Article

# A Gradually Linearizing Kalman Filter Bank Designing for Product-Type Strong Nonlinear Systems 

Chenglin Wen and Zhipeng Lin *

School of Affiliation, Hangzhou Dianzi University, Hangzhou 310018, China; wencl@hdu.edu.cn

* Correspondence: lin_zp1996@163.com

Citation: Wen, C.; Lin, Z. A Gradually Linearizing Kalman Filter Bank Designing for Product-Type

Strong Nonlinear Systems. Electronics 2022, 11, 714. https://doi.org/ 10.3390/electronics 11050714

Academic Editor: Teresa OrlowskaKowalska

Received: 12 January 2022
Accepted: 21 February 2022
Published: 25 February 2022
Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.


Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).


#### Abstract

Our study aimed to improve the poor performance of existing filters, such as EKF, UKF and CKF, that results from their weak approximation ability to nonlinear systems. This paper proposes a new extended Kalman filter bank focusing on a class of product-type strong nonlinear systems composed by system state variables, time-varying parameters and non-linear basic functions. Firstly, the non-linear basic functions are defined as hidden variables corresponding to system state variables, and then the strong nonlinear systems are described simplistically. Secondly, we discuss building two dynamic models between their future values of parameters, as well as hidden variables and their current values based on the given prior information. Thirdly, we recount how an extended Kalman filter bank was designed by gradually linearizing the strong nonlinear systems about system state variables, time-varying parameters and hidden variables, respectively. The first extended Kalman filter about future hidden variables was designed by using these estimates of the state variables and parameters, as well as hidden variables at current. The second extended Kalman filter about future parameters variables was designed by using these estimates of the current state variables and parameters, as well as future hidden variables. The third extended Kalman filter about future state variables was designed by using these estimates of the current state variables, as well as future parameters and hidden variables. Fourthly, we used digital simulation experiments to verify the effectiveness of this method.


Keywords: product type; time-varying parameters; hidden variables; strong nonlinear systems; dynamic models; gradually linearizing

## 1. Introduction

With the development of the times, filtering theory has played an important role in various fields at the domestic and international level, especially in national defense, military and other fields, such as tracking navigation, signal processing, automatic control, target tracking, etc. [1-6]. Kalman (R.E. Kalman) proposed a dynamic recursive state estimation method called the Kalman filter (KF) in 1960 [7]. The KF takes into account the statistical characteristics of the estimated and observed measures, and it designs an optimal filter based on the minimum mean square error to solve the problem of state estimation and target tracking in linear Gaussian systems [8,9]. However, in practical applications, KF is difficult to use to solve nonlinear models [10,11]. Bucy proposed the extended Kalman filter (EKF) method to solve nonlinear problems [12]. EKF is a function linear approximation method for nonlinear system models. It retains the Taylor expansion of the nonlinear function to the first-order term, thus successfully solving the nonlinear Kalman filter problem [13]. The EKF method has been successfully applied to state estimation and target tracking in the fields of vehicle control [14], traffic dynamics [15] and power systems [16]. However, since all the higher-order terms in the Taylor expansion are rounded, a large truncation error will occur in EKF [17]. Moreover, as the nonlinearity of the system increases, the filtering performance will gradually decrease or even diverge. Since, Zhou et al. proposed a Strong Tracking Filter (STF)in order to make up for the limitations of EKF. This method introduces
a fading factor and calculates the optimal value of the fading factor by forcing the residual sequence to be orthogonal. This method improves the estimation accuracy and robustness of filtering. Subsequently, Julier et al. [18] proposed the unscented Kalman filter (UKF) method. In contrast to the function approximation method, UKF uses a sampling method to approximate the probability distribution of a non-linear function, thereby solving the problem of non-linear filtering. However, the UKF method can only achieve the secondorder polynomial approximation to the nonlinear model at most, and the filtering effect will gradually decline as the degree of nonlinearity increases. Then the cubature Kalman filter (CKF) was developed on the basis of UKF [19,20]. Its sampling point selection method is optimized on the basis of UKF. It is mainly based on the idea of sampling points on a spherical surface. The two important steps of transforming into spherical radial integral form and third-order spherical radial criterion are realized. However, the error covariance matrix in CKF must be positive definite; otherwise, CKF will not be able to proceed. In short, the existing nonlinear methods mainly have two approaches: Taylor expansion and sampling approximation. For the overall nonlinear function, these two methods have less than the second-order approximation ability.

The joint estimation of state and parameters to identify the time-varying parameters in the model is also an important topic. In the existing methods, people often consider the method of expanding variables for the introduction of parameters in the system, but this method undoubtedly increases the degree of non-linearity of the original non-linear function. Therefore, the abovementioned nonlinear Kalman filtering methods, such as EKF, STF, and UKF, have obvious insufficient filtering capabilities for such situations, and they have encountered greater difficulties in estimation performance. Therefore, Jiang et al. [21] used the estimation result of the least square method as the initial value of the extended filtering algorithm and proposed a joint estimation method of the state and parameters of the polynomial system based on nonlinear filtering. Wen et al. [22] put the strong tracking filter combining the theory with multi-sensor data fusion technology, the result of state estimation based on global information fusion is obtained, and a new method of multisensor state and parameter joint estimation based on strong tracking filter is established. The abbreviations of Kalman Filtering methods in the article are shown in Table 1.

Table 1. Abbreviation of Kalman filter methods.

| Kalman Filter | Extended Kalman Filter | Strong Tracking Filter | Unscented Kalman Filter | Three Stage-Kalman Filter |
| :---: | :---: | :---: | :---: | :---: |
| Abbreviation | EKF | STF | UKF | TSKF |

In addition, the research team of Wen and Sun et al. has achieved a series of results in the design of high-order extended Kalman filter in References [23-25]. Wang et al. also proposed a high-order extended Kalman filter based on maximum correlation entropy Mann filter design method [26]. Liu et al. proposed a design method of higher-order extended Kalman filter based on Kronecker product transform [27].

In this paper, for a class of nonlinear system model with multiplicative hidden variables and parameters, a step-by-step state and parameter estimation method is proposed that is different from the above two methods. Through the step-by-step linearization, each variable is solved step by step with the help of the Three-Stage Kalman Filter (TSKF), the step-by-step estimation of system state and parameters is realized, and the filtering accuracy is effectively improved.

The main contributions of this paper are as follows: (1) adopting the idea of splitting, the multiplicative basic function term is defined as the hidden variable of the system, and it is regarded as the time-varying parameter of the system to simplify the system model formally; and (2) adopting the idea of gradual linearization, respectively designing Kalman filters on hidden variables, time-varying parameters and state variables, thus improving the filtering effect of each weakly nonlinear factor.

## 2. A Strong Tracking Filter (STF) Method for Joint Estimation of the State and Parameters in a Nonlinear System

The important role of parameter estimation is for fault detection and diagnosis. In the actual system-controlled process, the process parameters often change randomly over time and cannot get the law of its change. Estimating and testing these time-varying parameters with high precision in order to ensure the normal operation of the system. Therefore, it is very necessary to establish a joint estimation method of state and parameters in a nonlinear time-varying stochastic system. Under normal circumstances, people usually use the method of expanding state variables to estimate the state and parameters jointly based on the standard extended Kalman filter.

### 2.1. Description of Nonlinear Systems with Real Varying Parameters to Be Estimated

Consider the following type of nonlinear system:

$$
\begin{gather*}
x(k+1)=f(x(k), \theta(k))+w_{x}(k)  \tag{1}\\
y(k+1)=h(k+1, x(k+1), \theta(k+1))+v(k+1) \tag{2}
\end{gather*}
$$

where the $n$-dimensional state vector is as follows:

$$
x(k)=\left[x_{1}(k), \cdots, x_{i}(k), \cdots, x_{n}(k)\right]^{T}
$$

$m$-dimensional state vector:

$$
y(k+1)=\left[y_{1}(k+1), \cdots, y_{i}(k+1), \cdots, y_{m}(k+1)\right]^{T}
$$

$n$-dimensional nonlinear function vector:

$$
f(x(k), \theta(k))=\left[f_{1}(x(k), \theta(k)), \cdots, f_{n}(x(k), \theta(k))\right]^{T}
$$

$p$-dimensional time-varying parameter vector to be estimated:

$$
\theta(k)=\left[\theta_{1}(k), \cdots, \theta_{i}(k), \cdots, \theta_{p}(k)\right]^{T}
$$

Among the above variables, state variables, $x(k)$, and parametric variables, $\theta(k)$, need to be estimated; $f(x(k), \theta(k))$ in Equation (1) is the state transition nonlinear function with time-varying parameters $\theta(k)$, and $h(x(k+1), \theta(k+1))$ in the Equation (2) is the statetransition nonlinear function with time-varying parameters $\theta(k+1)$. Moreover, $w_{x}(k)$ and $v(k+1)$ are the independent Gaussian white noise of the corresponding dimension.

Firstly, in order to realize real-time estimation of state $x(k)$ and parameters $\theta(k)$, we introduce the following auxiliary equation:

$$
\begin{equation*}
\theta(k+1)=\theta(k)+w_{\theta}(k) \tag{3}
\end{equation*}
$$

We then establish the corresponding expansion vector:

$$
\begin{aligned}
x_{e}(k) & =\left[x^{T}(k), \theta^{T}(k)\right]^{T} \\
w_{e}(k) & =\left[w_{x}^{T}(k), w_{\theta}^{T}(k)\right]^{T}
\end{aligned}
$$

Then the system (1)-(3) can be written in the following equivalent form:

$$
\begin{gather*}
x_{e}(k+1)=f_{e}\left(k, u(k), x_{e}(k)\right)+w_{e}(k)  \tag{4}\\
y(k+1)=h_{e}(k+1, x(k+1))+v(k+1) \tag{5}
\end{gather*}
$$

where

$$
\begin{align*}
f_{e}\left(x_{e}(k), k\right) & =\left[\begin{array}{c}
f(x(k), \theta(k), k) \\
\theta(k)
\end{array}\right]  \tag{6}\\
h_{e}\left(k+1, x_{e}(k+1)\right) & =h(k+1, x(k+1), \theta(k+1)) \tag{7}
\end{align*}
$$

It is assumed that the initial value of the state $x_{e}(0)$ has the following statistical characteristics:

$$
\begin{equation*}
E\left\{x_{e}(0)\right\}=\hat{x}_{e 0} ; E\left\{\left[x_{e}(0)-\hat{x}_{e 0}\right]\left[x_{e}(0)-\hat{x}_{e 0}\right]^{T}\right\}=P_{0} \tag{8}
\end{equation*}
$$

The initial values of noise $w_{e}(k), v(k+1)$, and the extended dimension state $x_{e}(0)$ have the following characteristics:

$$
\begin{align*}
& E\left\{w_{e}(k)\right\}=0 ; E\left\{w_{e}(k) w_{e}^{T}(k)\right\}=Q_{e}(k) \\
& E\{v(k+1)\}=0 ; E\left\{v(k+1) v^{T}(k+1)\right\}=R(k+1)  \tag{9}\\
& E\left\{w_{e}(k) v^{T}(k+1)\right\}=0, E\left\{w_{e}(k) x_{e}^{T}(0)\right\}=0 ; E\left\{v(k+1) x_{e}^{T}(0)\right\}=0
\end{align*}
$$

Remark 1. The strongly nonlinear system mentioned in this paper refers to the system in which the second-order and above derivatives of the system equation still affect it.

### 2.2. Design of STF for Joint Estimation of State and Parameter in the Nonlinear System

This section introduces the strong tracking filter (STF) of the joint estimation of state and parameters for the nonlinear system Equations (4) and (5) after dimension expansion. Moreover, suppose that the estimated value of the expanded state variable, $x_{e}(k)$, and the estimated error covariance matrix were obtained:

$$
\begin{align*}
\hat{x}_{e}(k \mid k) & =E\left\{x_{e}(k) \mid \hat{x}_{e 0}, y(1), y(2), \cdots, y(k)\right\} \\
P_{e}(k \mid k) & =E\left\{\left[x_{e}(k)-\hat{x}_{e}(k \mid k)\right]\left[x_{e}(k)-\hat{x}_{e}(k \mid k)\right]^{T}\right\} \tag{10}
\end{align*}
$$

and assume that the observed value $y(k+1)$ at time $k+1$ is known. The goal of this section is to obtain the estimated value of the state and the estimated error covariance matrix, as follows:

$$
\begin{equation*}
\hat{x}_{e}(k \mid k), P_{e}(k \mid k) \xrightarrow{y(k+1)} \hat{x}_{e}(k+1 \mid k+1), P_{e}(k+1 \mid k+1) \tag{11}
\end{equation*}
$$

Therefore, design the following nonlinear system (4) and (5) of the state and parameter joint estimation of the strong tracking filter (STF):

$$
\begin{equation*}
\hat{x}_{e}(k+1 \mid k+1)=\hat{x}_{e}(k+1 \mid k)+K_{e}(k+1) \widetilde{y}_{e}(k+1 \mid k) \tag{12}
\end{equation*}
$$

where the measurement forecast estimation error is as follows:

$$
\begin{align*}
\tilde{y}_{e}(k+1 \mid k) & =y(k+1)-\hat{y}_{e}(k+1 \mid k)  \tag{13}\\
& =y(k+1)-h_{e}\left(\hat{x}_{e}(k+1 \mid k)\right)
\end{align*}
$$

The key to the design of the strong tracking filter (STF) is to obtain the forcibly obtained gain matrix, $K(k+1)$, which must satisfy the following orthogonality principle:

$$
\begin{gather*}
E\left\{\left[x_{e}(k)-\hat{x}_{e}(k \mid k)\right]\left[x_{e}(k)-\hat{x}_{e}(k \mid k)\right]^{T}\right\}=\min  \tag{14}\\
E\left\{\widetilde{y}(k+j) \widetilde{y}^{T}(k)\right\}=0, k=1,2, \cdots, N, j=1,2, \cdots, N \tag{15}
\end{gather*}
$$

The Equation (14) is the performance index of the extended Kalman filter. The Equation (15) expresses the orthogonality of innovation, and the residual sequence at different times is required to be orthogonal to each other everywhere; that is, the orthogonal principle

Equation (15) is introduced on the basis of EKF. When the model is determined, the filter will operate normally. At this time, the STF degenerates to an EKF based on the performance index (14), and (15) will not have any adjustment effect. When the model is uncertain, the state estimate of the filter will deviate from the true state of the system, which will definitely be reflected in the output residual. At this time, the gain matrix will be adjusted online, and the forced Equation (15) is still valid, so that the residual sequence remains orthogonal to each other, and the filter can maintain accurate tracking of the actual system state.

The specific steps to establish STF are as follows:

$$
\begin{gather*}
\hat{x}_{e}(k+1 \mid k)=f_{e}\left(\hat{x}_{e}(k \mid k)\right)  \tag{16}\\
P_{e}(k+1 \mid k)=\lambda(k) A_{e}\left(\hat{x}_{e}(k \mid k)\right) P_{e}(k \mid k) A^{T}\left(\hat{x}_{e}(k \mid k)\right)+Q_{e}(k)  \tag{17}\\
K_{e}(k+1)=P_{e}(k+1 \mid k) H_{e}^{T}\left(\hat{x}_{e}(k+1 \mid k)\right) \\
\times\left[H_{e}\left(\hat{x}_{e}(k+1 \mid k)\right) P_{e}(k+1 \mid k) H_{e}^{T}\left(\hat{x}_{e}(k+1 \mid k)\right)\right]^{-1} \tag{18}
\end{gather*}
$$

$\lambda(k)$ in Equation (17) is the adaptive fading factor of the system, and we have

$$
\lambda(k)=\left\{\begin{array}{l}
\lambda_{1}(k), \lambda_{1}(k) \geq 1  \tag{19}\\
1, \lambda_{1}(k)<1
\end{array}\right.
$$

where

$$
\begin{gather*}
\lambda_{1}(k)=\frac{\operatorname{Tr}(N(k))}{\operatorname{Tr}(M(k))}  \tag{20}\\
N(k)=V_{1}(k)-K_{e}\left(\hat{x}_{e}(k+1 \mid k)\right) Q_{e}(k) H_{e}^{T}\left(\hat{x}_{e}(k+1 \mid k)\right)-l(k) R_{e}(k+1)  \tag{21}\\
M(k)=H_{e}\left(\hat{x}_{e}(k+1 \mid k)\right) A_{e}(\hat{x}(k \mid k)) P_{e}(k \mid k) \\
\times A_{e}^{T}\left(\hat{x}_{e}(k \mid k)\right) H_{e}^{T}\left(\hat{x}_{e}(k+1 \mid k)\right)  \tag{22}\\
A_{e}\left(\hat{x}_{e}(k \mid k)\right)=\left.\frac{\partial f_{e}\left(x_{e}(k)\right)}{\partial x_{e}}\right|_{x_{e}(k)=\hat{x}_{e}(k \mid k)}  \tag{23}\\
H_{e}\left(\hat{x}_{e}(k+1 \mid k)\right)=\left.\frac{\partial h_{e}(x(k+1))}{\partial x_{e}}\right|_{x_{e}(k+1)=\hat{x}_{e}(k+1 \mid k)}  \tag{24}\\
V_{1}(k)=\left\{\begin{array}{l}
\widetilde{y}_{e}(1) \widetilde{y}_{e}^{T}(1), k=0 \\
\frac{\rho V_{1}(k-1)+\widetilde{y}_{e}(k) \tilde{y}_{e}^{T}(k)}{1+\rho}, k \geq 1
\end{array}\right. \tag{25}
\end{gather*}
$$

where $\rho$ is the forgetting factor of STF, and $l(k)$ is the weakening factor.
The core of the principle of orthogonality is Equation (15). When Equation (14) is replaced by other performance indicators, other similar orthogonality principles can be obtained, giving it the properties of a strong tracking filter.

Remark 2. Degradation mechanism analysis: STF->EKF->KF.
(1) A strong tracking filter (STF) is composed of Equations (14)-(27), when $\lambda(k) \equiv 1$ in Equation (19), and STF degenerates the traditional EKF.
(2) The nonlinear functions in Equations (1) and (2) degenerate into the following:

$$
\begin{align*}
f\left(x_{e}(k), k\right) & :=A(k) x(k)  \tag{26}\\
h\left(x_{e}(k+1),(k+1)\right) & :=H(k+1) x(k+1) \tag{27}
\end{align*}
$$

Then STF degenerates the traditional Kalman filter (KF).

## 3. A Class of Nonlinear Systems Described by the Product of States, Parameters and Nonlinear Functions

### 3.1. Description of Non-Linear System

The following is a strong nonlinear system described by the product of states, parameters and nonlinear functions:

$$
\begin{gather*}
x_{i}(k+1)=\sum_{j=1}^{n} a_{i j} x_{j}(k) \times \theta_{i}(k) \times f_{i}(x(k))+w_{i}(k) ; i=1,2, \cdots, n  \tag{28}\\
z_{i}(k+1)=  \tag{29}\\
h_{i}^{(1)}(x(k+1)) \times h_{i}^{(2)}\left(\theta_{i}(k+1)\right) \times h_{i}^{(3)}\left(f_{i}(x(k))\right) \\
\\
+v_{i}(k+1) ; i=1,2, \cdots, m
\end{gather*}
$$

where $k$ is the discrete time, $x(k) \in R^{n}$ is the system state vector, $\lambda(k) \in R^{p}$ is the timevarying parameter to be estimated and $z(k+1) \in R^{m}$ is the observation vector of the state. Moreover, $f_{j}(x(k))$ is the basic multiplicative factor about the state variable; $h_{i}^{(1)}(x(k+1))$ is the nonlinear function about the parameter $x(k+1), h_{i}^{(2)}(\theta(k+1))$ is the nonlinear function about the parameter $\theta(k+1)$ and $h_{i}^{(3)}(f(x(k+1))$ is the compound nonlinear function about the state $x(k+1)$.

Suppose 1. $w(k)$ and $v(k+1)$ are uncorrelated white noise sequences, and they meet the following conditions

$$
\begin{gather*}
E\{w(k)\}=0, E\{v(k)\}=0 \\
E\left\{w(k) w^{T}(j)\right\}=Q(k) \delta_{k, j}  \tag{30}\\
E\left\{v(k) v^{T}(j)\right\}=R(k+1) \delta_{k, j}
\end{gather*}
$$

While $k=j, \delta_{k, j}=1 ;$ otherwise, $\delta_{k, j}=0$.
Suppose 2. The initial state value and the system noise are independent of each other, and they meet the following conditions:

$$
\begin{gathered}
E\left\{w(k) v^{T}(k)\right\}=0, E\left\{x(0) v^{T}(k)\right\}=0, E\left\{x(0) w_{1}^{T}(k)\right\}=0 \\
E\{x(0)\}=\hat{x}_{0}, E\left\{\left[x(0)-\hat{x}_{0}\right]\left[x(0)-\hat{x}_{0}\right]^{T}\right\}=P_{0}, P_{0} \geq 0
\end{gathered}
$$

### 3.2. Gradually Linearizing of the Strongly Nonlinear Systems

First, define the basic nonlinear function $f_{j}(x(k), k)$ as the hidden variable about the system state variable, $x(k)$.

$$
\begin{equation*}
\alpha_{i}(k):=f_{j}(x(k), k), i=1,2, \cdots, n \tag{31}
\end{equation*}
$$

Then Equations (28) and (29) can be written as follows:

$$
\begin{gather*}
x_{i}(k+1)=\sum_{j=1}^{n} a_{i j} x_{j}(k) \times \lambda_{i}(k) \times \alpha_{i}(k)+w_{i}(k) ; i=1,2, \cdots, n  \tag{32}\\
z_{i}(k+1)=h_{i}^{(1)}(x(k+1)) \times h_{i}^{(2)}(\theta(k+1)) \times h_{i}^{(3)}(\alpha(k+1))+v_{i}(k+1) ; i=1,2, \cdots, m \tag{33}
\end{gather*}
$$

(1) Pseudo-linearized representation for hidden variables:

$$
\begin{gather*}
x(k+1)=A_{\alpha}(x(k), \theta(k)) \times \alpha(k)+w(k)  \tag{34}\\
z(k+1)=H_{\alpha}\left(h^{(1)}(x(k+1)), h^{(2)}(\theta(k+1))\right) \times h^{(3)}(\alpha(k+1))+v_{\alpha}(k+1) \tag{35}
\end{gather*}
$$

where

$$
\begin{align*}
& A_{\alpha}(x(k), \theta(k))=\operatorname{diag}\left\{\sum_{j=1}^{n} a_{1 j} x_{j}(k) \theta_{1}(k), \cdots, \sum_{j=1}^{n} a_{i j} x_{j}(k) \theta_{i}(k), \cdots, \sum_{j=1}^{n} a_{n j} x_{j}(k) \theta_{n}(k)\right\}  \tag{36}\\
& H_{\alpha}\left(h^{(1)}\left(x(k+1), h^{(2)}(\theta(k+1))\right)=\operatorname{diag}\left\{h_{1}^{(1)}(x(k+1)) \times h_{1}^{(2)}(\theta(k+1)),\right.\right.  \tag{37}\\
& \left.\quad \cdots, h_{i}^{(1)}(x(k+1)) \times h_{i}^{(2)}(\theta(k+1)), \cdots, h_{m}^{(1)}(x(k+1)) \times h_{m}^{(2)}(\theta(k+1))\right\}
\end{align*}
$$

(2) Pseudo-linear representation for time-varying parameters:

$$
\begin{gather*}
x(k+1)=A_{\theta}(x(k), \alpha(k)) \times \theta(k)+w_{\theta}(k)  \tag{38}\\
z(k+1)=H_{\theta}\left(h^{(1)}(x(k+1)), h^{(3)}(\alpha(k+1))\right) \times h^{(2)}(\theta(k+1))+v(k+1) \tag{39}
\end{gather*}
$$

where

$$
\begin{gather*}
A_{\theta}(x(k), \alpha(k))=\operatorname{diag}\left\{\sum_{j=1}^{m} a_{1 j} x_{j}(k) \alpha_{1}(k), \cdots, \sum_{j=1}^{m} a_{i j} x_{j}(k) \alpha_{i}(k), \cdots, \sum_{j=1}^{m} a_{n j} x_{j}(k) \alpha_{n}(k)\right\}  \tag{40}\\
H_{\theta}\left(h^{(1)}\left(x(k+1), h^{(3)}(\alpha(k+1))\right)=\operatorname{diag}\left\{h_{1}^{(1)}(x(k+1)) \times h_{1}^{(3)}(\alpha(k+1)),\right.\right.  \tag{41}\\
\left.\quad \cdots, h_{i}^{(1)}(x(k+1)) \times h_{i}^{(3)}(\alpha(k+1)), \cdots, h_{m}^{(1)}(x(k+1)) \times h_{m}^{(3)}(\alpha(k+1))\right\}
\end{gather*}
$$

(3) Pseudo-linearized representation of state variables:

$$
\begin{gather*}
x(k+1)=A_{x}(\theta(k), \alpha(k)) \times x(k)+w(k)  \tag{42}\\
z(k+1)=H_{x}\left(h^{(2)}\left(\theta(k+1), h^{(3)}(\alpha(k+1))\right) \times h^{(1)}\left(x(k+1)+v_{x}(k+1)\right.\right. \tag{43}
\end{gather*}
$$

where

$$
\begin{gather*}
A_{x}(\theta(k), \alpha(k))= \\
{\left[\begin{array}{ccccc}
a_{11} \theta_{1}(k) \alpha_{1}(k) & \cdots & a_{1 i} \theta_{1}(k) \alpha_{1}(k) & \cdots & a_{1 n} \theta_{1}(k) \alpha_{1}(k) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
a_{i 1} \theta_{i}(k) \alpha_{i}(k) & \cdots & a_{i i} \theta_{i}(k) \alpha_{i}(k) & \vdots & a_{i n} \theta_{i}(k) \alpha_{i}(k) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
a_{n 1} \theta_{n}(k) \alpha_{n}(k) & \cdots & a_{n i} \theta_{n}(k) \alpha_{n}(k) & \cdots & a_{n n} \theta_{n}(k) \alpha_{n}(k)
\end{array}\right]}  \tag{44}\\
H_{x}\left(h^{(2)}\left(\theta(k+1), h^{(3)}(\alpha(k+1))\right)=\operatorname{diag}\left\{h_{1}^{(2)}(\theta(k+1)) \times h_{1}^{(3)}(\alpha(k+1)),\right.\right.  \tag{45}\\
\left.\cdots, h_{i}^{(2)}(\theta(k+1)) \times h_{i}^{(3)}(\alpha(k+1)), \cdots, h_{m}^{(2)}(\theta(k+1)) \times h_{m}^{(3)}(\alpha(k+1))\right\}
\end{gather*}
$$

Remark 3. The abovementioned linear processes for hidden variables, $\alpha(k)$; time-varying parameter variables, $\theta(k)$; and system state variables, $x(k)$, are all pseudo-linear processes. This is because the abovementioned processes are only equivalent transformation processes in the original system space and do not make any substantial changes. Therefore, they need to be re-linearized in the expanded space.
(4) The linear dynamic modeling of parametric variables $\theta(k)$ and hidden state variables $\alpha(k)$.

This subsection establishes following the linear dynamic correlation model with time as the variable between $\theta(k+1), \alpha(k+1), \theta(k), \alpha(k)$ and $x(k)$.

$$
\begin{align*}
\alpha(k+1) & =S_{\alpha}(k) x(k)+G_{\alpha}(k) \alpha(k)+L_{\alpha}(k) \theta(k)+w_{\alpha}(k)  \tag{46}\\
\theta(k+1) & =S_{\theta}(k) x(k)+G_{\theta}(k) \alpha(k)+L_{\theta}(k) \theta(k)+w_{\theta}(k) \tag{47}
\end{align*}
$$

where

$$
\begin{equation*}
S_{\theta}(k), G_{\theta}(k), L_{\theta}(k) ; S_{\alpha}(k), G_{\alpha}(k), L_{\alpha}(k) \tag{48}
\end{equation*}
$$

They are the matrix of the system to be identified with the appropriate dimensions; $w_{\theta}(k)$ and $w_{\alpha}(k)$ are the error vectors of the model to be modeled.

Remark 4. (1) Based on the strong tracking filter (STF) introduced in Section 2, the estimated sequence of state variables, $x(k)$; time-varying parameter variables, $\theta(k)$; and hidden variables, $\alpha(k)$, can be obtained as follows:

$$
\begin{equation*}
\hat{x}(k \mid k), \hat{\theta}(k \mid k), \hat{\alpha}(k \mid k)) ; k=1,2, \cdots, M \tag{49}
\end{equation*}
$$

(1) Use the least-square method to determine the statistical characteristics of model parameters and modeling errors:

Based on the Equations (46) and (47), the estimated value of the matrix parameter to be estimated can be obtained by using the linear least-squares method and is written as follows:

$$
\begin{equation*}
\hat{S}_{\lambda}, \hat{G}_{\lambda}, \hat{L}_{\lambda} ; \hat{S}_{\alpha} \hat{G}_{\alpha}, \hat{L}_{\alpha} \tag{50}
\end{equation*}
$$

Using Equations (46), (47) and (50), we can obtain the following:

$$
\begin{align*}
\widetilde{\theta}(k+1) & =\theta(k+1)-\hat{\theta}(k+1) \\
& =\theta(k+1)-\hat{S}_{\theta} x(k)-\hat{G}_{\theta} \alpha(k)-\hat{L}_{\theta} \theta(k)  \tag{51}\\
\widetilde{\alpha}(k+1) & =\alpha(k+1)-\hat{\alpha}(k+1) \\
& =\alpha(k+1)-\hat{S}_{\alpha} x(k)-\hat{G}_{\alpha} \alpha(k)-\hat{L}_{\alpha} \theta(k) \tag{52}
\end{align*}
$$

Then the statistical characteristics of modeling errors of $w_{\theta}(k)$ and $w_{\alpha}(k)$ can be determined separately with the help of the sequence $\{\widetilde{\theta}(k)\},\{\widetilde{\alpha}(k)\}$.
(2) Use Multidimensional Taylor Network to determine the statistical characteristics of model parameters and modeling errors.

The Multidimensional Taylor Network that has been effectively used in practice can be borrowed to determine the estimated value of the parameters $S_{\theta}(k), G_{\theta}(k), L_{\theta}(k)$, $S_{\alpha}(k), G_{\alpha}(k)$ and $L_{\alpha}(k)$ in (46) and (47). Then we can use the Equations (51) and (52) method to determine the statistical characteristics of the modeling error of $w_{\theta}(k)$ and $w_{\alpha}(k)$, respectively.
(3) Use the idea of random walk to determine the statistical characteristics of model parameters and modeling errors.

In the absence of any prior information, set it as follows:

$$
\begin{equation*}
S_{\theta}(k)=0, G_{\theta}(k)=0, L_{\theta}(k)=I ; S_{\alpha}(k)=0, \quad G_{\alpha}(k)=I, L_{\alpha}(k)=0 \tag{53}
\end{equation*}
$$

both $\theta(k)$ and $\alpha(k)$ are regarded as random processes conforming to random walks, and the methods of Equations (51) and (52) are used to determine the statistical characteristics of modeling errors of $w_{\theta}(k)$ and $w_{\alpha}(k)$.

## 4. The Design of Three-Stage Kalman Filter for Estimating State, Parameters and Hidden Variables

Assuming the initial values $x(0), \lambda(0), \alpha(0)$ of estimated state, parameters and hidden variables, there are the following statistical characteristics:

$$
\begin{align*}
& E\{x(0)\}=\hat{x}_{0} ; E\left\{\left[x(0)-\hat{x}_{0}\right]\left[x(0)-\hat{x}_{0}\right]^{T}\right\}=P_{x 0} \\
& E\{\theta(0)\}=\hat{\theta}_{0} ; E\left\{\left[\theta(0)-\hat{\lambda}_{0}\right]\left[\theta(0)-\hat{\theta}_{0}\right]^{T}\right\}=P_{\theta 0}  \tag{54}\\
& E\{\alpha(0)\}=\hat{\alpha}_{0} ; E\left\{\left[\alpha(0)-\hat{\alpha}_{0}\right]\left[\alpha(0)-\hat{\alpha}_{0}\right]^{T}\right\}=P_{\alpha 0}
\end{align*}
$$

The estimated process noise, $w_{x}(k), w_{\theta}(k)$ and $w_{\alpha}(k)$; observation noise, $v_{x}(k), v_{\theta}(k)$, and $v_{\alpha}(k)$; and initial value, $x(0), \lambda(0)$, and $\alpha(0)$, are all statistically independent of each other.

Obtained sequence of observations:

$$
\begin{equation*}
y(1), y(2), \cdots, y(k) \tag{55}
\end{equation*}
$$

further gain

$$
\begin{align*}
& \hat{x}(k \mid k)=E\left\{x(k) \mid \hat{x}_{0}, y(1), y(2), \cdots, y(k)\right\} \\
& P_{x}(k \mid k)=E\left\{[x(k)-\hat{x}(k \mid k)][x(k)-\hat{x}(k \mid k)]^{T}\right\}  \tag{56}\\
& \hat{\theta}(k \mid k)=E\left\{\theta(k) \mid \hat{\theta}_{0}, y(1), y(2), \cdots, y(k)\right\} \\
& P_{\theta}(k \mid k)=E\left\{\left[\theta(k-\hat{\theta}(k \mid k)][\theta(k)-\hat{\theta}(k \mid k)]^{T}\right\}\right.  \tag{57}\\
& \hat{\alpha}(k \mid k)=E\left\{\alpha(k) \mid \hat{\alpha}_{0}, y(1), y(2), \cdots, y(k)\right\} \\
& P_{\alpha}(k \mid k)=E\left\{[\alpha(k)-\hat{\alpha}(k \mid k)][\alpha(k)-\hat{\alpha}(k \mid k)]^{T}\right\} \tag{58}
\end{align*}
$$

The goal of this section is as follows:

$$
\begin{align*}
& \hat{x}(k \mid k), P_{x}(k \mid k) ; \hat{\theta}(k \mid k), P_{\hat{\theta}}(k \mid k) ; \hat{\alpha}(k \mid k), P_{\alpha}(k \mid k) \\
& \xrightarrow{\xrightarrow[\hat{\alpha}(k+1 \mid k+1)]{y(k+1)}} \hat{\alpha}(k+1 \mid k+1), P_{\alpha}(k+1 \mid k+1) \\
& \xrightarrow[\hat{\alpha}(k+1 \mid k+1), \hat{\theta}(k+1 \mid k+1)]{y(k+1)} \hat{\theta}(k+1 \mid k+1), P_{\theta}(k+1 \mid k+1)  \tag{59}\\
&
\end{align*}
$$

### 4.1. The Linear Kalman Filter Design for Solving Hidden Variables $\alpha(k+1)$

The pseudo linearization system composed of Equations (46) and (35) is rewritten as follows:

$$
\begin{gather*}
\alpha(k+1)=S_{\alpha}(k) x(k)+G_{\alpha}(k) \alpha(k)+L_{\alpha}(k) \lambda(k)+w_{\alpha}(k)  \tag{60}\\
z(k+1)=H_{\alpha}\left(h^{(1)}(x(k+1)), h^{(2)}(\lambda(k+1))\right) \times h^{(3)}(\alpha(k+1))+v_{\alpha}(k+1) \tag{61}
\end{gather*}
$$

and suppose that the initial value $\alpha(0)$ of the parameter has the following statistical characteristics:

$$
\begin{equation*}
E\{\alpha(0)\}=\hat{\alpha}_{0} ; E\left\{\left[\alpha(0)-\hat{\alpha}_{0}\right]\left[\alpha(0)-\hat{\alpha}_{0}\right]^{T}\right\}=P_{\alpha 0} \tag{62}
\end{equation*}
$$

The initial values of noise $v_{\alpha}(k+1), w_{\alpha}(k)$ and parameters $\alpha(0)$ have the following characteristics:

$$
\begin{align*}
& E\left\{w_{\alpha}(k)\right\}=0 ; E\left\{v_{\alpha}(k+1)\right\}=0 ; E\left\{w_{\alpha}(k) v_{\alpha}^{T}(k+1)\right\}=0 \\
& E\left\{w_{\alpha}(k) \alpha^{T}(0)\right\}=0 ; E\left\{v_{\alpha}(k+1) \alpha^{T}(0)\right\}=0 \tag{63}
\end{align*}
$$

The goal of this section is to achieve the following:

$$
\begin{align*}
& \hat{x}(k \mid k), P_{x}(k \mid k) ; \hat{\theta}(k \mid k), P_{\theta}(k \mid k) ; \hat{\alpha}(k \mid k), P_{\alpha}(k \mid k) \\
& \underset{\text { formula }(4.2 .2)}{y(k+1)} \alpha(k+1 \mid k+1), P_{\alpha}(k+1 \mid k+1) \tag{64}
\end{align*}
$$

Under the conditions of known $\hat{x}(k \mid k), \hat{\theta}(k \mid k), \hat{\alpha}(k \mid k)$ and online observations, Equation (61) is used to describe the linearization of hidden variables $\alpha(k+1)$ to construct the Kalman filter to find the following:

$$
\begin{equation*}
\hat{\alpha}(k+1 \mid k+1), P_{\alpha}(k+1 \mid k+1) \tag{65}
\end{equation*}
$$

The Design of the Linear Kalman Filter to Estimate Value of the Hidden Variable $\alpha(k+1)$
To design a linear filter estimator for solving hidden variables, we have the following:

$$
\begin{align*}
\hat{\alpha}(k+1 \mid k+1) & =E\{\alpha(k+1) \mid \hat{x}(k \mid k), \hat{\theta}(k \mid k), \hat{\alpha}(k \mid k), y(k+1)\} \\
& =\hat{\alpha}(k+1 \mid k)+K_{\alpha}(k+1)\left[y(k+1)-\hat{y}_{\alpha}(k+1 \mid k)\right]  \tag{66}\\
\hat{\alpha}(k+1 \mid k+1) & =E\left\{[\alpha(k+1)-\hat{\alpha}(k+1 \mid k+1)][\alpha(k+1)-\hat{\alpha}(k+1 \mid k+1)]^{T}\right\}
\end{align*}
$$

where there are Equations (91)-(107) in sequence.
(1) Hidden variable forecast estimates and forecast estimate error covariance matrix:

Based on Equation (84), the predicted estimated value of the hidden variable can be obtained as follows:

$$
\begin{equation*}
\hat{\alpha}(k+1 \mid k)=S_{\alpha}(k) \hat{x}(k \mid k)+G_{\alpha}(k) \hat{\alpha}(k \mid k)+L_{\alpha}(k) \hat{\theta}(k \mid k) \tag{67}
\end{equation*}
$$

and forecast estimate error:

$$
\begin{align*}
\hat{\alpha}(k+1 \mid k)= & \alpha(k+1)-\hat{\alpha}(k+1 \mid k) \\
= & S_{\alpha}(k) x(k)+G_{\alpha}(k) \alpha(k)+L_{\alpha}(k) \theta(k)+w_{\alpha}(k)  \tag{68}\\
& -S_{\alpha}(k) \hat{x}(k \mid k)-G_{\alpha}(k) \hat{\alpha}(k \mid k)-L_{\alpha}(k) \hat{\theta}(k \mid k) \\
= & S_{\alpha}(k) \widetilde{x}(k \mid k)+G_{\alpha}(k) \widetilde{\alpha}(k \mid k)+L_{\alpha}(k) \widetilde{\theta}(k \mid k)+w_{\alpha}(k)
\end{align*}
$$

and predictive estimate error covariance matrix:

$$
\begin{align*}
P_{\alpha}(k+1 \mid k)= & E\left\{\widetilde{\alpha}(k+1 \mid k) \widetilde{\alpha}^{T}(k+1 \mid k)\right\} \\
= & \theta_{\alpha}(k)\left[S_{\alpha}(k) P_{x}(k \mid k) S_{\alpha}^{T}(k)+G_{\alpha}(k) P_{\alpha}(k \mid k) G_{\alpha}^{T}(k)\right.  \tag{69}\\
& \left.+L_{\alpha}(k) P_{\alpha}(k \mid k) L_{\alpha}^{T}(k)\right]+Q_{\alpha}(k)
\end{align*}
$$

(2) Prediction estimates and prediction estimation errors of measured values of hidden variables:

Based on Equation (61), the predicted estimated value for the measured value of the hidden variable can be obtained as follows:

$$
\begin{equation*}
\left.\hat{z}(k+1 \mid k)=H_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k))\right), h^{(2)}(\hat{\theta}(k+1 \mid k))\right) \times h^{(2)}(\hat{\alpha}(k+1 \mid k)) \tag{70}
\end{equation*}
$$

and forecast estimation error vector:

$$
\begin{align*}
\widetilde{z}_{\alpha}(k+1 \mid k)= & z(k+1)-\hat{z}_{\alpha}(k+1 \mid k) \\
= & H_{\alpha}\left(h^{(1)}(x(k+1)), h^{(2)}(\theta(k+1)) \times h^{(3)}(\alpha(k+1))+v(k+1)\right. \\
& -H_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\theta}(k+1 \mid k+1)) \times h^{(3)}(\hat{\alpha}(k+1 \mid k))\right. \\
\approx & H_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\theta}(k+1 \mid k))\right. \\
& \times\left[h^{(3)}(\alpha(k+1))-h^{(3)}(\hat{\alpha}(k+1 \mid k))\right]+v(k+1) \\
= & H_{\alpha}(\hat{x}(k+1 \mid k), \hat{\theta}(k+1 \mid k))  \tag{71}\\
& \times H_{\alpha}^{(3)}(\hat{\alpha}(k+1 \mid k))[\alpha(k+1)-\hat{\alpha}(k+1 \mid k)]+v(k+1) \\
\approx & H_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\theta}(k+1 \mid k))\right. \\
& \times H_{\alpha}^{(3)}(\hat{\alpha}(k+1 \mid k)) \widetilde{\alpha}(k+1 \mid k)+v_{\alpha}(k+1) \\
= & \bar{H}_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\theta}(k+1 \mid k)), \hat{\alpha}(k+1 \mid k)\right) \\
& \times \widetilde{\alpha}(k+1 \mid k)+v_{\alpha}(k+1)
\end{align*}
$$

and predictive estimate error covariance matrix:

$$
\begin{align*}
& P_{\alpha, z z}(k+1 \mid k) \\
& =E\left\{\widetilde{z}_{\alpha}(k+1 \mid k) \widetilde{z}_{\alpha}^{T}(k+1 \mid k)\right\} \\
& =\left(\bar{H}_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\theta}(k+1 \mid k)), \hat{\alpha}(k+1 \mid k)\right)\right) P_{\alpha}(k+1 \mid k)  \tag{72}\\
& \times\left(\bar{H}_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\theta}(k+1 \mid k)), \hat{\alpha}(k+1 \mid k)\right)\right)^{T}+R_{\alpha}(k+1)
\end{align*}
$$

where

$$
\begin{align*}
& \bar{H}_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\theta}(k+1 \mid k)), \hat{\alpha}(k+1 \mid k)\right) \\
& \quad:=H_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\theta}(k+1 \mid k)) H_{\alpha}^{(3)}(\hat{\alpha}(k+1 \mid k))\right. \tag{73}
\end{align*}
$$

Remark 5. It is explained as follows in the derivation process of (70).
(1) The first " $\approx$ " is due to the approximate approximation used:

$$
\begin{equation*}
H_{\alpha}\left(h ^ { ( 1 ) } \left(x(k+1), h^{(2)}(\theta(k+1)) \approx H_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\theta}(k+1 \mid k))\right.\right.\right. \tag{74}
\end{equation*}
$$

(2) The second " $\approx$ " sign is due to the approximation of the first-order Taylor expansion:

$$
\begin{equation*}
h^{(3)}(\alpha(k+1))-h^{(3)}(\hat{\alpha}(k+1 \mid k)) \approx H_{\alpha}^{(3)}(\hat{\alpha}(k+1 \mid k)) \widetilde{\alpha}(k+1 \mid k) \tag{75}
\end{equation*}
$$

where

$$
\begin{equation*}
H_{\alpha}^{(3)}(\hat{\alpha}(k+1 \mid k))=\left.\frac{\partial h^{(3)}(\alpha(k+1))}{\partial \alpha(k+1)}\right|_{\alpha(k+1)=\hat{\alpha}(k+1 \mid k)} \tag{76}
\end{equation*}
$$

(3) The noise term changes as it follows the derivation process of (122):

$$
\begin{equation*}
v(k+1) \xrightarrow[\text { second" } \approx "]{\text { first" } \approx "} v_{\alpha}(k+1) \tag{77}
\end{equation*}
$$

this is caused by the introduction of new modeling errors in the two " $\approx$ ".
(4) In Equation (69), the parameter $\lambda_{\alpha}(k)$ is the fading factor of Equation (17) determined by Equation (19) in Section 2.

At this point, having completed the linearization process of using the strong tracking filter (STF) to solve the hidden variable estimated value and the estimated error covariance matrix, and the corresponding STF will be designed below.
(3) Filter design for real-time estimation of hidden variables:

$$
\begin{equation*}
\hat{\alpha}(k+1 \mid k+1)=\hat{\alpha}(k+1 \mid k)+K_{\alpha}(k+1)\left[y(k+1)-\hat{y}_{\alpha}(k+1 \mid k)\right] \tag{78}
\end{equation*}
$$

Calculate the estimated error of the hidden variable as follows:

$$
\begin{align*}
& \widetilde{\alpha}(k+1 \mid k+1) \\
& =\alpha(k+1)-\hat{\alpha}(k+1 \mid k+1) \\
& =\alpha(k+1)-\hat{\alpha}(k+1 \mid k)-K_{\alpha}(k+1)\left[y(k+1)-\hat{y}_{\alpha}(k+1 \mid k)\right] \\
& =\widetilde{\alpha}(k+1 \mid k)-K_{\alpha}(k+1)\left[y(k+1)-\hat{y}_{\alpha}(k+1 \mid k)\right] \\
& =\widetilde{\alpha}(k+1 \mid k)-K_{\alpha}(k+1)\left[\begin{array}{c}
\left.H_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\lambda}(k+1 \mid k+1)), \hat{\alpha}(k+1 \mid k)\right)\right) \\
\left.\quad \times \widetilde{\alpha}(k+1 \mid k)+v_{\alpha}(k+1)\right]
\end{array}\right.
\end{align*}
$$

a direct sum decomposition of the observation vector:

$$
\begin{align*}
z(k+1)= & \hat{z}_{\alpha}(k+1 \mid k)+\widetilde{z}_{\alpha}(k+1 \mid k) \\
= & \hat{z}_{\alpha}(k+1 \mid k)+\bar{H}_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\lambda}(k+1 \mid k+1)), \hat{\alpha}(k+1 \mid k)\right) \\
& \times \widetilde{\alpha}(k+1 \mid k)+v_{\alpha}(k+1) \tag{80}
\end{align*}
$$

utilize the Orthogonal Principle:

$$
\begin{equation*}
\left\{\widetilde{\alpha}(k+1 \mid k+1) z^{T}(k+1)\right\}=0 \tag{81}
\end{equation*}
$$

Then solve the gain matrix $K_{\alpha}(k+1)$ in Equation (78).

$$
\begin{align*}
& K_{\alpha}(k+1) \\
& =P_{\alpha}(k+1 \mid k)\left(\bar{H}_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\lambda}(k+1 \mid k+1)), \hat{\alpha}(k+1 \mid k)\right)\right)^{T} \\
& \quad \times\left[\left(\bar{H}_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\lambda}(k+1 \mid k+1)), \hat{\alpha}(k+1 \mid k)\right)\right) P(k+1 \mid k)\right.  \tag{82}\\
& \left.\quad \times\left(\bar{H}_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\lambda}(k+1 \mid k+1)), \hat{\alpha}(k+1 \mid k)\right)\right)^{T}+R_{\alpha}(k+1)\right]^{-1}
\end{align*}
$$

(4) Calculate hidden variables and estimate the error covariance matrix in real time:

$$
\begin{align*}
& P_{\alpha}(k+1 \mid k+1) \\
& =E\left\{[\alpha(k+1)-\hat{\lambda}(k+1 \mid k)][\alpha(k+1)-\hat{\lambda}(k+1 \mid k)]^{T}\right\} \\
& =\left[I-K_{\alpha}(k+1)\left(\bar{H}_{\alpha}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(2)}(\hat{\lambda}(k+1 \mid k+1)), \hat{\alpha}(k+1 \mid k)\right)\right)\right]  \tag{83}\\
& \quad \times P_{\alpha}(k+1 \mid k)
\end{align*}
$$

### 4.2. The Design of Linear Kalman Filter for Solving Parameters $\theta(k+1)$

Based on the pseudo linearization system composed of Equations (47) and (39), it is rewritten as follows:

$$
\begin{gather*}
\theta(k+1)=S_{\theta}(k) x(k)+G_{\theta}(k) \alpha(k)+L_{\theta}(k) \theta(k)+w_{\theta}(k)  \tag{84}\\
z(k+1)=H_{\theta}\left(h^{(1)}(x(k+1)), h^{(3)}(\alpha(k+1))\right) \times h^{(1)}(\theta(k+1))+v(k+1) \tag{85}
\end{gather*}
$$

and suppose that the initial value of the parameter $\theta(0)$ has the following statistical characteristics:

$$
\begin{equation*}
E\{\theta(0)\}=\hat{\theta}_{0} ; E\left\{\left[\theta(0)-\hat{\theta}_{0}\right]\left[\theta(0)-\hat{\theta}_{0}\right]^{T}\right\}=P_{\theta 0} \tag{86}
\end{equation*}
$$

The initial values of noise $v_{\theta}(k+1)$ and $w_{\theta}(k)$ and parameter $\theta(0)$ have the following characteristics:

$$
\begin{align*}
& E\left\{w_{\theta}(k)\right\}=0 ; E\left\{v_{\theta}(k+1)\right\}=0 ; E\left\{w_{\theta}(k) v_{\theta}^{T}(k+1)\right\}=0 \\
& E\left\{w_{\theta}(k) \theta^{T}(0)\right\}=0 ; E\left\{v_{\theta}(k+1) \theta^{T}(0)\right\}=0 \tag{87}
\end{align*}
$$

The goal of this section is to achieve the following:

$$
\begin{gather*}
\hat{x}(k \mid k), P_{x}(k \mid k) ; \hat{\theta}(k \mid k), P_{\lambda}(k \mid k) ; \hat{\alpha}(k+1 \mid k+1), P_{\alpha}(k+1 \mid k+1) \\
\underset{\text { formula(4.2.2) }}{y(k+1)}  \tag{88}\\
\theta
\end{gather*}(k+1 \mid k+1), P_{\theta}(k+1 \mid k+1)
$$

Under the conditions of known $\hat{x}(k \mid k), \hat{\theta}(k \mid k), \hat{\alpha}(k+1 \mid k+1)$ and online observations $y(k+1)$, we use the description of the linearization of hidden variables $\theta(k+1)$ in Equation (85) to construct the Kalman filter to find the following:

$$
\begin{equation*}
\hat{\theta}(k+1 \mid k+1), P_{\theta}(k+1 \mid k+1) \tag{89}
\end{equation*}
$$

The Design of Linear Filter for the Estimated Value of Parameter Variable $\theta(k+1)$
It is planned to design a linearized filter estimator to solve the parameter variables $\theta(k+1)$ :

$$
\begin{align*}
\hat{\theta}(k+1 \mid k+1) & =E\{\lambda(k+1) \mid \hat{x}(k \mid k), \hat{\alpha}(k \mid k), \hat{\theta}(k \mid k), y(k+1)\} \\
& =\hat{\theta}(k+1 \mid k)+K_{\theta}(k+1)\left[y(k+1)-\hat{y}_{\theta}(k+1 \mid k)\right]  \tag{90}\\
P_{\theta}(k+1 \mid k+1) & =E\left\{[\theta(k+1)-\hat{\theta}(k+1 \mid k+1)][\theta(k+1)-\hat{\theta}(k+1 \mid k+1)]^{T}\right\}
\end{align*}
$$

where there are Equations (91)-(107) in sequence.
(1) The predicted value of the parameter variable $\theta(k+1)$ and the predicted estimation error covariance matrix:

Based on Equation (84), we can get the predicted estimated value of the parameter variable $\theta(k+1)$ as follows:

$$
\begin{equation*}
\hat{\theta}(k+1 \mid k)=S_{\theta}(k) \hat{x}(k \mid k)+G_{\theta}(k) \hat{\alpha}(k \mid k)+L_{\theta}(k) \hat{\theta}(k \mid k) \tag{91}
\end{equation*}
$$

and forecast estimated error:

$$
\begin{align*}
\tilde{\theta}(k+1 \mid k)= & \theta(k+1)-\hat{\theta}(k+1 \mid k) \\
= & S_{\theta}(k) x(k)+G_{\theta}(k) \alpha(k)+L_{\theta}(k) \theta(k)+w_{\theta}(k) \\
& -S_{\theta}(k) \hat{x}(k \mid k)-G_{\theta}(k) \hat{\alpha}(k \mid k)-L_{\theta}(k) \hat{\theta}(k \mid k)  \tag{92}\\
= & S_{\theta}(k) \widetilde{x}(k \mid k)+G_{\theta}(k) \widetilde{\alpha}(k \mid k)+L_{\theta}(k) \widetilde{\theta}(k \mid k)+w_{\theta}(k)
\end{align*}
$$

and predictive estimate error covariance matrix:

$$
\begin{align*}
P_{\theta}(k+1 \mid k)= & E\left\{\widetilde{\theta}(k+1 \mid k) \widetilde{\theta}^{T}(k+1 \mid k)\right\} \\
= & \lambda_{\theta}(k)\left[S_{\theta}(k) P_{x}(k \mid k) S_{\theta}^{T}(k)+G_{\theta}(k) P_{\alpha}(k \mid k) G_{\lambda}^{T}(k)\right.  \tag{93}\\
& \left.+L_{\theta}(k) P_{\theta}(k \mid k) L_{\theta}^{T}(k)\right]+Q_{\theta}(k)
\end{align*}
$$

(2) Prediction estimates and prediction estimation errors of measured values of parameter variables $\theta(k+1)$ :

Based on Equation (85), for the measured value of the parameter variable, we have the predicted estimated value:

$$
\begin{align*}
\hat{z}_{\theta}(k+1 \mid k)= & H_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right)  \tag{94}\\
& \times h^{(2)}(\hat{\theta}(k+1 \mid k))
\end{align*}
$$

and forecast estimation error vector:

$$
\begin{align*}
\widetilde{z}_{\theta}(k+1 \mid k) & =z(k+1)-\hat{z}_{\theta}(k+1 \mid k) \\
= & H_{\theta}\left(h^{(1)}(x(k+1)), h^{(3)} \alpha(k+1)\right) \times h^{(2)}(\theta(k+1))+v(k+1) \\
& -H_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}((\hat{\alpha}(k+1 \mid k+1))) \times h^{(2)}(\hat{\theta}(k+1 \mid k))\right. \\
\approx & H_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}((\hat{\alpha}(k+1 \mid k+1)))\right. \\
& \times\left[h^{(2)}(\lambda(k+1))-h^{(2)}(\hat{\theta}(k+1 \mid k))\right]+v(k+1) \\
\approx & H_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}((\hat{\alpha}(k+1 \mid k+1)))\right.  \tag{95}\\
& \times H_{\theta}^{(2)}(\hat{\lambda}(k+1 \mid k))[\theta(k+1)-\hat{\theta}(k+1 \mid k)]+v(k+1) \\
= & H_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}((\hat{\alpha}(k+1 \mid k+1)))\right. \\
& \times H_{\theta}^{(2)}(\hat{\theta}(k+1 \mid k)) \widetilde{\theta}(k+1 \mid k)+v_{\theta}(k+1) \\
= & \bar{H}_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}((\hat{\alpha}(k+1 \mid k+1))), \hat{\theta}(k+1 \mid k)\right) \\
& \times \widetilde{\theta}(k+1 \mid k)+v_{\theta}(k+1)
\end{align*}
$$

and forecast estimation error covariance matrix:

$$
\begin{align*}
& P_{\theta, z z}(k+1 \mid k) \\
& =E\left\{\widetilde{z}_{\theta}(k+1 \mid k) \widetilde{z}_{\theta}^{T}(k+1 \mid k)\right\} \\
& =\left(\bar{H}_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}((\hat{\alpha}(k+1 \mid k+1))), \hat{\theta}(k+1 \mid k)\right)\right) P_{\theta}(k+1 \mid k)  \tag{96}\\
& \quad \times\left(\bar{H}_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}((\hat{\alpha}(k+1 \mid k+1))), \hat{\theta}(k+1 \mid k)\right)\right)^{T}+R_{\theta}(k+1)
\end{align*}
$$

where

$$
\begin{align*}
& \bar{H}_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}((\hat{\alpha}(k+1 \mid k+1))), \hat{\theta}(k+1 \mid k)\right)  \tag{97}\\
& \quad=H_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}((\hat{\alpha}(k+1 \mid k+1))) \times H_{\theta}^{(2)}(\hat{\lambda}(k+1 \mid k))\right.
\end{align*}
$$

Remark 6. In the derivation process of (94), it is explained as follows:
(1) The first " $\approx$ " is due to the approximate approximation used.

$$
\begin{align*}
& H_{\lambda}\left(h^{(1)}(x(k+1)), h^{(3)}(\alpha(k+1))\right) \\
& \quad \approx H_{\lambda}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right) \tag{98}
\end{align*}
$$

(2) The second " $\approx$ " sign is due to the approximation of Taylor expansion.

$$
\begin{equation*}
h^{(2)}(\theta(k+1))-h^{(2)}(\hat{\theta}(k+1 \mid k)) \approx H_{\theta}^{(2)}(\hat{\theta}(k+1 \mid k)) \times \widetilde{\theta}(k+1 \mid k) \tag{99}
\end{equation*}
$$

where

$$
\begin{equation*}
H_{\theta}^{(2)}(\hat{\theta}(k+1 \mid k))=\left.\frac{\partial h^{(2)}(\theta(k+1))}{\partial \theta(k+1)}\right|_{\theta(k+1)=\hat{\theta}(k+1 \mid k)} \tag{100}
\end{equation*}
$$

(3) In the derivation process of (94), the noise term changes as follows:

$$
\begin{equation*}
v(k+1) \xrightarrow[\text { second" }{ }^{\prime \prime} \approx \prime]{\text { first" }} v_{\lambda}(k+1) \tag{101}
\end{equation*}
$$

this is caused by the introduction of new modeling errors in the two " $\approx$ ".
(4) The parameter $\lambda_{\theta}(k)$ in Equation (93) is the fading factor of Equation (17) determined by Equation (19) in Section 2.

So far, we completed the linearization process of using the strong tracking filter (STF) to solve the estimation value and the estimation error covariance matrix of parameter $\theta(k+1)$, and the corresponding STF is designed below.
(3) The filter design for real-time estimation of parameter variables:

$$
\begin{equation*}
\hat{\theta}(k+1 \mid k+1)=\hat{\theta}(k+1 \mid k)+K_{\theta}(k+1)\left[y(k+1)-\hat{y}_{\theta}(k+1 \mid k)\right] \tag{102}
\end{equation*}
$$

Calculate real-time estimates of parameter variables:

$$
\begin{align*}
& \widetilde{\theta}(k+1 \mid k+1) \\
& =\theta(k+1)-\hat{\theta}(k+1 \mid k+1) \\
& =\theta(k+1)-\hat{\theta}(k+1 \mid k)-K_{\theta}(k+1)\left[y(k+1)-\hat{y}_{\theta}(k+1 \mid k)\right] \\
& =\widetilde{\theta}(k+1 \mid k)-K_{\theta}(k+1)\left[y(k+1)-\hat{y}_{\theta}(k+1 \mid k)\right]  \tag{103}\\
& =\widetilde{\theta}(k+1 \mid k)-K_{\theta}(k+1)\left[\bar{H}_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}((\hat{\alpha}(k+1 \mid k+1))), \hat{\theta}(k+1 \mid k)\right)\right. \\
& \left.\quad \times \widetilde{\theta}(k+1 \mid k)+v_{\theta}(k+1)\right]
\end{align*}
$$

a direct sum decomposition of observations:

$$
\begin{align*}
z(k+1)= & \hat{z}_{\theta}(k+1 \mid k)+\widetilde{z}_{\theta}(k+1 \mid k) \\
= & \hat{z}_{\theta}(k+1 \mid k)+\bar{H}_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1)), \hat{\theta}(k+1 \mid k)\right)  \tag{104}\\
& \times \widetilde{\theta}(k+1 \mid k)+v_{\theta}(k+1)
\end{align*}
$$

utilize the Orthogonal Principle:

$$
\begin{equation*}
\left\{\widetilde{\theta}(k+1 \mid k+1) z^{T}(k+1)\right\}=0 \tag{105}
\end{equation*}
$$

solve the gain matrix, $K_{\theta}(k+1)$, in Equation (102):

$$
\begin{align*}
& K_{\theta}(k+1)= \\
& P_{\theta}(k+1 \mid k)\left(\bar{H}_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1)), \hat{\theta}(k+1 \mid k)\right)\right)^{T} \\
& \times\left[\left(\bar{H}_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}(\hat{\theta}(k+1 \mid k+1)), \hat{\lambda}(k+1 \mid k)\right)\right) P(k+1 \mid k)\right.  \tag{106}\\
& \left.\quad \times\left(\bar{H}_{\theta}\left(h^{(1)}(\hat{x}(k+1 \mid k)), h^{(3)}(\hat{\theta}(k+1 \mid k+1)), \hat{\lambda}(k+1 \mid k)\right)\right)^{T}+R_{\theta}(k+1)\right]^{-1}
\end{align*}
$$

(4) Calculate the real-time estimated error covariance matrix of parameter variables.

According to the definition (90), we have the following:

$$
\begin{align*}
P_{\theta}(k+1 \mid k+1)= & {\left[I-K_{\theta}(k+1)\left(\bar{H}_{\theta}(\hat{x}(k+1 \mid k), \hat{\alpha}(k+1 \mid k), \hat{\theta}(k+1 \mid k))\right)\right] }  \tag{107}\\
& \times P_{\theta}(k+1 \mid k)
\end{align*}
$$

### 4.3. The Linear Kalman Filter Design for Solving State Variables $x(k+1)$

Based on the pseudo linearization system composed of Equations (42) and (43), it is rewritten as follows:

$$
\begin{gather*}
x(k+1)=A_{x}(\theta(k), \alpha(k)) \times x(k)+w(k)  \tag{108}\\
z(k+1)=H_{x}\left(h ^ { ( 2 ) } \left(\theta(k+1), h^{(3)}(\alpha(k+1)) \times h^{(1)}\left(x(k+1)+v_{x}(k+1)\right.\right.\right. \tag{109}
\end{gather*}
$$

and suppose that the initial value of the state $x(0)$ has the following statistical characteristics:

$$
\begin{equation*}
E\{x(0)\}=\hat{x}_{0} ; E\left\{\left[x(0)-\hat{x}_{0}\right]\left[x(0)-\hat{x}_{0}\right]^{T}\right\}=P_{x 0} \tag{110}
\end{equation*}
$$

The initial values of noise $v_{x}(k+1), w_{x}(k)$ and parameters have the following characteristics:

$$
\begin{align*}
& E\left\{w_{x}(k)\right\}=0 ; E\left\{v_{x}(k+1)\right\}=0 ; E\left\{w_{x}(k) v_{x}^{T}(k+1)\right\}=0  \tag{111}\\
& E\left\{w_{x}(k) x^{T}(0)\right\}=0 ; E\left\{v_{x}(k+1) x^{T}(0)\right\}=0
\end{align*}
$$

The goal of this section is as follows:

$$
\begin{gather*}
\hat{x}(k \mid k), P_{x}(k \mid k) ; \hat{\alpha}(k+1 \mid k+1), P_{\alpha}(k+1 \mid k+1) ; \hat{\theta}(k+1 \mid k+1), P_{\theta}(k+1 \mid k+1) \\
\xrightarrow[\text { formula(4.3.3) }]{y(k+1)} \hat{x}(k+1 \mid k+1), P_{x}(k+1 \mid k+1) \tag{112}
\end{gather*}
$$

Through the obtained state, parameter variable and hidden variable estimated value and covariance matrix, we obtain the following

$$
\hat{x}(k \mid k), P_{x}(k \mid k) ; \hat{\alpha}(k+1 \mid k+1), P_{\alpha}(k+1 \mid k+1) ; \hat{\theta}(k+1 \mid k+1), P_{\theta}(k+1 \mid k+1)
$$

When linearized by Equation (109) to obtain the estimated value of the state variable $x(k+1)$ and the estimated error covariance matrix, it achieves the following:

$$
\begin{equation*}
\hat{x}(k+1 \mid k+1), P_{x}(k+1 \mid k+1) \tag{113}
\end{equation*}
$$

The goal of this section is to achieve the following:

$$
\begin{gather*}
\hat{x}(k \mid k), P_{x}(k \mid k) ; \hat{\theta}(k+1 \mid k+1), P_{\theta}(k+1 \mid k+1) ; \hat{\alpha}(k+1 \mid k+1), P_{\alpha}(k+1 \mid k+1) \\
\underset{\text { formula(4.3.2) }}{y(k+1)} \hat{x}(k+1 \mid k+1), P_{x}(k+1 \mid k+1) \tag{114}
\end{gather*}
$$

Under the condition of known $\hat{x}(k \mid k), \hat{\theta}(k+1 \mid k+1), \hat{\alpha}(k+1 \mid k+1)$ and obtained online observations, through the description of the linearization of state variables in Equation (85), construct the Kalman filter for the following:

$$
\begin{equation*}
\hat{x}(k+1 \mid k+1), P_{x}(k+1 \mid k+1) \tag{115}
\end{equation*}
$$

The Linear Filter Design of State Variable Estimates $x(k+1)$

$$
\begin{align*}
\hat{x}(k+1 \mid k+1) & =E\{x(k+1) \mid \hat{x}(k \mid k), y(k+1), \hat{\alpha}(k+1 \mid k+1), \hat{\theta}(k+1 \mid k+1)\} \\
& =\hat{x}(k+1 \mid k)+K_{x}(k+1)\left[y(k+1)-\hat{y}_{x}(k+1 \mid k)\right]  \tag{116}\\
P_{x}(k+1 \mid k+1) & =E\left\{[x(k+1)-\hat{x}(k+1 \mid k+1)][x(k+1)-\hat{x}(k+1 \mid k+1)]^{T}\right\}
\end{align*}
$$

where there are Equations (117)-(133) in sequence.
(1) The predictive estimate of the state variable $x(k+1)$ and the predictive estimate error covariance matrix

Based on Equation (108), we get the predicted estimated value of the state variable $x(k+1)$

$$
\begin{equation*}
\hat{x}(k+1 \mid k)=A_{x}(\hat{\theta}(k \mid k), \hat{\alpha}(k \mid k)) \times \hat{x}(k \mid k) \tag{117}
\end{equation*}
$$

and forecast estimation error:

$$
\begin{align*}
& \widetilde{x}(k+1 \mid k) \\
& =x(k+1)-\hat{x}(k+1 \mid k) \\
& =A_{x}(\theta(k), \alpha(k)) \times x(k)+w(k)-A_{x}(\hat{\theta}(k \mid k), \hat{\alpha}(k \mid k)) \times \hat{x}(k \mid k)  \tag{118}\\
& \approx A_{x}(\hat{\theta}(k \mid k), \hat{\alpha}(k \mid k))[x(k)-\hat{x}(k \mid k)]+w(k) \\
& =A_{x}(\hat{\theta}(k \mid k), \hat{\alpha}(k \mid k)) \widetilde{x}(k \mid k)+w_{x}(k)
\end{align*}
$$

and forecast estimation error covariance matrix:

$$
\begin{align*}
P_{x}(k+1 \mid k) & =E\left\{\widetilde{x}(k+1 \mid k) \widetilde{x}^{T}(k+1 \mid k)\right\}  \tag{119}\\
& =A_{x}(\hat{\theta}(k \mid k), \hat{\alpha}(k \mid k)) P_{x}(k \mid k) A_{x}^{T}(\hat{\theta}(k \mid k), \hat{\alpha}(k \mid k))+Q_{x}(k)
\end{align*}
$$

Remark 7. In the derivation process of (92), it is explained as follows.
(1) The first " $\approx$ " is due to the approximate approximation used.

$$
\begin{equation*}
A_{x}(\theta(k), \alpha(k)) \approx A_{x}(\hat{\theta}(k \mid k), \hat{\alpha}(k \mid k)) \tag{120}
\end{equation*}
$$

(2) In the derivation process of (92), the noise term changes as follows:

$$
\begin{equation*}
w(k) \xrightarrow{\text { first" } \approx "} w_{x}(k) \tag{121}
\end{equation*}
$$

this is caused by the introduction of new modeling errors this time.
(2) Prediction estimates and prediction estimation errors of measured values of state variables:

Based on Equation (109), for the measured variable $x(k+1)$, we have a predictive estimate:

$$
\begin{equation*}
\hat{z}_{x}(k+1 \mid k)=H_{x}\left(h^{(1)}(\hat{\theta}(k+1 \mid k+1)), h^{(2)}(\hat{\alpha}(k+1 \mid k+1)) \times \hat{x}(k+1 \mid k)\right. \tag{122}
\end{equation*}
$$

and forecast estimation error vector:

$$
\begin{align*}
& \widetilde{z}_{x}(k+1 \mid k) \\
& =z(k+1)-\hat{z}_{x}(k+1 \mid k) \\
& =H_{x}\left(h^{(2)}(\theta(k+1)), h^{(2)}(\alpha(k+1)) \times h^{(1)}(x(k+1))+v(k+1)\right. \\
& \quad-H_{x}\left(h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1)) \times h^{(1)}(\hat{x}(k+1 \mid k))\right. \\
& \approx H_{x}\left(h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right. \\
& \quad \times\left[h^{(1)}(x(k+1))-h^{(1)}(\hat{x}(k+1 \mid k))\right]+v(k+1)  \tag{123}\\
& \approx H_{x}\left(h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k))\right. \\
& \left.\quad \times H_{x}^{(1)}(\hat{x}(k+1 \mid k)) \widetilde{x}(k+1 \mid k)\right)+v(k+1) \\
& =\bar{H}_{x}\left(\hat{x}(k+1 \mid k), h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right. \\
& \quad \times \widetilde{x}(k+1 \mid k)+v_{x}(k+1)
\end{align*}
$$

and forecast estimation error covariance matrix:

$$
\begin{align*}
& P_{x, z z}(k+1 \mid k) \\
& =E\left\{\widetilde{z}_{x}(k+1 \mid k) \widetilde{z}_{x}^{T}(k+1 \mid k)\right\} \\
& =\left(\bar{H}_{x}\left(\hat{x}(k+1 \mid k), h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right) P_{x}(k+1 \mid k)\right.  \tag{124}\\
& \quad \times\left(\bar{H}_{x}\left(\hat{x}(k+1 \mid k), h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right)^{T}+R_{x}(k+1)\right.
\end{align*}
$$

where

$$
\begin{align*}
& \bar{H}_{x}\left(\hat{x}(k+1 \mid k), h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right. \\
& \quad:=H_{x}\left(h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k)) H_{x}^{(1)}(\hat{x}(k+1 \mid k))\right. \tag{125}
\end{align*}
$$

Remark 8. In the derivation process of (123), it is explained as follows:
(1) The first " $\approx$ " is due to the approximate approximation used.

$$
\begin{align*}
& H_{x}\left(h^{(1)}(\theta(k+1)), h^{(2)}(\alpha(k+1))\right.  \tag{126}\\
& \quad \approx H_{x}\left(h^{(1)}(\hat{\theta}(k+1 \mid k+1)), h^{(2)}(\hat{\alpha}(k+1 \mid k+1))\right.
\end{align*}
$$

(2) In the derivation process of (123), the noise term changes as follows:

$$
\begin{equation*}
v(k+1) \xrightarrow{\text { first" } \approx "} v_{x}(k+1) \tag{127}
\end{equation*}
$$

this is caused by the new module error introduced by " $\approx$ " twice.
(3) The filter design for real-time estimation of state variables $x(k+1)$ :

$$
\begin{equation*}
\hat{x}(k+1 \mid k+1)=\hat{x}(k+1 \mid k)+K_{x}(k+1)\left[y(k+1)-\hat{y}_{x}(k+1 \mid k)\right] \tag{128}
\end{equation*}
$$

calculate the estimated error of the state variable $x(k+1)$ :

$$
\begin{align*}
& \widetilde{x}(k+1 \mid k+1) \\
& =x(k+1)-\hat{x}(k+1 \mid k+1) \\
& =x(k+1)-\hat{x}(k+1 \mid k)-K_{\alpha}(k+1)\left[y(k+1)-\hat{y}_{x}(k+1 \mid k)\right] \\
& =\widetilde{x}(k+1 \mid k)-K_{x}(k+1)\left[y(k+1)-\hat{y}_{x}(k+1 \mid k)\right]  \tag{129}\\
& =\widetilde{x}(k+1 \mid k)-K_{x}(k+1)\left[\overline { H } _ { x } \left(\hat{x}(k+1 \mid k), h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right.\right. \\
& \left.\quad \times \widetilde{x}(k+1 \mid k)+v_{x}(k+1)\right]
\end{align*}
$$

a direct sum decomposition of observations:

$$
\begin{align*}
z(k+1)= & \hat{z}_{x}(k+1 \mid k)+\widetilde{z}_{x}(k+1 \mid k) \\
= & \hat{z}_{x}(k+1 \mid k)+\bar{H}_{x}\left(\hat{x}(k+1 \mid k), h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right.  \tag{130}\\
& \times \widetilde{x}(k+1 \mid k)+v_{x}(k+1)
\end{align*}
$$

utilize the Orthogonal Principle as follows:

$$
\begin{equation*}
\left\{\widetilde{x}(k+1 \mid k+1) z^{T}(k+1)\right\}=0 \tag{131}
\end{equation*}
$$

solve the gain matrix, $K_{x}(k+1)$, in Equation (128):

$$
\begin{align*}
& K_{x}(k+1) \\
& =P_{x}(k+1 \mid k)\left(\bar{H}_{x}\left(\hat{x}(k+1 \mid k), h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right)^{T}\right. \\
& \quad \times\left[\left(\bar{H}_{x}\left(\hat{x}(k+1 \mid k), h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right) P_{x}(k+1 \mid k)\right.\right.  \tag{132}\\
& \quad \times\left(\bar{H}_{x}\left(\hat{x}(k+1 \mid k), h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right)^{T}+R_{x}(k+1)\right]^{-1}
\end{align*}
$$

(4) Calculate state variables $x(k+1)$ and estimate the error covariance matrix in real time:

$$
\begin{align*}
& P_{x}(k+1 \mid k+1) \\
& =E\left\{[x(k+1)-\hat{x}(k+1 \mid k)][x(k+1)-\hat{x}(k+1 \mid k)]^{T}\right\} \\
& =E\left\{\widetilde{x}(k+1 \mid k) \widetilde{x}^{T}(k+1 \mid k)\right\}  \tag{133}\\
& =\left[I-K_{x}(k+1)\left(\bar{H}_{x}\left(\hat{x}(k+1 \mid k), h^{(2)}(\hat{\theta}(k+1 \mid k+1)), h^{(3)}(\hat{\alpha}(k+1 \mid k+1))\right)\right]\right. \\
& \quad \times P_{x}(k+1 \mid k)
\end{align*}
$$

## 5. Simulation

Simulation 1. Consider the following one-dimensional nonlinear system:

$$
\begin{gathered}
x(k+1)=x(k) \alpha(k) \sin (x(k))+w(k) \\
z(k+1)=x(k+1) \alpha^{3}(k+1)+v(k)
\end{gathered}
$$

where $Q=[0.001], R_{1}=[0.000001]$, and $R_{2}=[0.0002]$; the change rule of time-varying parameter is unknown, but the normal value is known $\alpha_{0}=-0.58$. The data and curves in the table are the statistical results of 100 Monte Carlo simulations, which are taken during simulation:

$$
\begin{equation*}
x(0)=0, \hat{x}(0 \mid 0)=0, \hat{\alpha}(0 \mid 0)=-0.58, P(0 \mid 0)=\operatorname{diag}[0.3,0.3] ; \tag{134}
\end{equation*}
$$

Randomly select the following parameters $\alpha(k)$ with random noise $w_{2}(k) \sim N[0,0.0001]$ :

$$
\alpha(k+1)= \begin{cases}\alpha(k)+w_{2}(k), & 0 \leq k \leq 28  \tag{135}\\ \alpha(k)+\alpha_{0} / 10+w_{2}(k), & k=29 \\ \alpha(k)+w_{2}(k), & 30 \leq k \leq 58 \\ \alpha(k)-\alpha_{0} / 500+w_{2}(k), & 59 \leq k \leq 89 \\ \alpha(k)+w_{2}(k), & k=90 \\ \alpha(k)+\alpha_{0} / 5+w_{2}(k), & 91 \leq k \leq 100\end{cases}
$$

For a one-dimensional nonlinear system, when the system contains time-varying parameters, Figure 1 and Table 2 show the filtering effects of several filtering methods. Combining the data in Table 2, it can be seen that the estimation accuracy of the algorithm in this paper has been significantly improved regardless of the estimation of the state or the estimation of the parameters.


Figure 1. Filtering effect on parameter $\alpha$ and state $x$ : (a) true value and estimated value of $\alpha$, (b) true value and estimated value of $x,(c)$ estimation error of $\alpha$ and (d) estimation error of $x$.

Table 2. Performance comparison of different algorithms.

|  | MSE | EKF | STF | UKF | TSKF | VS EKF | VS STF |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| state | $x$ | 0.0065 | 0.0063 | $3.3783 \times 10^{-4}$ | $1.6557 \times 10^{-4}$ | $97.4528 \%$ | $97.3719 \%$ | $50.9901 \%$ |
| parameter | $\alpha$ | 0.0225 | $1.4868 \times 10^{-4}$ | $1.0504 \times 10^{-4}$ | $8.7584 \times 10^{-5}$ | $99.6107 \%$ | $41.0923 \%$ | $16.6184 \%$ |

Simulation 2. Consider the following two-dimensional nonlinear system model:

$$
\begin{gathered}
x_{1}(k+1)=0.8 \alpha_{1}(k) x_{1}(k) \sin \left(\beta_{1} x_{2}(k)\right)+w_{1}(k) \\
x_{2}(k+1)=0.8 \alpha_{2}(k) x_{2}(k) \cos \left(\beta_{2} x_{1}(k)\right)+w_{2}(k) \\
z_{1}(k+1)=\gamma x_{1}^{2}(k+1)+x_{2}(k+1)+v_{1}(k+1) \\
z_{2}(k+1)=x_{1}(k+1)+v_{2}(k+1)
\end{gathered}
$$

where $k$ is the discrete time and takes the value. The parameter $\beta$ indicates the degree of nonlinearity, and the parameter $\gamma$ indicates the proportion of nonlinearity. The model of the parameter $\alpha$ is as follows:

$$
\begin{aligned}
& \alpha_{1}(k+1)=\alpha_{1}(k) \sin \left(x_{1}(k)\right) \\
& \alpha_{2}(k+1)=\alpha_{2}(k) \cos \left(x_{2}(k)\right)
\end{aligned}
$$

Set the parameters $\beta$ and $\gamma$ as follows when $40 \leq k \leq 60$ :

$$
\beta_{1}=\operatorname{rand}(1), \beta_{2}=\operatorname{rand}(1), \gamma=\operatorname{rand}(1)
$$

otherwise $\beta_{1}=\beta_{2}=\gamma=0.8$. Process noise and observation noise have the following characteristics $w(k) \sim N[0, Q(k)], v(k+1) \sim N[0, R(k+1)]$. The data and curves in the Figures 2 and 3 and Tables 3 and 4 are the statistical results of 100 Monte Carlo simulations. During simulation, $x(0)=[0.8,0.6]^{T} ; \hat{x}(0 \mid 0)=[0.8,0.6]^{T}, P(0 \mid 0)=\operatorname{diag}[1,1]$.


Figure 2. Filtering effect on parameter $x$ : (a) true value and estimated value of $x_{1}$, (b) true value and estimated value of $x_{2}$, (c) estimation error of $x_{1}$ and (d) estimation error of $x_{2}$.


Figure 3. Filtering effect on parameter $\alpha$ : (a) true value and estimated value of $\alpha_{1}$, (b) true value and estimated value of $\alpha_{2}$; (c) estimation error of $\alpha_{1}$ and (d) estimation error of $\alpha_{2}$.

Table 3. Performance comparison of different algorithms.

|  | MSE | EKF | STF | UKF | TSKF | VS EKF | VS STF | VS UKF |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| state | $x_{1}$ | 0.0091 | 0.0076 | 0.0024 | 0.0010 | $89.0110 \%$ | $86.8421 \%$ | $58.3333 \%$ |
|  | $x_{2}$ | 0.0035 | 0.0019 | $9.9575 \times 10^{-4}$ | $8.8605 \times 10^{-4}$ | $74.6843 \%$ | $53.3658 \%$ | $11.0168 \%$ |

Table 4. Performance comparison of different algorithms.

|  | MSE | EKF | STF | UKF | TSKF | VS EKF | VS STF | VS UKF |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| parameter | $\alpha_{1}$ | 0.0032 | 0.0030 | 0.0019 | 0.0011 | $65.6250 \%$ | $63.3333 \%$ | $42.1053 \%$ |
|  | $\alpha_{2}$ | 0.0037 | 0.0031 | 0.0015 | 0.0009 | $75.6757 \%$ | $70.9677 \%$ | $40.0000 \%$ |

The data in Table 3 show that the estimated mean square error of the TSKF algorithm in this paper is 0.0010 and $8.8605 \times 10^{-4}$, respectively, which are significantly improved compared to the estimation accuracy of EKF, STF and UKF. Among them, the most improved was EKF: $x_{1}=89.0110 \%$ and $x_{2}=74.6843 \%$.

In the two-dimensional system of Simulation Experiment 5.2, time-varying parameters with stronger nonlinearity are introduced to verify the effectiveness of the algorithm in the multidimensional system. Table 4 shows the estimated mean square error of the TSKF algorithm in this paper is 0.0011 and 0.0009 , respectively, which is significantly improved compared to the estimation accuracy of EKF, STF and UKF. Among them, the most improved was EKF: $\alpha_{1}=65.6250 \%$ and $\alpha_{2}=70.9677 \%$. The experimental data show that the algorithm in this paper has achieved the expected effect.

## 6. Conclusions

Aiming at a system composed of linear terms and factorizable nonlinear terms accumulation, a nonlinear Kalman filtering method based on multiplicative latent variable state and parameter stepwise estimation was established. Firstly, the decomposed nonlinear multiplier was defined as a hidden variable, and the hidden variable was regarded as a new system variable, and a dynamic correlation model between the hidden variable and the original state was established: the linear state model and linear measurement model of the state. Finally, the Three-Stage Kalman Filter was designed to estimate the hidden variables, parameters and states in turn. The Monte Carlo simulation experiment shows that, compared with the typical nonlinear filtering of EKF, UKF and STF, this method has achieved better estimation results, and the filtering accuracy has been significantly improved. Moreover, the parameter variables can be estimated very well, which provides a new solution to the problem of state and parameter fusion estimation.

Although the method in this paper has achieved good filtering results, there are still areas for improvement. For example, how to solve the problem of state and parameter estimation for nonlinear systems with complex parameters by introducing hidden variables; and how to solve the problem of parameter identification in nonlinear input and output systems with time-varying or complex parameters. These are the key points and difficulties of the next step of research.

In addition, in the follow-up research work, we can cooperate with other members of our research team to combine the method in this paper with other fields, so that it can further achieve more practical applications; for example, it can be used to solve the parameters of neural networks [28,29], or used in fault diagnosis [30,31]. For the parametersolving problem in neural network and knowledge transfer, the existing methods can not realize the adaptive updating of parameters. However, the method in this paper can solve the parameters of more general models; it has greater advantages in solving such problems.

Author Contributions: Conceptualization, C.W. and Z.L.; methodology, C.W.; software, Z.L.; validation, C.W. and Z.L.; formal analysis, C.W.; writing-original draft preparation, C.W.; writing-review and editing, Z.L.; visualization, Z.L. All authors have read and agreed to the published version of the manuscript.
Funding: National Natural Science Foundation of China key project (No. 61751304), intelligent diagnosis, prediction and maintenance of abnormal conditions of large petrochemical plants.

Institutional Review Board Statement: Not Applicable.
Informed Consent Statement: Not Applicable.
Data Availability Statement: Not Applicable.
Conflicts of Interest: The authors declare no conflict of interest.

## References

1. Zhu, S.; Liu, W.; Cui, H. Multiple resolvable groups tracking using the GLMB filter. Acta Autom. Sin. 2017, 43, 2178-2189.
2. Hu, Y.; Jin, Z.; Xue, X.; Sun, C. Fault diagnosis for networked systems by asynchronous IMM fusion filtering. Acta Autom. Sin. 2017, 43, 1329-1338.
3. Thomas, C.; Thach, N.D.; Julien, M.; Wang, Z.; Tarek, R. Zonotopic Kalman Filter-Based Interval Estimation for Discrete-Time Linear Systems With Unknown Inputs. IEEE Control Syst. Lett. 2022, 6, 806-811.
4. Shen, T.; Xue, A.; Zhou, Z. Multi-sensor Gaussian mixture PHD fusion for multi-target tracking. Acta Autom. Sin. 2017, 43, 1028-1037.
5. Liu, J.; Sun, L.; Pu, J.; Hu, Z.; Wang, Y. Cooperative Localization in a Team of Two Mobile Robots Based on Rigid Constraints. Acta Electron. Sin. 2020, 48, 1777-1785.
6. Alfakih, M.; Keche, M.; Benoudnine, H. A Kalman-filter-based fusion method for accurate urban localisation. IET Commun. 2021, 15, 653-663. [CrossRef]
7. Kalman, R.E. A new approach to linear filter and prediction. Trans. ASME Ser. D J. Basic Eng. 1960, 82, 35-45. [CrossRef]
8. Qiao, S.; Han, N.; Zhu, X.; Shu, H. A Dynamic Trajectory Prediction Algorithm Based on Kalman Filter. Acta Electron. Sin. 2018, 46, 418-423.
9. Wang, X. Power Systems Dynamic State Estimation With the Two-Step Fault Tolerant Extended Kalman Filtering. IEEE Acesses 2021, 9, 137211-137223. [CrossRef]
10. Li, H.; Zhao, S. The Optimal Distributed Kalman Filtering Fusion With Linear Equality Contraint. IEEE Acesses 2021, 9, 106283-106292. [CrossRef]
11. Wen, C.; Cheng, X.; Xu, D.; Wen, C. Filter design based on characteristic functions for one class of multi-dimensional nonlinear non-Gaussian systems. Automatica 2017, 82, 171-180. [CrossRef]
12. Wen, C.; Wang, Z.; Hu, J.; Liu, Q.; Fuad, E.A. Recursive filtering for state-saturated systems with randomly occurring nonlinearities and missing measurements. Int. J. Robust Nonlinearity Control 2018, 28, 1715-1727. [CrossRef]
13. Kalman, R.E.; Bucy, R.S. New results in linear filtering and prediction theory. J. Basic Eng. 1961, 83, 95-108. [CrossRef]
14. Sunahara, Y. An approximate method of state estimation for nonlinear dynamical systems. Fluids Eng. 1970, 11, 957-972. [CrossRef]
15. Guo, H.; Chen, H.; Xu, F.; Wang, F.; Lu, G. Implementation of EKF for vehicle velocities estimation on FPGA. IEEE Trans. Ind. Electron. 2013, 60, 3823-3835. [CrossRef]
16. Ye, S.; Daniel, B.W. Scaling the Kalman filter for large-scale traffic estimation. IEEE Trans. Control Netw. Syst. 2018, 5, 968-980.
17. Zhao, J.; Marcos, N.; Lamine, M. A robust iterated extended Kalman filter for power system dyamic state estimation. IEEE Trans. Power Syst. 2017, 32, 3205-3216. [CrossRef]
18. Julier, S.J.; Uhlmann, J.K. Unscented filtering and nonlinear estimation. Proc. IEEE 2004, 92, 401-422. [CrossRef]
19. Arasaratnam, I.; Haykin, S. Cubature Kalman filters. IEEE Trans. Autom. Control 2009, 54, 1254-1269. [CrossRef]
20. Arasaratnam, I.; Haykin, S. Square-root quadrature Kalman filtering. IEEE Trans Signal Process. 2008, 56, 2589-2593. [CrossRef]
21. Jiang, Q.; Zhang, J. Nonlinear Filtering Based Joint Estimation of Parameters in Polynomial Systems. In Proceedings of the 26th China Control and Decision Conference, Changsha, China, 31 July 2014.
22. Wen, C.; Chen, Z.; Zhou, D. Joint State and Parameter Estimation for Multisensor Nonlinear Dynamic Systems on the Basis of Strong Tracking Filter. Acta Electron. Sin. 2002, 30, 1715-1717.
23. Sun, X.; Wen, C.; Wen, T. Maximum Correntropy High-Order Extended Kalman Filter. Chin. J. Electron. 2022, 31, 190-198. [CrossRef]
24. Sun, X.; Wen, C.; Wen, T. A Novel Step-by-Step High-Order Extended Kalman Filter Design for a Class of Complex Systems with Multiple Basic Multipliers. Chin. J. Electron. 2021, 30, 313-321. [CrossRef]
25. Sun, X.; Wen, C.; Wen, T. High-Order Extended Kalman Filter Design for a Class of Complex Dynamic Systems with Polynomial Nonlinearities. Chin. J. Electron. 2021, 30, 508-515. [CrossRef]
26. Wang, Q.; Sun, X.; Wen, C. Design Method for a Higher Order Extended Kalman Filter Based on Maximum Correlation Entropy and a Taylor Network System. Sensors 2021, 21, 5864. [CrossRef] [PubMed]
27. Liu, X.; Wen, C.; Sun, X. Design Method of High-Order Kalman Filter for Strong Nonlinear System Based on Kronecker Product Transform. Sensors 2022, 22, 653. [CrossRef] [PubMed]
28. Wen, T.; Xie, G.; Cao, Y.; Cai, B. A DNN-Based Channel Model for Network Planning in Train Control Systems. IEEE Trans. Intell. Transp. Syst. 2021, 1-8. [CrossRef]
29. Kong, Y.; Ma, X.; Wen, C. A New Method of Deep Convolutional Neural Network Image Classification Based on Knowledge Transfer in Small Label Sample Environment. Sensors 2022, 22, 898. [CrossRef]
30. Ye, L.; Ma, X.; Wen, C. Rotating Machinery Fault Diagnosis Method by Combining Time-Frequency Domain Features and CNN Knowledge Transfer. Sensor 2021, 21, 8168. [CrossRef]
31. Ma, X.; Wen, C.; Wen, T. An Asynchronous and Real-time Update Paradigm of Federated Learning for Fault Diagnosis. IEEE Trans. Ind. Inform. 2021, 17, 8531-8540. [CrossRef]
