\hat{x}_c)^{m_c - k_c} \right) P_{k_1,\ldots,k_{a-1},k_a+1,k_{a+1},\ldots,k_n} \right). \]

\[ \left( \sum_{m_i \leq M_i, 0 \leq k_i \leq c_i} h_{i;m_1,\ldots,m_n} \left( \prod_{c=1}^{c_i} \frac{m_c}{k_c} \right) (\hat{x}_c)^{m_c - k_c} \right) P_{k_1,\ldots,k_{a-1},k_a+1,k_{a+1},\ldots,k_n} \]

\[ - \sum_{a=1}^{n} \left( \alpha_a \sum_{i,j} r_{ij} \left( \sum_{m_i \leq M_i, 0 \leq k_i \leq m_i} h_{i;m_1,\ldots,m_n} \left( \prod_{b=1}^{n} \frac{m_b}{k_b} \right) (\hat{x}_b)^{m_a - k_a} \right) \right). \]

\[ P_{k_1,\ldots,k_{a-1},k_a+1,k_{a+1},\ldots,k_n}. \]
\[ \left( \sum_{m_j \leq M_j} \sum_{0 \leq k_b \leq m_b} h_{j;m_1,...,m_n} \left( \prod_{b=1}^{n} \binom{m_b}{k_b} (\widehat{x}_b)^{m_b-k_b} \right) \right) dt \\
- (dy - \widehat{h} dt)^{-1} \left( \sum_{a=1}^{n} \alpha_a \right) \left( \sum_{m_1 \leq M_1} \sum_{0 \leq k_a \leq m_a} h_{i;m_1,...,m_n} \left( \prod_{b=1}^{n} \binom{m_b}{k_b} (\widehat{x}_b)^{m_b-k_b} \right) P_{k_1,...,k_{a-1},k_{a+1},...,k_n} \right) \]

\[ \left( P_{a_1+k_1,...,a_n+k_n} - P_{a_1,...,a_n} P_{k_1,...,k_n} \right) \] (3.4.33)

Now we consider the last term of the equation about \( P_{N_1,...,N_n} \). From the above equation, it is given as following

\[ \left( \sum_{m_j \leq M_j} \sum_{0 \leq k_b \leq m_b} h_{j;m_1,...,m_n} \left( \prod_{b=1}^{n} \binom{m_b}{k_b} (\widehat{x}_b)^{m_b-k_b} \right) \right) dt \\
- (dy - \widehat{h} dt)^{-1} \left( \sum_{a=1}^{n} \alpha_a \right) \left( \sum_{m_1 \leq M_1} \sum_{0 \leq k_a \leq m_a} h_{i;m_1,...,m_n} \left( \prod_{b=1}^{n} \binom{m_b}{k_b} (\widehat{x}_b)^{m_b-k_b} \right) P_{k_1,...,k_{a-1},k_{a+1},...,k_n} \right) \]

\[ \left( P_{N_1+k_1,...,N_n+k_n} - P_{N_1,...,N_n} P_{k_1,...,k_n} \right) \]

\[ = - (dy - \widehat{h} dt)^{-1} \left( \sum_{m_1 \leq M_1} \sum_{0 \leq k_a \leq m_a} h_{i;m_1,...,m_n} \left( \prod_{b=1}^{n} \binom{m_b}{k_b} (\widehat{x}_b)^{m_b-k_b} \right) \sum_{a=1}^{n} \alpha_a \right) \left( \sum_{m_1 \leq M_1} \sum_{0 \leq k_b \leq m_b} h_{j;m_1,...,m_n} \left( \prod_{b=1}^{n} \binom{m_b}{k_b} (\widehat{x}_b)^{m_b-k_b} \right) \right) \]

\[ \left( P_{N_1,...,N_n} - P_{N_1+k_1,...,N_n+k_n} + P_{N_1,...,N_n} P_{k_1,...,k_n} \right) \] (3.4.34)
If we assume

\[
P_{N_1+k_1,\ldots,N_n+k_n} = \sum_{a=1}^{n} N_a P_{k_1,a-1,k_a+1,k_{a+1},\ldots,k_n} P_{N_1,a-1,N_a-1,N_{a+1},\ldots,N_n} + P_{N_1,\ldots,N_n P_{k_1,\ldots,k_n}}
\]

(3.4.35)

for \(0 \leq k_i \leq M_h, 1 \leq i \leq n\), then the term reduces to zero, and

for \(k_i = 0, 1 \leq i \leq n\), (3.4.35) is always true since \(P_{k_1,\ldots,k_n} = 0, k_i = 1, k_j = 0, j \neq i\), and \(P_{0,\ldots,0} = 1\)

for \(k_i = 1, k_j = 0, j \neq i\), \(P_{N_1+k_1,\ldots,N_n+k_n} = \sum_{a=1}^{n} N_a P_{k_1,a-1,k_a+1,k_{a+1},\ldots,k_n}\).

\(P_{N_1,\ldots,N_{a-1},N_a-1,N_{a+1},\ldots,N_n}\), since \(P_{k_1,\ldots,k_n} = 0, k_i = 1, k_j = 0, j \neq i\);

If we keep going, we can express all higher order \(P_{\alpha_1,\ldots,\alpha_n}, \alpha_i > N_i, 1 \leq i \leq n\) in terms of lower order \(P_{\alpha_1,\ldots,\alpha_n}\). However, notice the terms like \(P_{\alpha_1,\ldots,\alpha_n}, \alpha_i \leq N_i, \alpha_j > N_j, j \neq i\), we could not express them in terms of low order \(P_{\alpha_1,\ldots,\alpha_n}\). So we need to make the following assumption.

**Assumption** \(\alpha_i \leq N_i\)

\[
P_{\alpha_1+k_1,\ldots,\alpha_n+k_n} = \sum_{a=1}^{n} \alpha_a P_{k_1,a-1,k_a+1,k_{a+1},\ldots,k_n} P_{\alpha_1,\ldots,\alpha_{a-1},\alpha_{a+1},\ldots,\alpha_n} + P_{\alpha_1,\ldots,\alpha_n P_{k_1,\ldots,k_n}}
\]

(3.4.36)

holds for every condition below, separately:

(1) \(\alpha_i = N_i, 1 \leq i \leq n\) ((3.4.36) becomes (3.4.35)),

(2) \(\alpha_{ij} = N_i, k_{ij} = 0, \{i_1, \ldots, i_{n-1}\}\) is a subset of \(\{1, \ldots, n\}\),
(3) \( \alpha_i = N_{i,j}, k_{i,j} = 0, \{i_1, \ldots, i_{n-2}\} \) is a subset of \( \{1,\ldots,n\} \),

\[
\ldots
\]

(n) \( \alpha_{i_1} = N_{i_1}, k_{i_1} = 0, \{i_1\} \) is a subset of \( \{1,\ldots,n\} \).

With this assumption, we are able to express the terms \( P_{\alpha_1,\ldots,\alpha_n} \)'s, \( \alpha_i \leq N_i, 1 \leq i \leq N \), in terms of lower order \( P_{\alpha_1,\ldots,\alpha_n} \)'s.

Summing up the above discussions, we have the following theorem.

**Theorem 3.4.4.** For the continuous filtering problem given by the system (3.2.1) with \( f_i(x,t) \)'s, \( g_{i,j}(x,t) \)'s, \( h_i(x,t) \)'s being polynomial functions of \( x \) defined by (3.4.1)-(3.4.3), assuming that assumptions (3.4.36) holds, we can express all higher order \( P_{\alpha_1,\ldots,\alpha_n} \)'s (any \( \alpha_i > N_i \)), in terms of \( P_{\alpha_1,\ldots,\alpha_n} \)'s (every \( \alpha_i \leq N_i \)). Thus we can establish a closed system of equations about \( \hat{x}_i \)'s, \( P_{\alpha_1,\ldots,\alpha_n} \)'s, \( \alpha_i \leq N_i \). More Specifically, we solve the closed system of the equations: (3.4.15) for conditional mean \( \hat{x}_i \)'s, (3.4.33) for \( P_{\alpha_1,\ldots,\alpha_n} \)'s (every \( \alpha_i \leq N_i \), but some \( \alpha_i \neq N_i \)), and the equations for \( P_{N_1,\ldots,N_n} \) given as following

\[
\frac{dP_{N_1,\ldots,N_n}}{dt} = \left( -\sum_{a=1}^{n} N_a \sum_{m_1 \leq M_f}^{n} \sum_{0 \leq k_b \leq m_b} f_{a;m_1,\ldots,m_n} \left( \prod_{b=1}^{n} \left( \frac{m_b}{k_b} \right) (\hat{x}_b)^{m_b-k_b} \right) \right)
\]

\[
P_{k_1,\ldots,k_n} P_{N_1,\ldots,N_{a-1},N_a-1,N_{a+1},\ldots,N_n}
\]

\[
+ \frac{1}{2} \sum_{a=1}^{n} N_a (N_a - 1) \left( \sum_{i,j} r^{ij} \left( \sum_{m_1 \leq M_f}^{n} \sum_{0 \leq k_b \leq m_b} h_{i;j;m_1,\ldots,m_n} \left( \prod_{b=1}^{n} \left( \frac{m_b}{k_b} \right) (\hat{x}_b)^{m_b-k_b} \right) \right) \right)
\]

\[
P_{k_1,\ldots,k_{a-1},k_a+1,k_{a+1},\ldots,k_n}
\]

\[
\left( \sum_{m_1 \leq M_f}^{n} \sum_{0 \leq k_b \leq m_b} h_{j;m_1,\ldots,m_n} \left( \prod_{b=1}^{n} \left( \frac{m_b}{k_b} \right) (\hat{x}_b)^{m_b-k_b} \right) \right).
\]
\[
P_{k_1, \ldots, k_{a-1}, k_a+1, k_{a+1}, \ldots, k_n} P_{N_1, \ldots, N_{a-1}, N_a-2, N_{a+1}, \ldots, N_n} + \sum_{i=1}^{n} \sum_{m_i \leq M_f} \sum_{0 \leq k_a \leq m_a} \sum_{m_i \leq M_f} \sum_{0 \leq k_a \leq m_a} N_i f_{i; m_1, \ldots, m_n} \left( \prod_{a=1}^{n} \left( \frac{m_a}{k_a} \right) \left( \frac{m_a - k_a}{m_a - k_a} \right) \right) \]

\[
P_{N_1+k_1, \ldots, N_{a-1}+k_{a-1}, N_a+k_a-1, N_{a+1}+k_{a+1}, \ldots, N_n+k_n} + \sum_{i=1}^{n} \sum_{m_i \leq M_f} \sum_{0 \leq k_a \leq m_a} \sum_{m_i \leq M_f} \sum_{0 \leq k_a \leq m_a} N_i N_j g_{k_i; m_1, \ldots, m_n} g_{k_j; m_1, \ldots, m_n} \cdot \sum_{a=1}^{n} \left( \frac{m_a^1 + m_a^2}{k_a} \right) \left( \frac{m_a^1 + m_a^2 - k_a}{m_a^1 + m_a^2 - k_a} \right) \]

\[
P_{N_1+k_1, \ldots, N_{a-1}+k_{a-1}, N_a+k_a-1, N_{a+1}+k_{a+1}, \ldots, N_n+k_n} + \frac{1}{2} \sum_{i=1}^{n} \sum_{m_i \leq M_f} \sum_{0 \leq k_a \leq m_a} \sum_{m_i \leq M_f} \sum_{0 \leq k_a \leq m_a} N_i (N_i - 1) g_{k_i; m_1, \ldots, m_n} g_{k_i; m_1, \ldots, m_n} \cdot \sum_{a=1}^{n} \left( \frac{m_a^1 + m_a^2}{k_a} \right) \left( \frac{m_a^1 + m_a^2 - k_a}{m_a^1 + m_a^2 - k_a} \right) \]

\[
\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{m_i \leq M_f} \sum_{0 \leq k_c \leq m_i} h_{i; m_1, \ldots, m_n} \left( \prod_{c=1}^{n} \left( \frac{m_c}{k_c} \right) \left( \frac{m_c - k_c}{m_c - k_c} \right) \right) P_{k_1, \ldots, k_{a-1}, k_a+1, k_{a+1}, \ldots, k_n} \]

\[
\left( \sum_{m_j \leq M_f} \sum_{0 \leq k_c \leq m_j} h_{j; m_1, \ldots, m_n} \left( \prod_{c=1}^{n} \left( \frac{m_c}{k_c} \right) \left( \frac{m_c - k_c}{m_c - k_c} \right) \right) P_{k_1, \ldots, k_{a-1}, k_a+1, k_{a+1}, \ldots, k_n} \right) \]

\[
\sum_{a=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{m_i \leq M_f} \sum_{0 \leq k_b \leq m_i} h_{i; m_1, \ldots, m_n} \left( \prod_{b=1}^{n} \left( \frac{m_b}{k_b} \right) \left( \frac{m_b - k_b}{m_b - k_b} \right) \right) \]

\[
\left( \sum_{m_j \leq M_f} \sum_{0 \leq k_b \leq m_j} h_{j; m_1, \ldots, m_n} \left( \prod_{b=1}^{n} \left( \frac{m_b}{k_b} \right) \left( \frac{m_b - k_b}{m_b - k_b} \right) \right) \right) P_{N_1+k_1, \ldots, N_{a-1}+k_{a-1}, N_a+k_a-1, N_{a+1}+k_{a+1}, \ldots, N_n+k_n} - P_{N_1, \ldots, N_{a-1}, N_a-1, N_{a+1}, \ldots, N_n, P_{k_1, \ldots, k_n}} dt \tag{3.4.37} \]
3.5 Examples, numerical results and discussions

In this section, we shall give some examples to describe how to construct the suboptimal filter by using the method we introduced in the section above. We will also give our numerical results obtained by using the new suboptimal filter (NSF). We then compare our results with those obtained by using EKF, UKF, EnKF, and PF (SIR) for nonlinear filtering examples. The simulation procedures are given as following. First, based on the model equation (3.2.1), we generate the sample path of the state and observation by using the method introduced by [48], and we use simple rectangle rule integration with constant step size for the simulation for ODEs and SDEs. We descretized the filter model by Euler scheme when we use the discrete nonlinear filter methods. We take the time interval $[0, 10]$, time step $\Delta t = 0.001$, which means we take 10000 subintervals for the simulation. For the same state, we took an average of 20 runs for our new suboptimal filter, EKF, UKF, EnKF, and PF. We use 20 ensembles for EnKF, and 50 particles for PF.

A. Example 1: Linear Filtering Problems

In (3.2.1), if we choose $n = 1$, $f(x,t) = F(t)x$, $g(x,t) = G(t)$, and $h(x,t) = H(t)x$, $x_0$ follows a Gaussian distribution, then our new suboptimal filter reduces to Kalman Filter [4], i.e.

$$
\begin{align*}
    d\hat{x} &= F\hat{x} + \frac{1}{R}HP_2(dy - H\hat{x}dt) \\
    \frac{dP_2}{dt} &= 2FP_2 + G^2 - \frac{1}{R}(HP_2)^2 
\end{align*}
$$

We only need to assume that $P_3 = P_2P_1 = 0$, which is right since for linear cases, the density function is Gaussian.
B. Example 2: Cubic Sensor Problem

In (3.2.1), if we choose \( f(x,t) = 0 \), \( g(x,t) = 1 \), and \( h(x,t) = x^3 \), we get the cubic sensor problem. The cubic sensor problem is essential infinite-dimensional as proved by [1], [49], [50], and was studied by many authors, for example, [51], [52], [53]. From (3.3.34), we need to solve the following equations of conditional mean and variance for EKF,

\[
\begin{aligned}
\frac{d\hat{x}}{dt} &= \frac{3}{R} P_2 (\hat{x})^2 (dy - (\hat{x})^3 dt) \\
\frac{dP_2}{dt} &= 1 - \frac{9}{R} P_2 (\hat{x})^4
\end{aligned}
\]

(3.5.2)

Using the same assumption from example 1, we have our new suboptimal filter for conditional mean and variance is obtained as follow

\[
\begin{aligned}
\frac{d\hat{x}}{dt} &= \frac{1}{R} (P_2^2 + 3P_2 (\hat{x})^2) (dy - (3P_2 \hat{x} + (\hat{x})^3) dt) \\
\frac{dP_2}{dt} &= 1 - \frac{1}{R} (P_2^2 + 3P_2 (\hat{x})^2)^2
\end{aligned}
\]

(3.5.3)

We start with Figure 3.5.1, which shows that the true state and estimates provided by our new suboptimal filter and EKF. Figure 3.5.1-Figure 3.5.4 show the numerical results with \( R = 1 \), and initial state \( x_0 \sim N(2, 0.01) \). In this case, although our new suboptimal filter, EKF and UKF agree with the true state very well throughout the time period, UKF takes a much longer time than NSF and EKF (UKF takes 87.1 seconds while NSF takes 8.5 seconds and EKF takes 8.3 seconds separately). EnKF and PF fail to work as well as the other filters and also take too long time (EnKF takes 393.8 seconds and PF takes 1211.0 seconds). The comparisons of the five filters for the case with \( R = 0.1 \) and initial state \( x_0 \sim N(0, 0.001) \) are
given in Figure 3.5.5-Figure 3.5.8. As shown in this figure, our new suboptimal filter still works very well, but EKF completely fails, i.e. the estimate gets stuck at zero. UKF, EnKF and PF do not track the state very well and they take longer time (NSF, EKF, UKF, EnKF and PF take 8.5, 8.3, 90.0, 412.0, and 1224.1 seconds, separately).

Figure 3.5.1. NSF compared with EKF for cubic sensor problem, with $R = 1$ and initial condition $x_0 \sim N(2, 0.01)$. 
Figure 3.5.2. NSF compared with UKF for cubic sensor problem, with $R = 1$ and initial condition $x_0 \sim N(2, 0.01)$. 
Figure 3.5.3. NSF compared with EnKF for cubic sensor problem, with $R = 1$ and initial condition $x_0 \sim N(2, 0.01)$, 20 ensemble for EnKF.
Figure 3.5.4. NSF compared with PF for cubic sensor problem, with $R = 1$ and initial condition $x_0 \sim N(2, 0.01)$, 50 particles for PF.
Figure 3.5.5. NSF compared with EKF for cubic sensor problem, with $R = 0.1$ and initial condition $x_0 \sim N(0, 0.001)$. 
Figure 3.5.6. NSF compared with UKF for cubic sensor problem, with $R = 0.1$ and initial condition $x_0 \sim N(0, 0.001)$. 
Figure 3.5.7. NSF compared with EnKF for cubic sensor problem, with $R = 0.1$ and initial condition $x_0 \sim N(0, 0.001)$, 20 ensembles for EnKF.
Figure 3.5.8. NSF compared with PF for cubic sensor problem, with $R = 0.1$ and initial condition $x_0 \sim N(0, 0.001)$, 50 particles for PF.
C. Example 3

We choose

\[
\begin{align*}
  f_1 &= 0, & h_1 &= x_1 x_2, & g &= \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}, & R &= I_2, \\
  f_2 &= x_1^2, & h_2 &= x_2^2
\end{align*}
\]  

(3.5.4)

The initial state

\[
\begin{pmatrix} x_{10} \\ x_{20} \end{pmatrix} \sim N\left( \begin{pmatrix} 1.1 \\ 1.1 \end{pmatrix}, \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix} \right)
\]  

(3.5.5)

Then EKF reduces to solve

\[
\begin{align*}
  d\hat{x}_1 &= (dy_1 - \hat{x}_1 \hat{x}_2 dt)(P_{11} \hat{x}_1 + P_{20} \hat{x}_2) + (dy_2 - \hat{x}_2 dt)(2P_{11} \hat{x}_2) \\
  d\hat{x}_2 &= \hat{x}_1^2 dt + (dy_1 - \hat{x}_1 \hat{x}_2 dt)(P_{02} \hat{x}_1 + P_{11} \hat{x}_2) + (dy_2 - \hat{x}_2 dt)(2P_{02} \hat{x}_2) \\
  \frac{dP_{20}}{dt} &= -(P_{11} \hat{x}_1 + P_{20} \hat{x}_2)^2 + (2P_{11} \hat{x}_2)^2 + q_{11} \\
  \frac{dP_{11}}{dt} &= -(P_{11} \hat{x}_1 + P_{20} \hat{x}_2)(P_{02} \hat{x}_1 + P_{11} \hat{x}_2) + (2P_{11} \hat{x}_2)(2P_{02} \hat{x}_2) + 2P_{20} \hat{x}_1 \\
  \frac{dP_{02}}{dt} &= -(P_{02} \hat{x}_1 + P_{11} \hat{x}_2)^2 + (2P_{02} \hat{x}_2)^2 + q_{22}
\end{align*}
\]  

(3.5.6)

Using our new suboptimal filter, the equations of conditional mean and higher moments reduce to

\[
\begin{align*}
  d\hat{x}_1 &= (dy_1 - (P_{11} + \hat{x}_1 \hat{x}_2)dt)(P_{21} + P_{11} \hat{x}_1 + P_{20} \hat{x}_2) \\
  &\quad + (dy_2 - (P_{02} + \hat{x}_2 dt)(P_{12} + 2P_{11} \hat{x}_2) \\
  d\hat{x}_2 &= (P_{20} + \hat{x}_1^2 dt + (dy_1 - (P_{11} + \hat{x}_1 \hat{x}_2)dt)(P_{12} + P_{02} \hat{x}_1 + P_{11} \hat{x}_2)
\end{align*}
\]
\[ \begin{align*}
+ (dy_2 - (P_{02} + \hat{x}_2^2)dt)(2\hat{x}_2P_{02}) \\
\] 
\[ dP_{22} = (P_{02}((P_{21} + P_{11}\hat{x}_1 + P_{20}\hat{x}_2)^2 + (P_{12} + 2P_{11}\hat{x}_2)^2) + P_{20}((P_{12} + P_{02}\hat{x}_1 + P_{11}\hat{x}_2)^2
+ (2P_{02}\hat{x}_2)^2) + 4P_{11}((P_{21} + P_{11}\hat{x}_1 + P_{20}\hat{x}_2)(P_{12} + P_{02}\hat{x}_1 + P_{11}\hat{x}_2) \\
+ (P_{12} + 2P_{11}\hat{x}_2)(2P_{02}\hat{x}_2)) + P_{02}q_{11} + P_{20}q_{22} + 2(-P_{21}P_{20} + 2P_{20}P_{21} + 6P_{20}P_{11}\hat{x}_1)
- 2((P_{21} + P_{11}\hat{x}_1 + P_{20}\hat{x}_2)(2P_{12}P_{11} + 2P_{21}P_{02} + 3P_{11}P_{02}\hat{x}_1 + P_{22}\hat{x}_2 - P_{11}P_{21})
+ (P_{12} + 2P_{11}\hat{x}_2)(P_{12}P_{02} + 6P_{11}P_{02}\hat{x}_2 - P_{02}P_{12}))
- 2((P_{12} + P_{02}\hat{x}_1 + P_{11}\hat{x}_2)(2P_{12}P_{20} + 2P_{21}P_{11} + P_{22}\hat{x}_1 + 3P_{20}P_{11}\hat{x}_2 - P_{11}P_{21})
+ (2P_{02}\hat{x}_2)(2P_{12}P_{11} + 2P_{21}P_{02} + 2P_{22}\hat{x}_2 - P_{02}P_{21}))dt \\
\] 
\[ dP_{21} = ((-P_{20}P_{20} + P_{20}^2) - 2((P_{21} + P_{11}\hat{x}_1 + P_{20}\hat{x}_2)(P_{22} + P_{12}\hat{x}_1 + P_{21}\hat{x}_2 - P_{11}P_{11})
+ (P_{12} + 2P_{11}\hat{x}_2)(3P_{11}P_{02} + 2P_{12}\hat{x}_2 - P_{02}P_{11}))
- ((P_{12} + P_{02}\hat{x}_1 + P_{11}\hat{x}_2)(3P_{20}P_{11} + P_{21}\hat{x}_1 - P_{20}P_{11})
+ (2P_{02}\hat{x}_2)(P_{22} + 2P_{21}\hat{x}_2 - P_{02}P_{20}))dt
- (dy_1 - (P_{11} + \hat{x}_1\hat{x}_2)dt)(2P_{11}(P_{21} + P_{11}\hat{x}_1 + P_{20}\hat{x}_2)
- (2P_{12}P_{20} + 2P_{21}P_{11} + P_{22}\hat{x}_1 + 3P_{20}P_{11}\hat{x}_2 - P_{11}P_{21})
+ P_{20}(P_{12} + P_{02}\hat{x}_1 + P_{11}\hat{x}_2)) - (dy_2 - (P_{02} + \hat{x}_2^2)dt)(2P_{11}(P_{12} + 2P_{11}\hat{x}_2)
- (2P_{12}P_{11} + 2P_{21}P_{02} + 2P_{22}\hat{x}_2 - P_{02}P_{21}) + P_{20}(2P_{02}\hat{x}_2)) \\
\] 
\[ dP_{12} = (2(-P_{11}P_{20} + 3P_{20}P_{11} + 2P_{21}\hat{x}_1) -
(P_{21} + P_{11}\hat{x}_1 + P_{20}\hat{x}_2)(3P_{11}P_{02} + P_{12}\hat{x}_2 - P_{11}P_{02})
+ (P_{12} + 2P_{11}\hat{x}_2)(P_{02}^2 - P_{02}P_{02})) \
\]
Here we assume $P_{32} = 2P_{12}P_{20} + 2P_{21}P_{11}$, $P_{23} = 2P_{12}P_{11} + 2P_{21}P_{02}$, $P_{31} = 3P_{20}P_{11}$, $P_{13} = 3P_{11}P_{02}$, $P_{04} = P_{02}^2$, $P_{40} = P_{20}^2$, $P_{14} = 2P_{12}P_{02}$, $P_{41} = 2P_{20}P_{21}$, $P_{30} = P_{03} = 0$, (we choose

\begin{align*}
- 2((P_{12} + P_{02}\hat{x}_1 + P_{11}\hat{x}_2)(P_{22} + P_{12}\hat{x}_1 + P_{21}\hat{x}_2 - P_{11}P_{11}) \\
+ (2P_{02}\hat{x}_2)(3P_{11}P_{02} + 2P_{12}\hat{x}_2 - P_{02}P_{11}))dt \\
- (dy_1 - (P_{11} + \hat{x}_1\hat{x}_2)dt)(P_{02}(P_{21} + P_{11}\hat{x}_1 + P_{20}\hat{x}_2) \\
- (2P_{12}P_{11} + 2P_{21}P_{02} + 3P_{11}P_{02}\hat{x}_1 + P_{22}\hat{x}_2 - P_{11}P_{11}) \\
+ 2P_{11}(P_{12} + P_{02}\hat{x}_1 + P_{11}\hat{x}_2)) - (dy_2 - (P_{02} + \hat{x}_2^2)dt)(P_{02}(P_{12} + 2P_{11}\hat{x}_2) \\
- (2P_{12}P_{02} + 6P_{11}P_{02}\hat{x}_2 - P_{02}P_{12}) + 2P_{11}(2P_{02}\hat{x}_2)) \\
dP_{20} = -((P_{21} + P_{11}\hat{x}_1 + P_{20}\hat{x}_2)^2 + (P_{12} + 2P_{11}\hat{x}_2)^2) + q_{11})dt \\
- (dy_1 - (P_{11} + \hat{x}_1\hat{x}_2)dt)(-3P_{20}P_{11} + P_{21}\hat{x}_1 - P_{11}P_{20})) \\
- (dy_2 - (P_{02} + \hat{x}_2^2)dt)(-P_{22} + 2P_{21}\hat{x}_2 - P_{02}P_{20})) \\
dP_{11} = ((2P_{20}\hat{x}_1) - ((P_{21} + P_{11}\hat{x}_1 + P_{20}\hat{x}_2)(P_{12} + P_{02}\hat{x}_1 + P_{11}\hat{x}_2) \\
+ (P_{12} + 2P_{11}\hat{x}_2)(2P_{02}\hat{x}_2))dt \\
- (dy_1 - (P_{11} + \hat{x}_1\hat{x}_2)dt)(-P_{22} + P_{12}\hat{x}_1 + P_{21}\hat{x}_2 - P_{11}P_{11})) \\
- (dy_2 - (P_{02} + \hat{x}_2^2)dt)(-3P_{11}P_{02} + 2P_{12}\hat{x}_2 - P_{02}P_{11})) \\
dP_{02} = (2(P_{21} + 2P_{11}\hat{x}_1) - ((P_{12} + P_{02}\hat{x}_1 + P_{11}\hat{x}_2)^2 + (2P_{02}\hat{x}_2)^2) + q_{22})dt \\
- (dy_1 - (P_{11} + \hat{x}_1\hat{x}_2)dt)(-3P_{11}P_{02} + P_{12}\hat{x}_2 - P_{11}P_{02}) \\
- (dy_2 - (P_{02} + \hat{x}_2^2)dt)(-P_{02}^2 - P_{02}P_{02}))
\end{align*}
$N_1 = N_2 = 2)$. Numerical results for this example are given in Figure 3.5.9-Figure 3.5.12. One can see that our suboptimal filter track the state better than EKF. Although UKF works well comparing to NSF, NSF is much faster than UKF (NSF takes 15.4 seconds while UKF takes 163.4 seconds).

![Figure 3.5.9. NSF compared with EKF for example 3.](image-url)
Figure 3.5.10. NSF compared with UKF for example 3.
Figure 3.5.11. NSF compared with EnKF for example 3, 20 ensembles for EnKF.
Figure 3.5.12. NSF compared with PF for example 3, 50 particles for PF.
D. Discussion

From example 2, we know that EKF sometime completely fails. We first observe that those situations happen if the mean of initial condition is zero. A similar observation was also reported in [53]. As [53] pointed out, EKF gets stuck at zero if the estimator reaches this value. The possible reason is EKF is a linear approximation of the optimal filter and it neglects more terms than our new suboptimal filter.

3.6 CONCLUSIONS

We have discussed the most popular methods for nonlinear filtering problem. All of them have their own advantages and drawbacks. Due to the importance in the industry, we need to find better optimal or suboptimal nonlinear filter to solve the problem. Starting from the equations for generalized higher moments $P_{\alpha_1,\ldots,\alpha_n}$, we successfully constructed a new suboptimal filter for polynomial filter problems. The resulted new suboptimal filter consists of solving a closed system of equations about $P_{\alpha_1,\ldots,\alpha_n}$'s, $\alpha_i \leq N_i$. Numerical results show that our new suboptimal filter works perfectly for polynomial filtering problems.
CITED LITERATURE


VITA

Contact

Yang Jiao

322 Science and Engineering Offices (M/C 249)

851 S. Morgan Street, Chicago, IL 60607

jiaoyang0715@gmail.com

Education

• PH.D in Applied Mathematics, University of Illinois at Chicago. 2012

• M.S. in Computational Finance, University of Illinois at Chicago. 2008

• B.S. in Applied Mathematics, University of Science and Technology of China, China. 2006

Work Experience

• Assistant of Manage Editor of Journal of Algebraic Geometry

  Chicago, IL, USA, 2010.12-2012.8

• Research Assistant at University of Illinois at Chicago

  Chicago, IL, USA, 2010.5-2012.8

• Teaching Assistant at University of Illinois at Chicago

  Chicago, IL, USA, 2006.8-2010.5
Publications

