mathematics

MDPI

Article

A Hierarchical Approach for Joint Parameter and State Estimation of a Bilinear System with Autoregressive Noise

Xiao Zhang [1], Feng Ding [1,2,*], Ling Xu [1], Ahmed Alsaedi [3] and Tasawar Hayat [3]

1 Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China; xzhang13@126.com (X.Z.); lingxu0848@163.com (L.X.)
2 College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, China
3 Department of Mathematics, King Abdulaziz University, Jeddah 21589, Saudi Arabia; aalsaedi@hotmail.com (A.A.); fmgpak@gmail.com (T.H.)
* Correspondence: fding@jiangnan.edu.cn

Received: 7 March 2019; Accepted: 8 April 2019; Published: 17 April 2019

Abstract: This paper is concerned with the joint state and parameter estimation methods for a bilinear system in the state space form, which is disturbed by additive noise. In order to overcome the difficulty that the model contains the product term of the system input and states, we make use of the hierarchical identification principle to present new methods for estimating the system parameters and states interactively. The unknown states are first estimated via a bilinear state estimator on the basis of the Kalman filtering algorithm. Then, a state estimator-based recursive generalized least squares (RGLS) algorithm is formulated according to the least squares principle. To improve the parameter estimation accuracy, we introduce the data filtering technique to derive a data filtering-based two-stage RGLS algorithm. The simulation example indicates the efficiency of the proposed algorithms.

Keywords: bilinear system; hierarchical identification; parameter estimation; least squares; state estimator

1. Introduction

Nonlinear systems widely exist in practical industrial processes. Employing mathematical models of these processes becomes increasingly significant in system analysis [1–3], system control [4–6] and signal processing [7–10], so it is necessary to find efficient methods for system modeling [11–13]. Mathematical modeling methods are summarized into three categories, including mechanism modeling, experimental modeling, and their combination [14–17]. System identification is the field of getting appropriate models of dynamic systems directly from experimental data [18–21] and can be applied to many areas [22–26]. As a special class of nonlinear systems, bilinear systems are capable of approximately describing some complex dynamic processes such as a nuclear reactor, the heat transform process, and biology [27–29]. As a result, there are many active researches towards the parameter estimation problem of bilinear systems [30,31].

In the literature of bilinear system identification, Verdult and Verhaegen applied the subspace identification techniques for multi-input multi-output bilinear systems by using the separable least squares principle [32]. Larkowski et al. handled the parameter estimation problem of the diagonal bilinear errors-in-variables system and employed the bias-compensated least squares method to estimate the parameters [33]. Hizir et al. transformed a bilinear system into its equivalent linear model and used the observer Kalman filter identification algorithm to identify the bilinear system [34]. Vicario et al. extended the interaction matrices to bilinear systems to present the relationship between the system states and the measurement data and developed the intersection subspace algorithm for the bilinear system identification [35].

The recursive least squares (RLS) method is regarded as the common parameter estimation approach among many different parameter estimation techniques [36–38]. Gan et al. proposed an efficient variable projection method to solve nonlinear least squares problems based on the matrix decomposition [39]. Compared with the transfer function representation [40,41], the state space model can reflect the motion of the inner mechanism and cover state estimation [42–44]. Many estimation methods such as sequential Monte Carlo methods [45] and subspace identification methods [46] have been extensively applied in many areas. Stroud et al. proposed a Bayesian method for the joint state and parameter estimation of state space models with additive Gaussian noise using the ensemble Kalman filter [47]. Urteaga et al. presented a sequential Monte Carlo method for filtering and prediction of time-varying signals by fusing the information from candidate models instead of using model selection [48]. Martino et al. presented the parallel particle filters based on the Bayesian model averaging principle for sequential tracking and online model selection [49]. Schön et al. derived an expectation maximization algorithm under the maximum likelihood framework and gave the parameter estimates of nonlinear state space models [50]. Li and Liu derived the input-output representation of the bilinear system through eliminating the state variables and proposed the filtering-based least squares iterative algorithm [51]. However, the state variables in the system were eliminated, so the state estimation of the bilinear system was not studied.

The bilinear system involves the products of control inputs and state variables in the state equation. Thus, linear system identification methods cannot be utilized to compute the parameter estimates due to the unmeasurable states. Different from the work in [51], this paper focuses on the joint parameter and state estimation problem for a bilinear state space system with colored noise. The bilinear model can be regarded as a time-varying state space model [52]. Thus, this paper presents a bilinear state estimator based on the Kalman filtering algorithm for computing the unknown states. Then, the state estimator-based recursive generalized least squares (RGLS) identification algorithm is developed for estimating the system parameters according to the least squares principle. For the purpose of improving the parameter estimation accuracy, the data filtering technique is introduced by transforming the bilinear model into several filtered submodels with smaller dimensions. The main contributions of this paper are listed as follows.

- We present a bilinear state estimator on the basis of the Kalman filtering algorithm.
- We derive a state estimator-based RGLS algorithm for joint state and parameter estimation of bilinear state space models on the basis of the hierarchical identification principle.
- We derive a filtering-based two-stage RGLS (F-TS-RGLS) algorithm using the state estimates for improving the parameter estimation accuracy by introducing the data filtering technique.

The remainder of this paper is organized as follows. Section 2 describes the bilinear system and gives its identification model. Section 3 derives a state estimator-based RGLS identification algorithm. Section 4 presents a state estimator-based F-TS-RGLS algorithm by using the data filtering technique. Section 5 provides an illustrative example to demonstrate the effectiveness of the proposed algorithms. Finally, some conclusions are given in Section 6.

2. System Description

Some symbols are introduced for convenience. "$A =: X$" or "$X := A$" stands for "A is defined as X"; $\hat{\boldsymbol{\theta}}_t$ denotes the estimate of $\boldsymbol{\theta}$ at time t; the symbol $\boldsymbol{I}$ ($\boldsymbol{I}_n$) represents an identity matrix of an appropriate size ($n \times n$); z denotes a unit forward shift operator like $z\boldsymbol{x}_t = \boldsymbol{x}_{t+1}$ and $z^{-1}\boldsymbol{x}_t = \boldsymbol{x}_{t-1}$; the superscript T symbolizes the matrix/vector transpose.

Recently, a state observer-based multi-innovation stochastic gradient algorithm and a state observer-based recursive least squares identification algorithm were presented for a bilinear system with white noise [53]:

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}\boldsymbol{x}_t u_t + \boldsymbol{f} u_t, \\ y_t &= \boldsymbol{h}\boldsymbol{x}_t + v_t. \end{aligned}$$

Zhang et al. derived a state filtering-based hierarchical identification algorithm and a state filtering-based forgetting factor recursive least squares algorithm for a bilinear system with white noise [54]:

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}\boldsymbol{x}_t u_t + \boldsymbol{f} u_t, \\ y_t &= \boldsymbol{h}\boldsymbol{x}_t + v_t. \end{aligned}$$

On the basis of the data filtering technique, a bilinear state observer-based multi-innovation extended stochastic gradient algorithm and a data filtering-based multi-innovation extended stochastic gradient algorithm were developed for a bilinear system with moving average noise [55]:

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}\boldsymbol{x}_t u_t + \boldsymbol{f} u_t, \\ y_t &= \boldsymbol{h}\boldsymbol{x}_t + D(z) v_t. \end{aligned}$$

where $D(z) = 1 + d_1 z^{-1} + \cdots + d_{n_c} z^{-n_d}$.

Different from the work in [53–55], this paper considers a single-input single-output bilinear system with autoregressive noise:

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}\boldsymbol{x}_t u_t + \boldsymbol{f} u_t, \tag{1}$$
$$y_t = \boldsymbol{h}\boldsymbol{x}_t + w_t, \tag{2}$$

where $\boldsymbol{x}_t :=[x_{1,t}, x_{2,t}, \cdots, x_{n,t}]^{\mathrm{T}} \in \mathbb{R}^n$ is the state vector, $u_t \in \mathbb{R}$ and $y_t \in \mathbb{R}$ are the system input and output data, $w_t \in \mathbb{R}$ is the measurement noise, and $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times n}$, $\boldsymbol{f} \in \mathbb{R}^n$, and $\boldsymbol{h} \in \mathbb{R}^{1\times n}$ are the system parameters:

$$\boldsymbol{A} := \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -a_{n-1} & 0 & \cdots & 0 & 1 \\ -a_n & 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n\times n},$$
$$\boldsymbol{B} := \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_n \end{bmatrix} \in \mathbb{R}^{n\times n}, \quad \boldsymbol{b}_i \in \mathbb{R}^{1\times n}, \ \boldsymbol{f} := \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} \in \mathbb{R}^n, \quad \boldsymbol{h} := [1, 0, \cdots, 0] \in \mathbb{R}^{1\times n}.$$

Without loss of generality, assume that $u_t = 0$, $\boldsymbol{x}_t = 0$, $y_t = 0$, and $w_t = 0$ for $t \leqslant 0$.

Assumption 1. *This paper assumes that the stochastic noise $v_t \in \mathbb{R}$ is a Gaussian noise with zero mean and variance σ^2. The colored noise w_t can be a moving average process, an autoregressive process, or an autoregressive moving average process. In this paper, w_t is approximated by an autoregressive process:*

$$w_t := -c_1 w_{t-1} - \cdots - c_{n_c} w_{t-n_c} + v_t, \tag{3}$$

Define the polynomial $C(z) = 1 + c_1 z^{-1} + \cdots + c_{n_c} z^{-n_c}$. *Equation (3) can be written as* $w_t = \frac{1}{C(z)} v_t$.

Assumption 2. *The system stability is the basis of system identification. This paper assumes that the bilinear system in (1) and (2) is stable, that is to say the bounded input leads to the bounded output, and the system is observable and controllable.*

Assumption 3. *System identification contains the determination of the system dimension and parameter estimation. This paper assumes that the orders* n *and* n_c *of the system are known. The parameters* a_i, b_{ij}, f_i, *and* c_i *are to be identified from experimental data* u_t *and* y_t.

Referring to the method in [53] and from (1) and (2), we have:

$$x_{1,t} = -\sum_{i=1}^{n} a_i x_{1,t-i} + \sum_{i=1}^{n} \boldsymbol{b}_i \boldsymbol{x}_{t-i} u_{t-i} + \sum_{i=1}^{n} f_i u_{t-i}. \tag{4}$$

Define the system parameter vectors $\boldsymbol{\theta}$, $\boldsymbol{\theta}_{\mathrm{s}}$, and $\boldsymbol{\theta}_{\mathrm{n}}$ and the information vectors $\boldsymbol{\varphi}_t$, $\boldsymbol{\varphi}_{\mathrm{s},t}$, and $\boldsymbol{\varphi}_{\mathrm{n},t}$ as:

$$\begin{aligned}
\boldsymbol{\theta}: &= \begin{bmatrix} \boldsymbol{\theta}_{\mathrm{s}} \\ \boldsymbol{\theta}_{\mathrm{n}} \end{bmatrix} \in \mathbb{R}^{n^2+2n+n_c}, \\
\boldsymbol{\theta}_{\mathrm{s}}: &= [\boldsymbol{a}^{\mathrm{T}}, \boldsymbol{b}^{\mathrm{T}}, \boldsymbol{f}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{n^2+2n}, \\
\boldsymbol{a}: &= [a_1, a_2, \cdots, a_{n-1}, a_n]^{\mathrm{T}} \in \mathbb{R}^{n}, \\
\boldsymbol{b}: &= [\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_{n-1}, \boldsymbol{b}_n]^{\mathrm{T}} \in \mathbb{R}^{n^2}, \\
\boldsymbol{f}: &= [f_1, f_2, \cdots, f_{n-1}, f_n]^{\mathrm{T}} \in \mathbb{R}^{n}, \\
\boldsymbol{\theta}_{\mathrm{n}}: &= [c_1, c_2, \cdots, c_{n_c-1}, c_{n_c}]^{\mathrm{T}} \in \mathbb{R}^{n_c}, \\
\boldsymbol{\varphi}_t: &= \begin{bmatrix} \boldsymbol{\varphi}_{\mathrm{s},t} \\ \boldsymbol{\varphi}_{\mathrm{n},t} \end{bmatrix} \in \mathbb{R}^{n^2+2n+n_c}, \\
\boldsymbol{\varphi}_{\mathrm{s},t}: &= [\boldsymbol{\varphi}_{x,t}^{\mathrm{T}}, \boldsymbol{\varphi}_{xu,t}^{\mathrm{T}}, \boldsymbol{\varphi}_{u,t}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{n^2+2n}, \\
\boldsymbol{\varphi}_{x,t}: &= [-x_{1,t-1}, -x_{1,t-2}, \cdots, -x_{1,t-n}]^{\mathrm{T}} \in \mathbb{R}^{n}, \\
\boldsymbol{\varphi}_{xu,t}: &= [\boldsymbol{x}_{t-1}^{\mathrm{T}} u_{t-1}, \boldsymbol{x}_{t-2}^{\mathrm{T}} u_{t-2}, \cdots, \boldsymbol{x}_{t-n}^{\mathrm{T}} u_{t-n}]^{\mathrm{T}} \in \mathbb{R}^{n^2}, \\
\boldsymbol{\varphi}_{u,t}: &= [u_{t-1}, u_{t-2}, \cdots, u_{t-n+1}, u_{t-n}]^{\mathrm{T}} \in \mathbb{R}^{n}, \\
\boldsymbol{\varphi}_{\mathrm{n},t}: &= [-w_{t-1}, -w_{t-2}, \cdots, -w_{t-n_c}]^{\mathrm{T}} \in \mathbb{R}^{n_c}.
\end{aligned}$$

Inserting (4) into (2) gives:

$$\begin{aligned}
y_t &= x_{1,t} + w_t \\
&= \boldsymbol{\varphi}_{x,t}^{\mathrm{T}} \boldsymbol{a} + \boldsymbol{\varphi}_{xu,t}^{\mathrm{T}} \boldsymbol{b} + \boldsymbol{\varphi}_{u,t}^{\mathrm{T}} \boldsymbol{f} + w_t \\
&= \boldsymbol{\varphi}_{\mathrm{s},t}^{\mathrm{T}} \boldsymbol{\theta}_{\mathrm{s}} + w_t.
\end{aligned} \tag{5}$$

Then, Equation (3) can be written as:

$$\begin{aligned}
w_t &= -c_1 w_{t-1} - c_2 w_{t-2} - \cdots - c_{n_c} w_{t-n_c} + v_t \\
&= \boldsymbol{\varphi}_{\mathrm{n},t}^{\mathrm{T}} \boldsymbol{\theta}_{\mathrm{n}} + v_t.
\end{aligned} \tag{6}$$

Substituting (6) into (5), we obtain:

$$\begin{aligned} y_t &= x_{1,t} + w_t \\ &= \boldsymbol{\varphi}_{\mathrm{s},t}^{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{s}} + w_t \\ &= \boldsymbol{\varphi}_{\mathrm{s},t}^{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{s}} + \boldsymbol{\varphi}_{\mathrm{n},t}^{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{n}} + v_t \\ &= \boldsymbol{\varphi}_t^{\mathrm{T}}\boldsymbol{\theta} + v_t. \end{aligned} \tag{7}$$

Equation (7) is the identification model of the bilinear system in (1) and (2). The objective of this paper is to derive a bilinear state estimator to obtain the state estimates and explore an efficient parameter identification method to generate highly-accurate parameter estimates and to improve the computational efficiency by using available experimental data.

3. The RGLS Algorithm Using the Bilinear State Estimates

In this section, a state estimator-based RGLS algorithm is employed to estimate the unknown states and parameters of the considered bilinear state space system.

Define a cost function:

$$J(\boldsymbol{\theta}) := \sum_{j=1}^{t}[y_j - \boldsymbol{\varphi}_j^{\mathrm{T}}\boldsymbol{\theta}]^2.$$

Using the least squares principle and minimizing $J(\boldsymbol{\theta})$ give the recursive least squares algorithm:

$$\hat{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_{t-1} + \boldsymbol{L}_t[y_t - \boldsymbol{\varphi}_t^{\mathrm{T}}\hat{\boldsymbol{\theta}}_{t-1}], \tag{8}$$

$$\boldsymbol{L}_t = \boldsymbol{P}_{t-1}\boldsymbol{\varphi}_t[1 + \boldsymbol{\varphi}_t^{\mathrm{T}}\boldsymbol{P}_{t-1}\boldsymbol{\varphi}_t]^{-1}, \tag{9}$$

$$\boldsymbol{P}_t = [\boldsymbol{I} - \boldsymbol{L}_t\boldsymbol{\varphi}_t^{\mathrm{T}}]\boldsymbol{P}_{t-1}. \tag{10}$$

The first identification difficulty occurs because only the observation data u_t and y_t are available, but $\boldsymbol{\varphi}_t$, $\boldsymbol{\varphi}_{x,t}$ and $\boldsymbol{\varphi}_{xu,t}$ contain the unknown $\boldsymbol{x}_t$.

Remark 1. *To solve this problem, one method is to eliminate the state vector $\boldsymbol{x}_t$ and to obtain a new identification model that involves only the system input and output variables [51]. However, this cannot work because the parameter matrix $\boldsymbol{B}$ here is not the special matrix with many zero entries like in [51]. Thus, it is necessary to study new state estimation methods to obtain the state estimate.*

The second identification difficulty occurs because $\boldsymbol{\varphi}_t$ also contains the noise variable $\{w_{t-i}, i = 1, 2, \ldots, n_c\}$, so the parameter estimates $\hat{\boldsymbol{\theta}}_t$ cannot be computed by (8)–(10).

Remark 2. *The solution here is to introduce the hierarchical identification method for obtaining the parameter estimates and state estimates, respectively. Based on the auxiliary model identification idea, we replace the unknown disturbance w_{t-i} and the unknown state $\{\boldsymbol{x}_{t-i}, i = 1, 2, \ldots, n\}$ with their estimates $\hat{w}_{t-i}$ and $\hat{\boldsymbol{x}}_{t-i}$ in the identification algorithms and define the estimated information vectors $\hat{\boldsymbol{\varphi}}_t$, $\hat{\boldsymbol{\varphi}}_{\mathrm{s},t}$, $\hat{\boldsymbol{\varphi}}_{x,t}$, $\hat{\boldsymbol{\varphi}}_{xu,t}$, $\hat{\boldsymbol{\varphi}}_{\mathrm{n},t}$ of $\boldsymbol{\varphi}_t$, $\boldsymbol{\varphi}_{\mathrm{s},t}$, $\boldsymbol{\varphi}_{x,t}$, $\boldsymbol{\varphi}_{xu,t}$, $\boldsymbol{\varphi}_{\mathrm{n},t}$ as:*

$$\hat{\boldsymbol{\varphi}}_t := \begin{bmatrix} \hat{\boldsymbol{\varphi}}_{\mathrm{s},t} \\ \hat{\boldsymbol{\varphi}}_{\mathrm{n},t} \end{bmatrix} \in \mathbb{R}^{n^2+2n+n_c}, \tag{11}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{s},t} := [\hat{\boldsymbol{\varphi}}_{x,t}^{\mathrm{T}}, \hat{\boldsymbol{\varphi}}_{xu,t}^{\mathrm{T}}, \boldsymbol{\varphi}_{u,t}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{n^2+2n}, \tag{12}$$

$$\hat{\boldsymbol{\varphi}}_{x,t} := [-\hat{x}_{1,t-1}, -\hat{x}_{1,t-2}, \cdots, -\hat{x}_{1,t-n}]^{\mathrm{T}} \in \mathbb{R}^{n}, \tag{13}$$

$$\hat{\boldsymbol{\varphi}}_{xu,t} := [\hat{\boldsymbol{x}}_{t-1}^{\mathrm{T}}u_{t-1}, \hat{\boldsymbol{x}}_{t-2}^{\mathrm{T}}u_{t-2}, \cdots, \hat{\boldsymbol{x}}_{t-n}^{\mathrm{T}}u_{t-n}]^{\mathrm{T}} \in \mathbb{R}^{n^2}, \tag{14}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n},t} := [-\hat{w}_{t-1}, -\hat{w}_{t-2}, \cdots, -\hat{w}_{t-n_c}]^{\mathrm{T}} \in \mathbb{R}^{n_c}, \tag{15}$$

and the estimated parameter vectors $\hat{\boldsymbol{\theta}}_t$, $\hat{\boldsymbol{\theta}}_{\mathrm{s},t}$, $\hat{\boldsymbol{a}}_t$, $\hat{\boldsymbol{b}}_t$, $\hat{\boldsymbol{f}}_t$, *and* $\hat{\boldsymbol{\theta}}_{\mathrm{n},t}$ *of* $\boldsymbol{\theta}$, $\boldsymbol{\theta}_{\mathrm{s}}$, $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{f}$, *and* $\boldsymbol{\theta}_{\mathrm{n}}$ *as:*

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_t &:= \begin{bmatrix} \hat{\boldsymbol{\theta}}_{\mathrm{s},t} \\ \hat{\boldsymbol{\theta}}_{\mathrm{n},t} \end{bmatrix} \in \mathbb{R}^{n^2+2n+n_c}, \\
\hat{\boldsymbol{\theta}}_{\mathrm{s},t} &:= [\hat{\boldsymbol{a}}_t^{\mathrm{T}}, \hat{\boldsymbol{b}}_t^{\mathrm{T}}, \hat{\boldsymbol{f}}_t^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{n^2+2n}, \\
\hat{\boldsymbol{a}}_t &:= [\hat{a}_{1,t}, \hat{a}_{2,t}, \cdots, \hat{a}_{n,t}]^{\mathrm{T}} \in \mathbb{R}^{n}, \\
\hat{\boldsymbol{b}}_t &:= [\hat{\boldsymbol{b}}_{1,t}, \hat{\boldsymbol{b}}_{2,t}, \cdots, \hat{\boldsymbol{b}}_{n,t}]^{\mathrm{T}} \in \mathbb{R}^{n^2}, \\
\hat{\boldsymbol{f}}_t &:= [\hat{f}_{1,t}, \hat{f}_{2,t}, \cdots, \hat{f}_{n,t}]^{\mathrm{T}} \in \mathbb{R}^{n}, \\
\hat{\boldsymbol{\theta}}_{\mathrm{n},t} &:= [\hat{c}_{1,t}, \hat{c}_{2,t}, \cdots, \hat{c}_{n_c,t}]^{\mathrm{T}} \in \mathbb{R}^{n_c}.
\end{aligned}$$

According to the structure of (5), the estimate $\hat{w}_t$ can be calculated through:

$$\begin{aligned}
\hat{w}_t &= y_t - \hat{\boldsymbol{\varphi}}_{x,t}^{\mathrm{T}}\hat{\boldsymbol{a}}_t - \hat{\boldsymbol{\varphi}}_{xu,t}^{\mathrm{T}}\hat{\boldsymbol{b}}_t - \boldsymbol{\varphi}_{u,t}^{\mathrm{T}}\hat{\boldsymbol{f}}_t \\
&= y_t - \hat{\boldsymbol{\varphi}}_{\mathrm{s},t}^{\mathrm{T}}\hat{\boldsymbol{\theta}}_{\mathrm{s},t}.
\end{aligned} \tag{16}$$

According to the method in [55], we construct the following bilinear state estimator:

$$\hat{\boldsymbol{x}}_{t+1} = \boldsymbol{A}\hat{\boldsymbol{x}}_t + \boldsymbol{B}\hat{\boldsymbol{x}}_t u_t + \boldsymbol{f}u_t + \boldsymbol{L}_{x,t}[y_t - \boldsymbol{h}\hat{\boldsymbol{x}}_t - \hat{\boldsymbol{\varphi}}_{\mathrm{n},t}^{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{n}}], \tag{17}$$

$$\boldsymbol{L}_{x,t} = [\boldsymbol{A} + \boldsymbol{B}u_t]\boldsymbol{P}_{x,t}\boldsymbol{h}^{\mathrm{T}}[\boldsymbol{h}\boldsymbol{P}_{x,t}\boldsymbol{h}^{\mathrm{T}} + 1]^{-1}, \tag{18}$$

$$\boldsymbol{P}_{x,t+1} = [\boldsymbol{A} + \boldsymbol{B}u_t]\boldsymbol{P}_{x,t}[\boldsymbol{A} + \boldsymbol{B}u_t]^{\mathrm{T}} - \boldsymbol{L}_{x,t}\boldsymbol{h}\boldsymbol{P}_{x,t}[\boldsymbol{A} + \boldsymbol{B}u_t]^{\mathrm{T}}. \tag{19}$$

Remark 3. *Another problem occurs because the parameters* $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{f}$ *are unknown. Therefore, the state estimation in (17)–(19) cannot be realized. The solution is to use* $\hat{\boldsymbol{a}}_t$, $\hat{\boldsymbol{b}}_t$, *and* $\hat{\boldsymbol{f}}_t$ *to construct the parameter estimates* $\hat{\boldsymbol{A}}_t$, $\hat{\boldsymbol{B}}_t$, *and* $\hat{\boldsymbol{f}}_t$ *of* $\boldsymbol{A}$, $\boldsymbol{B}$, *and* $\boldsymbol{f}$ *as:*

$$\hat{\boldsymbol{A}}_t = \begin{bmatrix} -\hat{a}_{1,t} & 1 & 0 & \cdots & 0 \\ -\hat{a}_{2,t} & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -\hat{a}_{n-1,t} & 0 & \cdots & 0 & 1 \\ -\hat{a}_{n,t} & 0 & \cdots & 0 & 0 \end{bmatrix}, \hat{\boldsymbol{B}}_t = \begin{bmatrix} \hat{\boldsymbol{b}}_{1,t} \\ \hat{\boldsymbol{b}}_{2,t} \\ \vdots \\ \hat{\boldsymbol{b}}_{n,t} \end{bmatrix}, \quad \hat{\boldsymbol{f}}_t = \begin{bmatrix} \hat{f}_{1,t} \\ \hat{f}_{2,t} \\ \vdots \\ \hat{f}_{n,t} \end{bmatrix}. \tag{20}$$

Then, replacing $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{f}$, and $\boldsymbol{\theta}_{\mathrm{n}}$ in (17)–(19) with their estimates $\hat{\boldsymbol{A}}_t$, $\hat{\boldsymbol{B}}_t$, $\hat{\boldsymbol{f}}_t$, and $\hat{\boldsymbol{\theta}}_{\mathrm{n},t}$ obtains the bilinear state estimator:

$$\hat{\boldsymbol{x}}_{t+1} = \hat{\boldsymbol{A}}_t\hat{\boldsymbol{x}}_t + \hat{\boldsymbol{B}}_t\hat{\boldsymbol{x}}_t u_t + \hat{\boldsymbol{f}}_t u_t + \boldsymbol{L}_{x,t}[y_t - \boldsymbol{h}\hat{\boldsymbol{x}}_t - \hat{\boldsymbol{\varphi}}_{\mathrm{n},t}^{\mathrm{T}}\hat{\boldsymbol{\theta}}_{\mathrm{n},t}], \tag{21}$$

$$\boldsymbol{L}_{x,t} = [\hat{\boldsymbol{A}}_t + \hat{\boldsymbol{B}}_t u_t]\boldsymbol{P}_{x,t}\boldsymbol{h}^{\mathrm{T}}[\boldsymbol{h}\boldsymbol{P}_{x,t}\boldsymbol{h}^{\mathrm{T}} + 1]^{-1}, \tag{22}$$

$$\boldsymbol{P}_{x,t+1} = [\hat{\boldsymbol{A}}_t + \hat{\boldsymbol{B}}_t u_t]\boldsymbol{P}_{x,t}[\hat{\boldsymbol{A}}_t + \hat{\boldsymbol{B}}_t u_t]^{\mathrm{T}} - \boldsymbol{L}_{x,t}\boldsymbol{h}\boldsymbol{P}_{x,t}[\hat{\boldsymbol{A}}_t + \hat{\boldsymbol{B}}_t u_t]^{\mathrm{T}}. \tag{23}$$

Replacing the unknown $\boldsymbol{\varphi}_t$ in (8)–(10) with its estimate $\hat{\boldsymbol{\varphi}}_t$, and combining the bilinear state estimator in (20)–(23) give:

$$\hat{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_{t-1} + \boldsymbol{L}_t[y_t - \hat{\boldsymbol{\varphi}}_t^{\mathrm{T}}\hat{\boldsymbol{\theta}}_{t-1}], \tag{24}$$

$$\boldsymbol{L}_t = \boldsymbol{P}_{t-1}\hat{\boldsymbol{\varphi}}_t[1 + \hat{\boldsymbol{\varphi}}_t^{\mathrm{T}}\boldsymbol{P}_{t-1}\hat{\boldsymbol{\varphi}}_t]^{-1}, \tag{25}$$

$$\boldsymbol{P}_t = [\boldsymbol{I} - \boldsymbol{L}_t\hat{\boldsymbol{\varphi}}_t^{\mathrm{T}}]\boldsymbol{P}_{t-1}. \tag{26}$$

Equations (11)–(26) form the recursive generalized least squares (RGLS) algorithm based on the state estimates.

Remark 4. *In order to improve the RGLS parameter estimation accuracy, the next section will construct a linear filter to filter the experimental data for estimating the system parameters.*

4. The F-TS-RGLS Algorithm Using the Bilinear State Estimator

For the purpose of improving the parameter estimation accuracy of the RGLS algorithm, this section presents an F-TS-RGLS algorithm for the bilinear system by constructing a linear filter $C(z)$ to filter the observation data of the system and to decompose the identification model in (7) into two sub-identification models. Multiplying both sides of (1) and (2) by $C(z)$ obtains:

$$\begin{aligned} C(z)\boldsymbol{x}_{t+1} &= C(z)\boldsymbol{A}\boldsymbol{x}_t + C(z)\boldsymbol{B}\boldsymbol{x}_t u_t + C(z)\boldsymbol{f}u_t, \quad (27)\\ C(z)y_t &= C(z)\boldsymbol{h}\boldsymbol{x}_t + v_t. \quad (28)\end{aligned}$$

Define the filtered input $u_{\mathrm{f},t}$, the filtered output $y_{\mathrm{f},t}$, and the filtered state $\boldsymbol{x}_{\mathrm{f},t}$ as:

$$\begin{aligned} u_{\mathrm{f},t} &:= C(z)u_t = u_t + \sum_{i=1}^{n_c} c_i u_{t-i}, \quad (29)\\ y_{\mathrm{f},t} &:= C(z)y_t = y_t + \sum_{i=1}^{n_c} c_i y_{t-i}, \quad (30)\\ \boldsymbol{x}_{\mathrm{f},t} &:= C(z)\boldsymbol{x}_t = \boldsymbol{x}_t + \sum_{i=1}^{n_c} c_i \boldsymbol{x}_{t-i}. \quad (31)\end{aligned}$$

Equations (27) and (28) can be transformed as:

$$\begin{aligned} \boldsymbol{x}_{\mathrm{f},t+1} &= \boldsymbol{A}\boldsymbol{x}_{\mathrm{f},t} + \boldsymbol{B}\boldsymbol{x}_{\mathrm{f},t}u_t + \boldsymbol{f}u_{\mathrm{f},t}, \quad (32)\\ y_{\mathrm{f},t} &= \boldsymbol{h}\boldsymbol{x}_{\mathrm{f},t} + v_t, \quad (33)\end{aligned}$$

where $\boldsymbol{x}_{\mathrm{f},t} := [x_{1\mathrm{f},t}, x_{2\mathrm{f},t}, \cdots, x_{(\mathrm{n}-1)\mathrm{f},t}, x_{\mathrm{nf},t}]^{\mathrm{T}} \in \mathbb{R}^n$. From (32) and (33), we have:

$$x_{1\mathrm{f},t} = -\sum_{i=1}^{n} a_i x_{1\mathrm{f},t-i} + \sum_{i=1}^{n} \boldsymbol{b}_i \boldsymbol{x}_{\mathrm{f},t-i} u_{t-i} + \sum_{i=1}^{n} f_i u_{\mathrm{f},t-i} = \boldsymbol{\varphi}_{\mathrm{f},t}^{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{s}}, \quad (34)$$

where the filtered information vector:

$$\begin{aligned} \boldsymbol{\varphi}_{\mathrm{f},t} := [& -x_{1\mathrm{f},t-1}, -x_{1\mathrm{f},t-2}, \cdots, -x_{1\mathrm{f},t-n+1}, -x_{1\mathrm{f},t-n},\\ & \boldsymbol{x}_{\mathrm{f},t-1}^{\mathrm{T}}u_{t-1}, \boldsymbol{x}_{\mathrm{f},t-2}^{\mathrm{T}}u_{t-2}, \cdots, \boldsymbol{x}_{\mathrm{f},t-n}^{\mathrm{T}}u_{t-n}, u_{\mathrm{f},t-1}, \quad (35)\\ & u_{\mathrm{f},t-2}, \cdots, u_{\mathrm{f},t-n+1}, u_{\mathrm{f},t-n}]^{\mathrm{T}} \in \mathbb{R}^{n^2+2n}.\end{aligned}$$

Equations (33) and (6) can be expressed as the following two filtered sub-identification models:

$$\begin{aligned} y_{\mathrm{f},t} &= \boldsymbol{\varphi}_{\mathrm{f},t}^{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{s}} + v_t, \quad (36)\\ w_t &= \boldsymbol{\varphi}_{\mathrm{n},t}^{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{n}} + v_t. \quad (37)\end{aligned}$$

On the basis of the two sub-identification models in (36) and (37), define two cost functions:

$$\begin{aligned} J_1(\boldsymbol{\theta}_{\mathrm{s}}) &:= \textstyle\sum_{j=1}^{t}[y_{\mathrm{f},j} - \boldsymbol{\varphi}_{\mathrm{f},j}^{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{s}}]^2,\\ J_2(\boldsymbol{\theta}_{\mathrm{n}}) &:= \textstyle\sum_{j=1}^{t}[w_j - \boldsymbol{\varphi}_{\mathrm{n},j}^{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{n}}]^2. \end{aligned} \quad (38)$$

Minimizing $J_1(\boldsymbol{\theta}_{\mathrm{s}})$ and $J_2(\boldsymbol{\theta}_{\mathrm{n}})$ gives the following recursive relations:

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_{\mathrm{s},t} &= \hat{\boldsymbol{\theta}}_{\mathrm{s},t-1} + \boldsymbol{L}_{\mathrm{s},t}[y_{\mathrm{f},t} - \boldsymbol{\varphi}_{\mathrm{f},t}^{\mathrm{T}}\hat{\boldsymbol{\theta}}_{\mathrm{s},t-1}], && (39)\\
\boldsymbol{L}_{\mathrm{s},t} &= \boldsymbol{P}_{\mathrm{s},t-1}\boldsymbol{\varphi}_{\mathrm{f},t}[1 + \boldsymbol{\varphi}_{\mathrm{f},t}^{\mathrm{T}}\boldsymbol{P}_{\mathrm{s},t-1}\boldsymbol{\varphi}_{\mathrm{f},t}]^{-1}, && (40)\\
\boldsymbol{P}_{\mathrm{s},t} &= [\boldsymbol{I} - \boldsymbol{L}_{\mathrm{s},t}\boldsymbol{\varphi}_{\mathrm{f},t}^{\mathrm{T}}]\boldsymbol{P}_{\mathrm{s},t-1}, && (41)\\
\hat{\boldsymbol{\theta}}_{\mathrm{n},t} &= \hat{\boldsymbol{\theta}}_{\mathrm{n},t-1} + \boldsymbol{L}_{\mathrm{n},t}[w_t - \boldsymbol{\varphi}_{\mathrm{n},t}^{\mathrm{T}}\hat{\boldsymbol{\theta}}_{\mathrm{n},t-1}], && (42)\\
\boldsymbol{L}_{\mathrm{n},t} &= \boldsymbol{P}_{\mathrm{n},t-1}\boldsymbol{\varphi}_{\mathrm{n},t}[1 + \boldsymbol{\varphi}_{\mathrm{n},t}^{\mathrm{T}}\boldsymbol{P}_{\mathrm{n},t-1}\boldsymbol{\varphi}_{\mathrm{n},t}]^{-1}, && (43)\\
\boldsymbol{P}_{\mathrm{n},t} &= [\boldsymbol{I} - \boldsymbol{L}_{\mathrm{n},t}\boldsymbol{\varphi}_{\mathrm{n},t}^{\mathrm{T}}]\boldsymbol{P}_{\mathrm{n},t-1}. && (44)
\end{aligned}$$

However, Equations (39)–(44) cannot generate the parameter estimates $\hat{\boldsymbol{\theta}}_{\mathrm{s},t}$ and $\hat{\boldsymbol{\theta}}_{\mathrm{n},t}$. Because the noise term w_t and $\boldsymbol{\varphi}_{\mathrm{n},t}$ are unmeasurable, the system state $\boldsymbol{x}_t$ in the system information vector $\boldsymbol{\varphi}_{\mathrm{s},t}$ is unmeasurable. In addition, $C(z)$ is unknown, then $y_{\mathrm{f},t}$, $u_{\mathrm{f},t}$, $\boldsymbol{x}_{\mathrm{f},t}$, and $\boldsymbol{\varphi}_{\mathrm{f},t}$ cannot be obtained.

Remark 5. *The difficulty is overcome by utilizing the auxiliary model identification idea to take the place of the unknown variables with their estimates. Use the estimate $\hat{w}_{t-i}$ of w_{t-i} to construct the estimate of $\boldsymbol{\varphi}_{\mathrm{n},t}$:*

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n},t} = [-\hat{w}_{t-1}, -\hat{w}_{t-2}, \cdots, -\hat{w}_{t-n_c}]^{\mathrm{T}}. \tag{45}$$

From (5), replacing $\boldsymbol{\varphi}_{\mathrm{s},t}$ and $\boldsymbol{\theta}_{\mathrm{s}}$ with their corresponding estimates $\hat{\boldsymbol{\varphi}}_{\mathrm{s},t}$ and $\hat{\boldsymbol{\theta}}_{\mathrm{s},t-1}$ obtains the estimate of w_t:

$$\hat{w}_t = y_t - \hat{\boldsymbol{\varphi}}_{\mathrm{s},t}^{\mathrm{T}}\hat{\boldsymbol{\theta}}_{\mathrm{s},t-1}, \tag{46}$$

where:

$$\begin{aligned}
\hat{\boldsymbol{\varphi}}_{\mathrm{s},t} &= [\hat{\boldsymbol{\varphi}}_{x,t}^{\mathrm{T}}, \hat{\boldsymbol{\varphi}}_{xu,t}^{\mathrm{T}}, \boldsymbol{\varphi}_{u,t}^{\mathrm{T}}]^{\mathrm{T}}, && (47)\\
\hat{\boldsymbol{\varphi}}_{x,t} &= [-\hat{x}_{1,t-1}, \cdots, -\hat{x}_{1,t-n}]^{\mathrm{T}}, && (48)\\
\hat{\boldsymbol{\varphi}}_{xu,t} &= [\hat{\boldsymbol{x}}_{t-1}^{\mathrm{T}}u_{t-1}, \hat{\boldsymbol{x}}_{t-2}^{\mathrm{T}}u_{t-2}, \cdots, \hat{\boldsymbol{x}}_{t-n}^{\mathrm{T}}u_{t-n}]^{\mathrm{T}}, && (49)\\
\boldsymbol{\varphi}_{u,t} &= [u_{t-1}, u_{t-2}, \cdots, u_{t-n}]^{\mathrm{T}}, && (50)\\
\hat{\boldsymbol{\theta}}_{\mathrm{s},t} &= [\hat{\boldsymbol{a}}_t^{\mathrm{T}}, \hat{\boldsymbol{b}}_t^{\mathrm{T}}, \hat{\boldsymbol{f}}_t^{\mathrm{T}}]^{\mathrm{T}}. && (51)
\end{aligned}$$

By virtue of the Kalman filter, a state estimator in the bilinear form can be established for generating the state estimates. Then, utilizing the parameter estimates:

$$\hat{\boldsymbol{\theta}}_{\mathrm{n},t} = [\hat{c}_{1,t}, \hat{c}_{2,t}, \cdots, \hat{c}_{n_c,t}] \tag{52}$$

to construct the estimates of $C(z)$ gives:

$$\hat{C}(t,z) := 1 + \hat{c}_{1,t}z^{-1} + \hat{c}_{2,t}z^{-2} + \cdots + \hat{c}_{n_c,t}z^{-n_c}.$$

Since the linear filter is unknown, we use its estimate to filter the system input u_t, the system output y_t, and the system state $\boldsymbol{x}_t$ to give the filtered estimates of $u_{\mathrm{f},t}$, $y_{\mathrm{f},t}$, and $\boldsymbol{x}_{\mathrm{f},t}$:

$$\begin{aligned}
\hat{u}_{\mathrm{f},t} &= \hat{C}(t,z)u_t = u_t + \sum_{i=1}^{n_c}\hat{c}_{i,t}u_{t-i}, && (53)\\
\hat{y}_{\mathrm{f},t} &= \hat{C}(t,z)y_t = y_t + \sum_{i=1}^{n_c}\hat{c}_{i,t}y_{t-i}, && (54)\\
\hat{\boldsymbol{x}}_{\mathrm{f},t} &= \hat{\boldsymbol{x}}_t + \hat{C}(t,z)\boldsymbol{x}_t = \hat{\boldsymbol{x}}_t + \sum_{i=1}^{n_c}\hat{c}_{i,t}\hat{\boldsymbol{x}}_{t-i}. && (55)
\end{aligned}$$

Referring to the derivation of (21)–(23), we can obtain the following state estimator for obtaining the estimate of the filtered state $\boldsymbol{x}_{\mathrm{f},t}$:

$$
\begin{aligned}
\hat{\boldsymbol{x}}_{\mathrm{f},t+1} &= \hat{\boldsymbol{A}}_t\hat{\boldsymbol{x}}_{\mathrm{f},t} + \hat{\boldsymbol{B}}_t\hat{\boldsymbol{x}}_{\mathrm{f},t}u_t + \hat{\boldsymbol{f}}_t u_{\mathrm{f},t} + \boldsymbol{L}_{\mathrm{f},t}[y_{\mathrm{f},t} - \boldsymbol{h}\hat{\boldsymbol{x}}_{\mathrm{f},t} - \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{n},t}\hat{\boldsymbol{\theta}}_{\mathrm{n},t}], && (56)\\
\boldsymbol{L}_{\mathrm{f},t} &= [\hat{\boldsymbol{A}}_t + \hat{\boldsymbol{B}}_t u_{\mathrm{f},t}]\boldsymbol{P}_{\mathrm{f},t}\boldsymbol{h}^{\mathrm{T}}[\boldsymbol{h}\boldsymbol{P}_{\mathrm{f},t}\boldsymbol{h}^{\mathrm{T}} + 1]^{-1}, && (57)\\
\boldsymbol{P}_{\mathrm{f},t+1} &= [\hat{\boldsymbol{A}}_t + \hat{\boldsymbol{B}}_t u_{\mathrm{f},t}]\boldsymbol{P}_{\mathrm{f},t}[\hat{\boldsymbol{A}}_t + \hat{\boldsymbol{B}}_t u_{\mathrm{f},t}]^{\mathrm{T}} - \boldsymbol{L}_{\mathrm{f},t}\boldsymbol{h}\boldsymbol{P}_{\mathrm{f},t}[\hat{\boldsymbol{A}}_t + \hat{\boldsymbol{B}}_t u_{\mathrm{f},t}]^{\mathrm{T}}, && (58)\\
\hat{\boldsymbol{A}}_t &= \begin{bmatrix} -\hat{a}_{1,t} & 1 & 0 & \cdots & 0\\ -\hat{a}_{2,t} & 0 & 1 & \ddots & 0\\ \vdots & \vdots & \ddots & \ddots & \vdots\\ -\hat{a}_{n-1,t} & 0 & \cdots & 0 & 1\\ -\hat{a}_{n,t} & 0 & \cdots & 0 & 0 \end{bmatrix},\ \hat{\boldsymbol{B}}_t = \begin{bmatrix} \hat{\boldsymbol{b}}_{1,t}\\ \hat{\boldsymbol{b}}_{2,t}\\ \vdots\\ \hat{\boldsymbol{b}}_{n,t}\end{bmatrix},\ \hat{\boldsymbol{f}}_t = \begin{bmatrix} \hat{f}_{1,t}\\ \hat{f}_{2,t}\\ \vdots\\ \hat{f}_{n,t}\end{bmatrix}. && (59)
\end{aligned}
$$

By replacing $x_{1\mathrm{f},t-i}$, $\boldsymbol{x}_{\mathrm{f},t-i}$ and $u_{\mathrm{f},t-i}$ in (35) with their estimates $\hat{x}_{1\mathrm{f},t-i}$, $\hat{\boldsymbol{x}}_{\mathrm{f},t-i}$ and $\hat{u}_{\mathrm{f},t-i}$, define the estimated filtered information vector:

$$
\begin{aligned}
\hat{\boldsymbol{\varphi}}_{\mathrm{f},t} := [& -\hat{x}_{1\mathrm{f},t-1}, -\hat{x}_{1\mathrm{f},t-2}, \cdots, -\hat{x}_{1\mathrm{f},t-n+1}, -\hat{x}_{1\mathrm{f},t-n},\\
& \hat{\boldsymbol{x}}^{\mathrm{T}}_{\mathrm{f},t-1}u_{t-1}, \hat{\boldsymbol{x}}^{\mathrm{T}}_{\mathrm{f},t-2}u_{t-2}, \cdots, \hat{\boldsymbol{x}}^{\mathrm{T}}_{\mathrm{f},t-n}u_{t-n}, \hat{u}_{\mathrm{f},t-1},\\
& \hat{u}_{\mathrm{f},t-2}, \cdots, \hat{u}_{\mathrm{f},t-n+1}, \hat{u}_{\mathrm{f},t-n}]^{\mathrm{T}} \in \mathbb{R}^{n^2+2n}. \qquad (60)
\end{aligned}
$$

Similarly, replacing the unmeasurable information vector $\boldsymbol{\varphi}_{\mathrm{f},t}$ and $\boldsymbol{\varphi}_{\mathrm{n},t}$ in (39)–(44) with $\hat{\boldsymbol{\varphi}}_{\mathrm{f},t}$ and $\hat{\boldsymbol{\varphi}}_{\mathrm{n},t}$ obtains:

$$
\begin{aligned}
\hat{\boldsymbol{\theta}}_{\mathrm{s},t} &= \hat{\boldsymbol{\theta}}_{\mathrm{s},t-1} + \boldsymbol{L}_{\mathrm{s},t}[\hat{y}_{\mathrm{f},t} - \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{f},t}\hat{\boldsymbol{\theta}}_{\mathrm{s},t-1}], && (61)\\
\boldsymbol{L}_{\mathrm{s},t} &= \boldsymbol{P}_{\mathrm{s},t-1}\hat{\boldsymbol{\varphi}}_{\mathrm{f},t}[1 + \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{f},t}\boldsymbol{P}_{\mathrm{s},t-1}\hat{\boldsymbol{\varphi}}_{\mathrm{f},t}]^{-1}, && (62)\\
\boldsymbol{P}_{\mathrm{s},t} &= [\boldsymbol{I} - \boldsymbol{L}_{\mathrm{s},t}\hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{f},t}]\boldsymbol{P}_{\mathrm{s},t-1}, && (63)\\
\hat{\boldsymbol{\theta}}_{\mathrm{n},t} &= \hat{\boldsymbol{\theta}}_{\mathrm{n},t-1} + \boldsymbol{L}_{\mathrm{n},t}[\hat{w}_t - \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{n},t}\hat{\boldsymbol{\theta}}_{\mathrm{n},t-1}], && (64)\\
\boldsymbol{L}_{\mathrm{n},t} &= \boldsymbol{P}_{\mathrm{n},t-1}\hat{\boldsymbol{\varphi}}_{\mathrm{n},t}[1 + \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{n},t}\boldsymbol{P}_{\mathrm{n},t-1}\hat{\boldsymbol{\varphi}}_{\mathrm{n},t}]^{-1}, && (65)\\
\boldsymbol{P}_{\mathrm{n},t} &= [\boldsymbol{I} - \boldsymbol{L}_{\mathrm{n},t}\hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{n},t}]\boldsymbol{P}_{\mathrm{n},t-1}. && (66)
\end{aligned}
$$

The procedures of calculating $\hat{\boldsymbol{\theta}}_{\mathrm{s},t}$ and $\hat{\boldsymbol{\theta}}_{\mathrm{n},t}$ by the filtering-based recursive generalized least squares algorithm in (45)–(66) are listed as follows.

1. To initialize: let $t = 1$, $\hat{\boldsymbol{\theta}}_{\mathrm{n},0} = \mathbf{1}_{n_c}/p_0$, $\hat{\boldsymbol{\theta}}_{\mathrm{s},0} = \mathbf{1}_{n^2+2n}/p_0$, $\hat{y}_{\mathrm{f},t-i} = 1/p_0$, $\hat{u}_{\mathrm{f},t-i} = 1/p_0$, $\hat{\boldsymbol{x}}_{\mathrm{f},t+1-i} = \mathbf{1}_n/p_0$, $\hat{\boldsymbol{x}}_{t-i} = \mathbf{1}_n/p_0$, $\hat{w}_{t-j} = 1/p_0$, $\boldsymbol{P}_{\mathrm{s},0} = p_0\boldsymbol{I}_{n^2+2n}$, $\boldsymbol{P}_{\mathrm{n},0} = p_0\boldsymbol{I}_{n_c}$, $\boldsymbol{P}_{\mathrm{f},1} = \boldsymbol{I}_n$, $p_0 = 10^6$, $i = 1, 2, \cdots, n$, $j = 1, 2, \cdots, n_c$.
2. Collect the measurement data u_t and y_t. Form $\hat{\boldsymbol{\varphi}}_{\mathrm{n},t}$ using (45) and $\hat{\boldsymbol{\varphi}}_{\mathrm{s},t}$ using (47)–(49).
3. Compute $\hat{w}_t$ using (46), the gain vector $\boldsymbol{L}_{\mathrm{n},t}$ using (65), and the covariance matrix $\boldsymbol{P}_{\mathrm{n},t}$ using (66).
4. Update the parameter estimates $\hat{\boldsymbol{\theta}}_{\mathrm{n},t}$ using (64). Read $\hat{c}_{j,t}$ from $\hat{\boldsymbol{\theta}}_{\mathrm{n},t}$ in (52).
5. Compute $\hat{y}_{\mathrm{f},t}$ and $\hat{u}_{\mathrm{f},t}$ from (53) and (54). Construct $\hat{\boldsymbol{\varphi}}_{\mathrm{f},t}$ using (60).
6. Compute the gain vector $\boldsymbol{L}_{\mathrm{s},t}$ using (62) and the covariance matrix $\boldsymbol{P}_{\mathrm{s},t}$ using (63). Update the parameter estimates $\hat{\boldsymbol{\theta}}_{\mathrm{s},t}$ using (61).
7. Read $\hat{a}_{i,t}$, $\hat{\boldsymbol{b}}_{i,t}$ and $\hat{f}_{i,t}$ from $\hat{\boldsymbol{\theta}}_{\mathrm{s},t}$ in (51). Construct $\hat{\boldsymbol{A}}_t$, $\hat{\boldsymbol{B}}_t$, $\hat{\boldsymbol{f}}_t$ using (59). Compute $\boldsymbol{L}_{\mathrm{f},t}$ and $\boldsymbol{P}_{\mathrm{f},t+1}$ using (57) and (58).
8. Compute $\hat{\boldsymbol{x}}_{\mathrm{f},t+1}$ and $\hat{\boldsymbol{x}}_t$ using (56) and (55).
9. Increase t by one, and turn to Step 2.

Remark 6. *By using the hierarchical identification principle, the unknown parameters and states can be estimated interactively.*

Remark 7. *By introducing a linear filter, the original identification model is decomposed into two filtered sub-identification models. Then, the system parameters and the noise parameters are estimated interactively.*

Remark 8. *For identification methods, the large dimensions of the parameter vector* $\boldsymbol{\theta}$ *cause a heavy computational cost. Thus, this paper uses the hierarchical identification principle for improving the computational efficiency of the proposed methods based on the decomposition.*

Remark 9. *The number of the multiplication and addition operations is used to measure the computational cost of an algorithm. The computational cost of the RGLS and F-TS-RGLS algorithms is shown in Table 1.*

Table 1. The computation efficiency of the recursive least squares (RLS) and filtering-based two-stage RGLS (F-TS-RGLS) algorithms.

Algorithms	Number of Multiplications	Number of Additions	Total Flop
RGLS	$2(n^2+2n+n_c)^2+$ $4(n^2+2n+n_c)+n^2+2n$	$2(n^2+2n+n_c)^2+$ $2(n^2+2n+n_c)+n^2+2n$	$N_1:=4(n^2+2n+n_c)^2+$ $6(n^2+2n+n_c)+2n^2+4n$
F-TS-RGLS	$2(n^2+2n)^2+5n^2+$ $10n+2n_c^2+6n_c$	$2(n^2+2n)^2+3n^2+$ $6n+2n_c^2+4n_c$	$N_2:=4(n^2+2n)^2+8n^2+$ $16n+4n_c^2+10n_c$

The difference of the computational cost of the RGLS and the F-TS-RGLS algorithms is:

$$\begin{aligned} N_1-N_2 &= 4(n^2+2n+n_c)^2+6(n^2+2n+n_c)+2n^2 \\ &\quad +4n-[4(n^2+2n)^2+8n^2+16n+4n_c^2+10n_c] \\ &= (8n^2+2n-4)n_c. \end{aligned}$$

Remark 10. *Since* $n \geqslant 1$ *and* $n_c \geqslant 1$*, the difference* N_1-N_2 *is positive. That is to say, the computational cost of the F-TS-RGLS algorithm is less than that of the RGLS algorithm.*

5. Numerical Example

This section provides an example to illustrate the effectiveness of the proposed algorithms. Here, we consider a bilinear state space model:

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \begin{bmatrix} 0.46 & 1 \\ -0.30 & 0 \end{bmatrix} \boldsymbol{x}_t + \begin{bmatrix} 0.08 & 0.17 \\ -0.07 & -0.22 \end{bmatrix} \boldsymbol{x}_t u_t + \begin{bmatrix} 0.40 \\ 0.15 \end{bmatrix} u_t, \\ y_t &= [1,\ 0]\boldsymbol{x}_t + w_t, \\ w_t &= -cw_{t-1} + v_t. \end{aligned}$$

The parameter vector to be identified is:

$$\begin{aligned} \boldsymbol{\theta} &= [a_1, a_2, b_{11}, b_{12}, b_{21}, b_{22}, f_1, f_2, c]^{\mathrm{T}} \\ &= [-0.46, 0.30, 0.08, 0.17, -0.07, -0.22, 0.40, 0.15, 0.08]^{\mathrm{T}}. \end{aligned}$$

In the simulation, the input signal $\{u_t\}$ was chosen as a pseudo-random binary sequence with zero mean and unit variance and $\{v_t\}$ as a white noise sequence with zero mean and variances σ^2. Set the data length $L=3000$, and employ the proposed algorithms to estimate the parameters and states of this example system. The simulation input and output data are shown in Figure 1, which indicate that the bounded input signal leads to the bounded output. The parameter estimates and their estimation errors of the RGLS algorithm and the F-TS-RGLS algorithm are shown in Tables 2 and 3. The RGLS estimation error δ versus t is shown in Figure 2 with $\sigma^2=0.10^2$ and $\sigma^2=0.15^2$. The F-TS-RGLS estimation error versus t is shown in Figure 3 with different variances. The predicted outputs and the true outputs are shown in Figures 4 and 5. The states $x_{i,t}$ and their estimates $\hat{x}_{i,t}$ versus t are shown in Figure 6.

Table 2. The parameter estimates and their errors with $\sigma^2 = 0.15^2$.

Algorithms	t	a_1	a_2	b_{11}	b_{12}	b_{21}	b_{22}	f_1	f_2	c	$\delta(\%)$
RGLS	100	−0.00316	0.05268	0.07544	0.05230	−0.06414	−0.05137	0.42414	0.34337	0.09662	77.79370
	200	−0.02327	0.12465	0.06667	0.02426	−0.06804	−0.03437	0.41314	0.31430	−0.00151	73.30282
	500	−0.16392	0.16651	0.03428	0.08494	−0.05452	−0.03641	0.42364	0.26468	0.02093	53.53272
	1000	−0.26324	0.20629	0.03025	0.11604	−0.04730	−0.07382	0.41937	0.22896	0.01985	38.31102
	2000	−0.36424	0.23406	0.03468	0.13072	−0.04674	−0.10435	0.41210	0.19167	0.03129	24.68460
	3000	−0.40788	0.25214	0.04352	0.12255	−0.04747	−0.12509	0.40931	0.18628	0.05481	18.65007
F-TS-RGLS	100	−0.17207	0.13225	0.08149	0.04762	−0.08050	−0.06387	0.38554	0.24812	−0.24016	67.37094
	200	−0.26379	0.21163	0.05110	0.10094	−0.07089	−0.08449	0.39815	0.22078	−0.20425	51.91655
	500	−0.41609	0.27122	0.04448	0.17605	−0.04003	−0.19822	0.40894	0.15549	−0.06009	20.84531
	1000	−0.44506	0.28907	0.05723	0.16636	−0.05408	−0.19663	0.40519	0.15223	−0.00605	12.54315
	2000	−0.45785	0.28984	0.06293	0.16491	−0.06176	−0.19095	0.40193	0.14834	0.02730	8.43565
	3000	−0.45744	0.29468	0.06717	0.15764	−0.06434	−0.19611	0.40149	0.14983	0.05888	4.92240
True values		−0.46000	0.30000	0.08000	0.17000	−0.07000	−0.22000	0.40000	0.15000	0.08000	

Table 3. The parameter estimates and their errors with $\sigma^2 = 0.10^2$.

Algorithms	t	a_1	a_2	b_{11}	b_{12}	b_{21}	b_{22}	f_1	f_2	c	$\delta(\%)$
RGLS	100	−0.01344	0.08368	0.04929	0.12366	−0.04534	−0.08160	0.42729	0.30373	0.01047	71.76858
	200	−0.04819	0.15699	0.03634	0.12767	−0.03469	−0.10328	0.41799	0.25865	−0.09361	65.81981
	500	−0.21525	0.21335	0.02114	0.18346	−0.02157	−0.15342	0.42234	0.19440	−0.08012	42.77206
	1000	−0.31044	0.24492	0.03458	0.18170	−0.03375	−0.16150	0.41530	0.17384	−0.06081	30.24311
	2000	−0.39697	0.26064	0.04731	0.17039	−0.04645	−0.16265	0.40883	0.15616	−0.02604	19.40043
	3000	−0.42738	0.27001	0.05540	0.15703	−0.05159	−0.16863	0.40674	0.15902	0.00337	14.22233
F-TS-RGLS	100	−0.17072	0.12873	0.08786	0.04669	−0.08179	−0.08320	0.39734	0.23700	−0.08772	56.19069
	200	−0.26673	0.21478	0.05515	0.09642	−0.07483	−0.09063	0.40715	0.20974	−0.07335	40.40498
	500	−0.45453	0.28197	0.04590	0.18588	−0.04140	−0.21202	0.41195	0.13243	0.04761	8.46768
	1000	−0.46578	0.29562	0.05933	0.17264	−0.05829	−0.20062	0.40718	0.14093	0.06403	4.89664
	2000	−0.47371	0.29198	0.06315	0.17165	−0.06534	−0.19084	0.40294	0.13951	0.06910	5.33512
	3000	−0.46830	0.29693	0.06704	0.16129	−0.06756	−0.19803	0.40240	0.14385	0.08923	4.02737
True values		−0.46000	0.30000	0.08000	0.17000	−0.07000	−0.22000	0.40000	0.15000	0.08000	

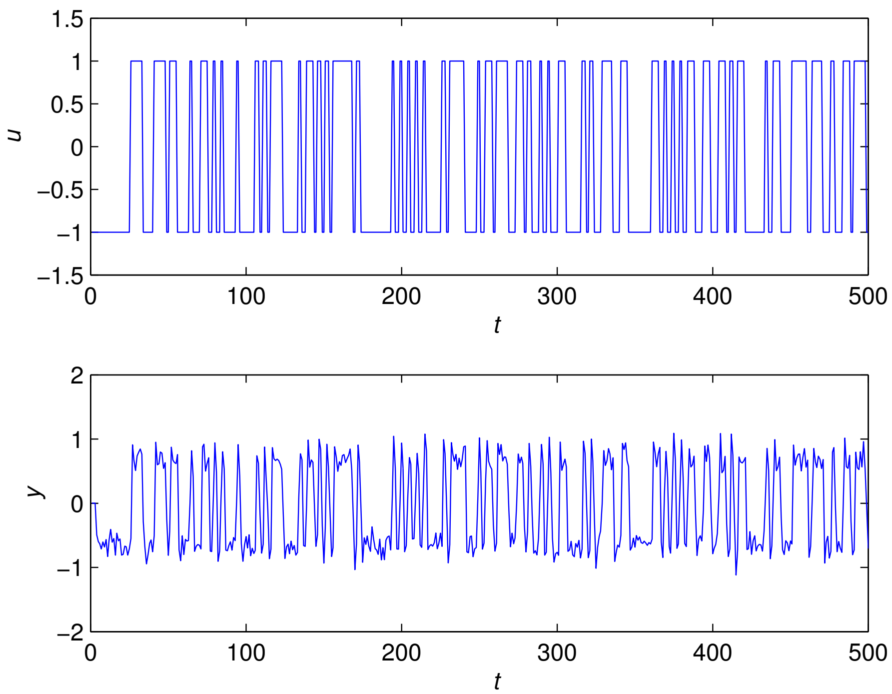

Figure 1. The simulated input-output data versus t.

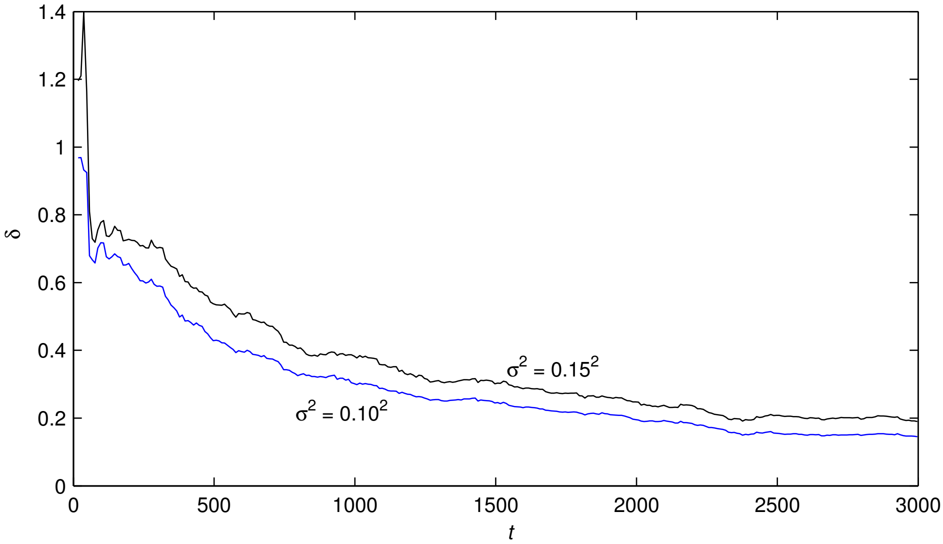

Figure 2. The RGLS parameter estimation errors δ versus t with different variances.

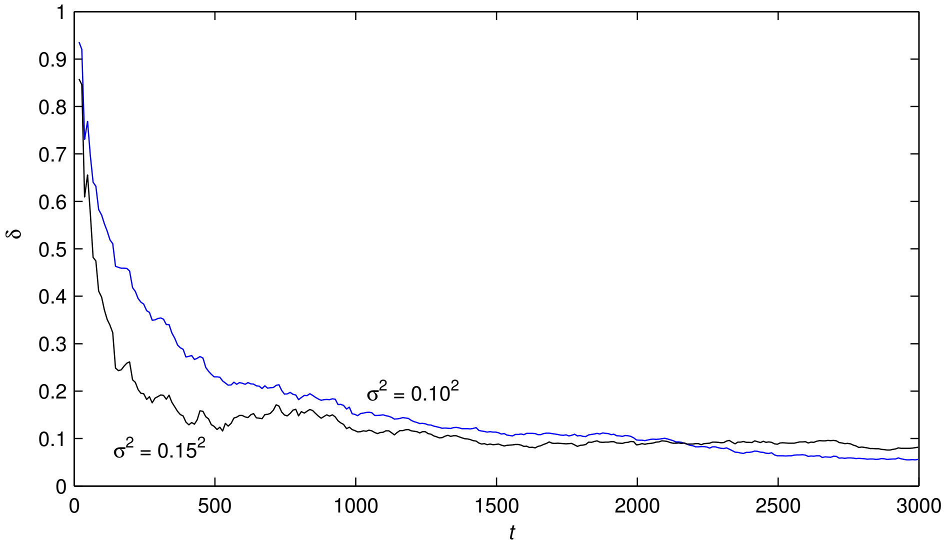

Figure 3. The F-TS-RGLS parameter estimation errors versus t with different variances.

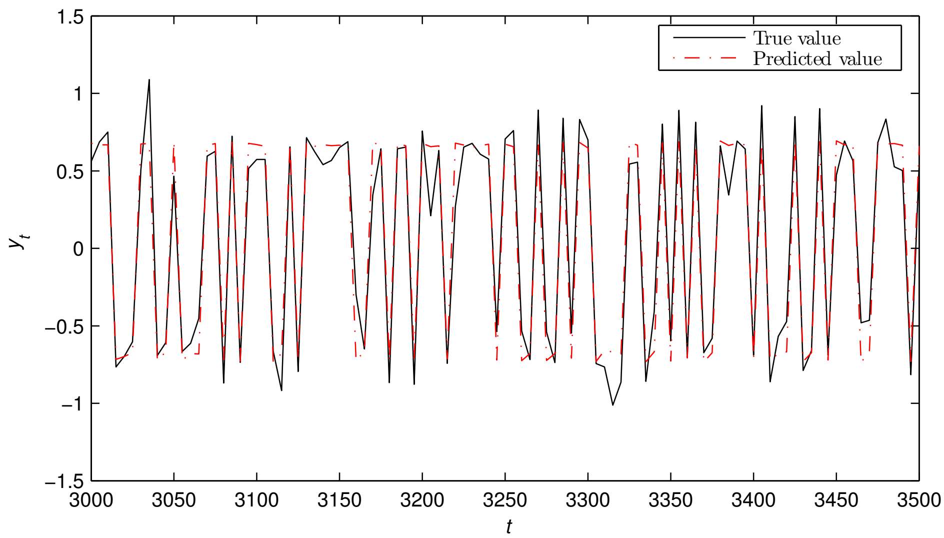

Figure 4. The true outputs and the RGLS predicted outputs.

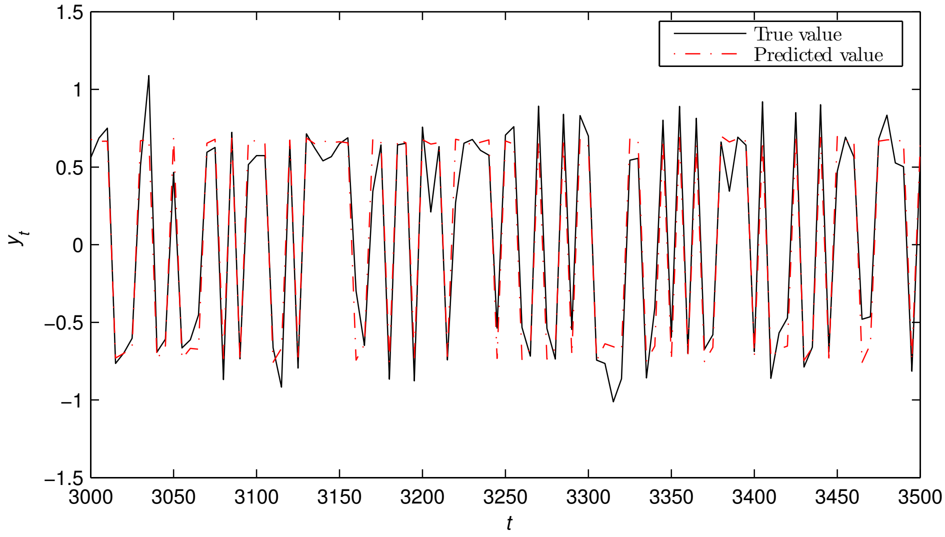

Figure 5. The true outputs and the F-TS-RGLS predicted outputs.

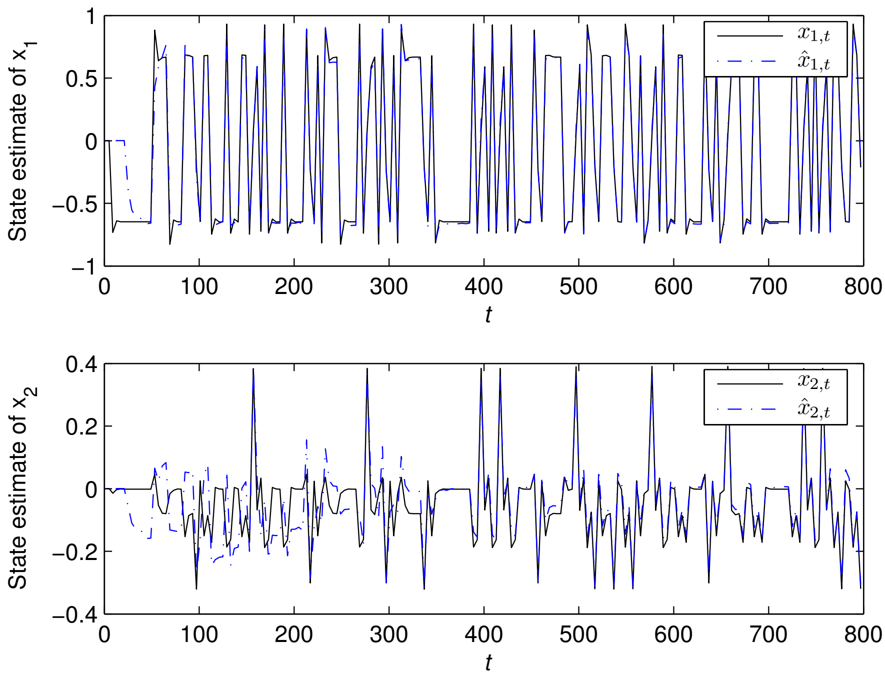

Figure 6. State $\boldsymbol{x}_t$ and its estimate $\hat{\boldsymbol{x}}_t$ versus t ($\sigma^2 = 0.10^2$).

From the simulation results in Tables 2 and 3 and Figures 1–6, we can draw the following conclusions.

- The estimation errors of the RGLS algorithm and the F-TS-RGLS algorithm became smaller with the data length increasing. This means that the proposed algorithms are effective.
- Under the same noise levels, the state estimator-based F-TS-RGLS algorithm could generate more accurate parameter estimates than the state estimator-based RGLS algorithm.
- The F-TS-RGLS algorithm could decrease the dimensions of the covariance matrices and improve the computational efficiency.
- The state estimates obtained from the bilinear state estimator could track their true values as t increased.

6. Conclusions

This paper presents a recursive parameter and state estimation algorithm on the basis of the hierarchical identification principle for bilinear systems. In this approach, the system parameters and states could be estimated interactively. The bilinear state estimator was derived based on the Kalman filter. Then, the state estimator-based RGLS algorithm and the state estimator-based F-TS-RGLS algorithm were proposed based on the data filtering technique and the decomposition-coordination principle. The simulation results showed that the state estimator-based F-TS-RGLS algorithm had higher parameter estimation accuracy compared with the state estimator-based RGLS algorithm. Moreover, the state estimator-based F-TS-RGLS algorithm could greatly improve the computational efficiency. The methods proposed in this paper could be combined the particle Monte Carlo methods [56,57] and some statistical methods to study the model selection and parameter estimation [58–60] for different systems [61–66] and could be applied to other fields [67–72] such as communication networks [73–75].

Author Contributions: Conceptualization and methodology, X.Z. and F.D.; software, X.Z. and L.X.; validation and analysis, A.A. and T.H. Finally, all the authors have read and approved the final manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (No. 61873111), the 111 Project (B12018), and the Postgraduate Research and Practice Innovation Program of Jiangsu Province (KYCX18_1854).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, K.; Jiang, B.; Shi, P. Adjustable parameter-based distributed fault estimation observer design for multiagent systems with directed graphs. *IEEE Trans. Cybern.* **2017**, *47*, 306–314. [CrossRef] [PubMed]
2. Xu, L. A proportional differential control method for a time-delay system using the Taylor expansion approximation. *Appl. Math. Comput.* **2014**, *236*, 391–399. [CrossRef]
3. Xu, L. Application of the Newton iteration algorithm to the parameter estimation for dynamical systems. *J. Comput. Appl. Math.* **2015**, *288*, 33–43. [CrossRef]
4. Xu, L.; Chen, L.; Xiong, W.L. Parameter estimation and controller design for dynamic systems from the step responses based on the Newton iteration. *Nonlinear Dyn.* **2015**, *79*, 2155–2163. [CrossRef]
5. Xu, L.; Ding, F. Parameter estimation for control systems based on impulse responses. *Int. J. Control Autom. Syst.* **2017**, *15*, 2471–2479.
6. Wen, Y.Z.; Yin, C.C. Solution of Hamilton-Jacobi-Bellman equation in optimal reinsurance strategy under dynamic VaR constraint. *J. Funct. Spaces* **2019**, *2019*, 6750892. [CrossRef]
7. Xu, L. The parameter estimation algorithms based on the dynamical response measurement data. *Adv. Mech. Eng.* **2017**, *9*, 1–12. [CrossRef]
8. Xu, L.; Ding, F. Iterative parameter estimation for signal models based on measured data. *Circuits Syst. Signal Process.* **2018**, *37*, 3046–3069. [CrossRef]
9. Xu, L.; Xiong, W.L.; Alsaedi, A.; Hayat, T. Hierarchical parameter estimation for the frequency response based on the dynamical window data. *Int. J. Control Autom. Syst.* **2018**, *16*, 1756–1764. [CrossRef]
10. Zhang, X.; Ding, F.; Xu, L.; Yang, E.F. Highly computationally efficient state filter based on the delta operator. *Int. J. Adapt. Control Signal Process.* **2019**. [CrossRef]
11. Chen, G.Y.; Gan, M.; Chen, C.L.P.; Li, H.X. A regularized variable projection algorithm for separable nonlinear least-squares problems. *IEEE Trans. Autom. Control* **2019**, *64*, 526–537. [CrossRef]
12. Chen, G.Y.; Gan, M.; Ding, F.; Chen, C.L.P. Modified Gram-Schmidt method-based variable projection algorithm for separable nonlinear models. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**. [CrossRef]
13. Gan, M.; Li, H.X. An efficient variable projection formulation for separable nonlinear least squares problems. *IEEE Trans. Cybern.* **2014**, *44*, 707–711. [CrossRef]
14. Ding, F.; Liu, X.P.; Liu, G. Gradient based and least-squares based iterative identification methods for OE and OEMA systems. *Digit. Signal Process.* **2010**, *20*, 664–677. [CrossRef]

15. Ding, F.; Liu, X.G.; Chu, J. Gradient-based and least-squares-based iterative algorithms for Hammerstein systems using the hierarchical identification principle. *IET Control Theory Appl.* **2013**, *7*, 176–184. [CrossRef]
16. Ding, F. Decomposition based fast least squares algorithm for output error systems. *Signal Process.* **2013**, *93*, 1235–1242. [CrossRef]
17. Ding, F. Two-stage least squares based iterative estimation algorithm for CARARMA system modeling. *Appl. Math. Model.* **2013**, *37*, 4798–4808. [CrossRef]
18. Liu, Y.J.; Wang, D.Q.; Ding, F. Least squares based iterative algorithms for identifying Box-Jenkins models with finite measurement data. *Digit. Signal Process.* **2010**, *20*, 1458–1467. [CrossRef]
19. Xu, H.; Ding, F.; Yang, E.F. Modeling a nonlinear process using the exponential autoregressive time series model. *Nonlinear Dyn.* **2019**, *95*, 2079–2092. [CrossRef]
20. Liu, Q.Y.; Ding, F. Auxiliary model-based recursive generalized least squares algorithm for multivariate output-error autoregressive systems using the data filtering. *Circuits Syst. Signal Process.* **2019**, *38*, 590–610. [CrossRef]
21. Ge, Z.W.; Ding, F.; Xu, L.; Alsaedi, A.; Hayat, T. Gradient-based iterative identification method for multivariate equation-error autoregressive moving average systems using the decomposition technique. *J. Frankl. Inst.* **2019**, *356*, 1658–1676. [CrossRef]
22. Tian, X.P.; Niu, H.M. A bi-objective model with sequential search algorithm for optimizing network-wide train timetables. *Comput. Ind. Eng.* **2019**, *127*, 1259–1272. [CrossRef]
23. Yang, F.; Zhang, P.; Li, X.X. The truncation method for the Cauchy problem of the inhomogeneous Helmholtz equation. *Appl. Anal.* **2019**, *98*, 991–1004. [CrossRef]
24. Zhao, N.; Liu, R.; Chen, Y.; Wu, M.; Jiang, Y.; Xiong, W.; Liu, C. Contract design for relay incentive mechanism under dual asymmetric information in cooperative networks. *Wirel. Netw.* **2018**, *24*, 3029–3044. [CrossRef]
25. Xu, G.H.; Shekofteh, Y.; Akgul, A.; Li, C.B.; Panahi, S. A new chaotic system with a self-excited attractor: Entropy measurement, signal encryption, and parameter estimation. *Entropy* **2018**, *20*, 86. [CrossRef]
26. Li, X.Y.; Li, H.X.; Wu, B.Y. Piecewise reproducing kernel method for linear impulsive delay differential equations with piecewise constant arguments. *Appl. Math. Comput.* **2019**, *349*, 304–313. [CrossRef]
27. Bruni, C.; Dipillo, G.; Koch, G. Bilinear systems: An appealing class of nearly linear systems in theory and applications. *IEEE Trans. Autom. Control* **1974**, *19*, 334–348. [CrossRef]
28. Williamson, D. Observation of bilinear systems with application to biological control. *Automatica* **1977**, *13*, 243–254. [CrossRef]
29. Yu, D.; Shields, D.N. A bilinear fault detection observer. *Automatica* **1996**, *32*, 1597–1602. [CrossRef]
30. Mohler, R.R.; Kolodziej, W.J. An overview of bilinear system theory and applications. *IEEE Trans. Syst. Man Cybern.* **2007**, *10*, 683–688.
31. Favoreel, W.; De Moor, B.; Van Overschee, P. Subspace identification of bilinear systems subject to white inputs. *IEEE Trans. Autom. Control* **1999**, *44*, 1157–1165. [CrossRef]
32. Verdult, V.; Verhaegen, M. Identification of multivariable bilinear state space systems based on subspace techniques and separable least squares optimization. *Int. J. Control* **2001**, *74*, 1824–1836. [CrossRef]
33. Larkowski, T.; Linden, J.G.; Vinsonneau, B.; Burnham, K.J. Frisch scheme identification for dynamic diagonal bilinear models. *Int. J. Control* **2009**, *82*, 1591–1604. [CrossRef]
34. Hizir, N.B.; Phan, M.Q.; Betti, R.; Longman, R.W. Identification of discrete-time bilinear systems through equivalent linear models. *Nonlinear Dyn.* **2012**, *69*, 2065–2078.
35. Vicario, F.; Phan, M.Q.; Betti, R.; Longman, R.W. Linear state representations for identification of bilinear discrete-time models by interaction matrices. *Nonlinear Dyn.* **2014**, *77*, 1561–1576. [CrossRef]
36. Xu, L. The damping iterative parameter identification method for dynamical systems based on the sine signal measurement. *Signal Process.* **2016**, *120*, 660–667. [CrossRef]
37. Xu, L.; Ding, F. Parameter estimation algorithms for dynamical response signals based on the multi-innovation theory and the hierarchical principle. *IET Signal Process.* **2017**, *11*, 228–237. [CrossRef]
38. Xu, L.; Ding, F. Recursive least squares and multi-innovation stochastic gradient parameter estimation methods for signal modeling. *Circuits Syst. Signal Process.* **2017**, *36*, 1735–1753. [CrossRef]
39. Gan, M.; Chen, C.L.P.; Chen, G.Y.; Chen, L. On some separated algorithms for separable nonlinear squares problems. *IEEE Trans. Cybern.* **2018**, *48*, 2866–2874. [CrossRef]
40. Ding, F.; Liu, G.; Liu, X.P. Parameter estimation with scarce measurements. *Automatica* **2011**, *47*, 1646–1655. [CrossRef]

41. Ding, F.; Liu, G.; Liu, X.P. Partially coupled stochastic gradient identification methods for non-uniformly sampled systems. *IEEE Trans. Autom. Control* **2010**, *55*, 1976–1981. [CrossRef]
42. Zhao, S.Y.; Shmaliy, Y.S.; Ahn, C.K.; Liu, F. Adaptive-horizon iterative UFIR filtering algorithm with applications. *IEEE Trans. Ind. Electron.* **2018**, *65*, 6393–6402. [CrossRef]
43. Zhao, S.Y.; Huang, B.; Liu, F. Linear optimal unbiased filter for time-variant systems without apriori information on initial conditions. *IEEE Trans. Autom. Control* **2017**, *62*, 882–887. [CrossRef]
44. Xu, L.; Ding, F.; Gu, Y.; Alsaedi, A.; Hayat, T. A multi-innovation state and parameter estimation algorithm for a state space system with d-step state-delay. *Signal Process.* **2017**, *140*, 97–103. [CrossRef]
45. Chopin, N.; Jacob, P.E.; Papaspiliopoulos, O. SMC2: An efficient algorithm for sequential analysis of state space models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2013**, *75*, 397–426. [CrossRef]
46. Van Wingerden, J.W.; Verhaegen, M. Subspace identification of bilinear and LPV systems for open-and closed-loop data. *Automatica* **2009**, *45*, 372–381. [CrossRef]
47. Stroud, J.R.; Katzfuss, M.; Wikle, C.K. A Bayesian adaptive ensemble Kalman filter for sequential state and parameter estimation. *Mon. Weather Rev.* **2018**, *46*, 373–386. [CrossRef]
48. Urteaga, I.; Bugallo, M.F.; Djurić, P.M. Sequential Monte Carlo methods under model uncertainty. In Proceedings of the 2016 IEEE Statistical Signal Processing Workshop (SSP), Palma de Mallorca, Spain, 26–29 June 2016.
49. Martino, L.; Read, J.; Elvira, V.; Louzada, F. Cooperative parallel particle filters for online model selection and applications to urban mobility. *Digit. Signal Process.* **2017**, *60*, 172–185. [CrossRef]
50. Schön, T.B.; Wills, A.; Ninness, B. System identification of nonlinear state-space models. *Automatica* **2011**, *47*, 39–49. [CrossRef]
51. Li, M.H.; Liu, X.M. The least squares based iterative algorithms for parameter estimation of a bilinear system with autoregressive noise using the data filtering technique. *Signal Process.* **2018**, *147*, 23–34. [CrossRef]
52. Phan, M.Q.; Vicario, F.; Longman, R.W.; Betti, R. Optimal bilinear observers for bilinear state-space models by interaction matrices. *Int. J. Control* **2015**, *88*, 1504–1522. [CrossRef]
53. Zhang, X.; Ding, F.; Alsaadi, F.E.; Hayat, T. Recursive parameter identification of the dynamical models for bilinear state space systems. *Nonlinear Dyn.* **2017**, *89*, 2415–2429. [CrossRef]
54. Zhang, X.; Ding, F.; Xu, L.; Yang, E.F. State filtering-based least squares parameter estimation for bilinear systems using the hierarchical identification principle. *IET Control Theory Appl.* **2018**, *12*, 1704–1713. [CrossRef]
55. Zhang, X.; Xu, L.; Ding, F.; Hayat, T. Combined state and parameter estimation for a bilinear state space system with moving average noise. *J. Frankl. Inst.* **2018**, *355*, 3079–3103. [CrossRef]
56. Martino, L.; Elvira, V.; Camps-Valls, G. Distributed Particle Metropolis-Hastings schemes. In Proceedings of the 2018 IEEE Statistical Signal Processing Workshop, Freiburg, Germany, 10–13 June 2018.
57. Carvalho, C.M.; Johannes, M.S.; Lopes, H.F.; Polson, N.G. Particle learning and smoothing. *Stat. Sci.* **2010**, *25*, 88–106. [CrossRef]
58. Ma, F.Y.; Yin, Y.K.; Li, M. Start-up process modelling of sediment microbial fuel cells based on data driven. *Math. Probl. Eng.* **2019**, *2019*, 7403732. [CrossRef]
59. Pan, J.; Ma, H.; Jiang, X.; Ding, W.; Ding, F. Adaptive gradient-based iterative algorithm for multivariate controlled autoregressive moving average systems using the data filtering technique. *Complexity* **2018**, *2018*, 9598307. [CrossRef]
60. Pan, J.; Li, W.; Zhang, H.P. Control algorithms of magnetic suspension systems based on the improved double exponential reaching law of sliding mode control. *Int. J. Control Autom. Syst.* **2018**, *16*, 2878–2887. [CrossRef]
61. Wu, M.H.; Li, X.; Liu, C.; Liu, M.; Zhao, N.; Wang, J.; Wan, X.; Rao, Z.; Zhu, L. Robust global motion estimation for video security based on improved k-means clustering. *J. Ambient Intell. Humaniz. Comput.* **2019**, *10*, 439–448. [CrossRef]
62. Wan, X.K.; Wu, H.; Qiao, F.; Li, F.C.; Li, Y.; Yan, Y.W.; Wei, J.X. Electrocardiogram baseline wander suppression based on the combination of morphological and wavelet transformation based filtering. *Comput. Math. Methods Med.* **2019**, *2019*, 7196156. [CrossRef]
63. Wang, T.; Liu, L.; Zhang, J.; Schaeffer, E.; Wang, Y. A M-EKF fault detection strategy of insulation system for marine current turbine. *Mech. Syst. Signal Process.* **2019**, *115*, 269–280. [CrossRef]

64. Wang, Y.; Si, Y.; Huang, B.; Lou, Z. Survey on the theoretical research and engineering applications of multivariate statistics process monitoring algorithms: 2008–2017. *Can. J. Chem. Eng.* **2018**, *96*, 2073–2085. [CrossRef]
65. Gu, Y.; Chou, Y.; Liu, J.; Ji, Y. Moving horizon estimation for multirate systems with time-varying time-delays. *J. Frankl. Inst.* **2019**, *356*, 2325–2345. [CrossRef]
66. Wang, Y.J.; Ding, F. Iterative estimation for a non-linear IIR filter with moving average noise by means of the data filtering technique. *IMA J. Math. Control Inf.* **2017**, *34*, 745–764. [CrossRef]
67. Cao, Y.; Lu, H.; Wen, T. A safety computer system based on multi-sensor data processing. *Sensors* **2019**, *19*, 818. [CrossRef]
68. Cao, Y.; Zhang, Y.; Wen, T.; Li, P. Research on dynamic nonlinear input prediction of fault diagnosis based on fractional differential operator equation in high-speed train control system. *Chaos* **2019**, *29*, 013130. [CrossRef]
69. Cao, Y.; Li, P.; Zhang, Y. Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing. *Future Gener. Comput. Syst.* **2018**, *88*, 279–283. [CrossRef]
70. Cao, Y.; Ma, L.C.; Xiao, S.; Zhang, X.; Xu, W. Standard analysis for transfer delay in CTCS-3. *Chin. J. Electron.* **2017**, *26*, 1057–1063. [CrossRef]
71. Jiang, C.M.; Zhang, F.F.; Li, T.X. Synchronization and antisynchronization of N-coupled fractional-order complex chaotic systems with ring connection. *Math. Methods Appl. Sci.* **2018**, *41*, 2625–2638. [CrossRef]
72. Zhang, W.H.; Xue, L.; Jiang, X. Global stabilization for a class of stochastic nonlinear systems with SISS-like conditions and time delay. *Int. J. Robust Nonlinear Control* **2018**, *28*, 3909–3926. [CrossRef]
73. Zhao, N.; Chen, Y.; Liu, R.; Wu, M.H.; Xiong, W. Monitoring strategy for relay incentive mechanism in cooperative communication networks. *Comput. Electr. Eng.* **2017**, 60, 14–29. [CrossRef]
74. Zhao, N.; Wu, M.H.; Chen, J.J. Android-based mobile educational platform for speech signal processing. *Int. J. Electr. Eng. Edu.* **2017**, 54, 3–16. [CrossRef]
75. Ji, Y.; Ding, F. Multiperiodicity and exponential attractivity of neural networks with mixed delays. *Circuits Syst. Signal Process.* **2017**, 36, 2558–2573. [CrossRef]