algorithms

MDPI

Article

The Bias Compensation Based Parameter and State Estimation for Observability Canonical State-Space Models with Colored Noise

Xuehai Wang [1,*], Feng Ding [2], Qingsheng Liu [1] and Chuntao Jiang [1]

1 College of Mathematics and Statistics, Xinyang Normal University, Xinyang 464000, China; qingsheng408@163.com (Q.L.); chuntao118@163.com (C.J.)
2 School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China; fding@jiangnan.edu.cn
* Correspondence: xuehaiwang735@163.com

Received: 4 September 2018; Accepted: 22 October 2018; Published: 1 November 2018

check for updates

Abstract: This paper develops a bias compensation-based parameter and state estimation algorithm for the observability canonical state-space system corrupted by colored noise. The state-space system is transformed into a linear regressive model by eliminating the state variables. Based on the determination of the noise variance and noise model, a bias correction term is added into the least squares estimate, and the system parameters and states are computed interactively. The proposed algorithm can generate the unbiased parameter estimate. Two illustrative examples are given to show the effectiveness of the proposed algorithm.

Keywords: recursive identification; least squares; bias compensation; state-space model; state estimation

---

1. Introduction

Mathematical models play a great role in adaptive control, online prediction and signal modeling [1–4]. The system models of physical plants can be constructed from first principle, but this method cannot work well when the physical mechanism of the true plants is not very clear [5,6]. Thus, data-driven-based black-box or gray-box state-space modeling can be used to approximate the law of motion [7–9]. The advantages of state-space modeling lie in that the model maps the relationship from inputs to outputs, and the states reflect the internal dynamic behavior of the plants. The estimation of state-space models may include both the unknown system parameters and the unmeasurable states [10,11]. It is well known that the Kalman filtering method captures the optimal state estimator based on the measurement data [12–14]. However, the model parameters of the system are generally assumed to be known in advance, and the statistical characteristics of measurement noise and process noise have to be chosen, which are hard to satisfy in practice [15–17].

On the identification of state-space models, the classical identification methods are the prediction-error methods, which require adequate knowledge of the system structures and parameters [18,19], and the subspace identification methods, which ignore the consistency of system parameter estimates [20–22]. Since the system states and parameters are involved in the state-space model, the simultaneous estimation of them is a feasible choice [23,24]. In this aspect, Pavelková and Kárný investigated the state and parameter estimation problem of the state-space model with bounded noise, providing a maximum estimator based on Bayesian theory [25]. By transforming the single-input multiple-output Hammerstein state-space model into the multivariable one, Ma et al. used the Kalman smoother to estimate the system states and the expectation maximization algorithm to compute parameter estimates [26]. Li et al. presented a maximum likelihood-based identification algorithm for the Wiener nonlinear system with a linear state-space subsystem [27]. Wang and Liu applied the recursive least squares algorithm to non-uniformly sampled state-space systems with the states being available and combined the singular value decomposition with the hierarchical identification principle for the identification of the system with the states being unavailable [28]. These studies were mainly focused on the state-space system with white noise.

In industrial processes, the observation data are often corrupted by colored noises [29–31]. As we have known, the parameter estimates generated by the recursive least squares algorithm for the system with colored noise are biased. To obtain the unbiased estimates, the bias compensation method was proposed [32]. The basic idea is to compensate the biased least squares estimate by adding a correction term [33–35]. Zhang investigated the parameter estimation of the multiple-input single-output linear system with colored noise and introduced a stable prefilter to preprocess the input data for the purpose of obtaining the unbiased estimate [36]. Zheng extended the bias compensation method to the errors-in-variables output-error system, where the input was contaminated by white noise and the output errors were modeled by colored noise [37]. On the identification of state-space models with colored noise, Wang et al. applied the Kalman filter algorithm to estimate the system states and developed a filtering-based recursive least squares algorithm for observer canonical state-space systems with colored noise [38]. On the basis of the work in [38], this paper discusses the state and parameter identification of the state-space model with colored noise and presents a bias compensation-based identification algorithm for jointly estimating the system parameters and states based on the bias compensation. The main contributions of the paper are as follows.

- By using the bias compensation, this paper derives the identification model and achieves the unbiased parameter estimation for observability canonical state-space models with colored noise.
- By employing the interactive identification, this paper explores the relationship between the noise parameters and variance and the bias correction term and realizes the simultaneous estimation of the system parameters, noise parameters and system states.

The rest of this paper is organized as follows. Section 2 demonstrates the problem formulation about the observability canonical state-space system and derives the identification model. Section 3 develops a bias compensation-based parameter and state estimation algorithm. Section 4 provides two illustrative examples to show that the proposed algorithm is effective. Finally, Section 5 offers some concluding remarks.

2. Problem Description and Identification Model

For narrative convenience, let us introduce some notation. The nomenclature is displayed in Table 1.

Table 1. The nomenclature for symbols.

Nomenclature			
T	matrix/vector transpose	$\boldsymbol{A}^{-1}$	inverse of the matrix $\boldsymbol{A}$
E	expectation operator	z	unit forward shift operator
$\\|\boldsymbol{X}\\|$	2-norm of a matrix $\boldsymbol{X}$	$\boldsymbol{I}_m$	$m \times m$ identity matrix
$\mathbf{1}_n$	n-dimensional column vector whose elements are 1		
$\hat{\boldsymbol{\theta}}(k)$	estimate of the parameter vector $\boldsymbol{\theta}$ at instant k		
$\boldsymbol{\theta}(1:m)$	a vector consisting of the first entry to the m-th entry of $\boldsymbol{\theta}$		

Consider the following state-space system with colored noise,

$$\boldsymbol{x}(k+1) = \boldsymbol{G}\boldsymbol{x}(k) + \boldsymbol{h}u(k), \tag{1}$$

$$y(k) = \boldsymbol{f}\boldsymbol{x}(k) + e(k), \tag{2}$$

where $\boldsymbol{x}(k) := [x_1(k), x_2(k), \cdots, x_n(k)]^{\mathrm{T}} \in \mathbb{R}^n$ is the state vector, $u(k) \in \mathbb{R}$ is the system input, $y(k) \in \mathbb{R}$ is the system output, $e(k) \in \mathbb{R}$ is a random noise and $\boldsymbol{G} \in \mathbb{R}^{n\times n}$, $\boldsymbol{h} \in \mathbb{R}^n$ and $\boldsymbol{f} \in \mathbb{R}^{1\times n}$ are the system parameter matrix and vectors, defined as:

$$\boldsymbol{G} := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ g_1 & g_2 & g_3 & \cdots & g_n \end{bmatrix} \in \mathbb{R}^{n\times n}, \quad \boldsymbol{h} := \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_{n-1} \\ h_n \end{bmatrix} \in \mathbb{R}^n,$$
$$\boldsymbol{f} := [1,\ 0, \cdots,\ 0,\ 0] \in \mathbb{R}^{1\times n}.$$

The external disturbance $e(k)$ can be fitted by a moving average process, an autoregressive process or an autoregressive moving average process. Without loss of generality, we consider $e(k)$ as a moving average noise process:

$$e(k) := (1 + e_1 z^{-1} + e_2 z^{-2} + \cdots + e_{n_e} z^{-n_e}) v(k), \tag{3}$$

where $v(k)$ is the white noise with zero mean and variance δ. Assume that $y(k) = 0$, $u(k) = 0$ and $v(k) = 0$ for $k \leqslant 0$. The objective is to estimate the parameters g_i, h_i and e_i and the system state $\boldsymbol{x}(k)$ from the available input-output data $\{u(k), y(k) :\ k = 0, 1, 2, \cdots\}$.

Note that the system in Equations (1) and (2) is an observability canonical form, and the observability matrix $\boldsymbol{T}$ is an identity matrix, i.e.,

$$\boldsymbol{T} = [\boldsymbol{f}^{\mathrm{T}}, (\boldsymbol{f}\boldsymbol{G})^{\mathrm{T}}, \cdots, (\boldsymbol{f}\boldsymbol{G}^{n-1})^{\mathrm{T}}]^{\mathrm{T}} = \boldsymbol{I}_n. \tag{4}$$

From Equations (1) and (2), we have:

$$y(k) = \boldsymbol{f}\boldsymbol{x}(k) + e(k), \tag{5}$$
$$\begin{aligned} y(k+1) &= \boldsymbol{f}\boldsymbol{x}(k+1) + e(k+1) \\ &= \boldsymbol{f}\boldsymbol{G}\boldsymbol{x}(k) + \boldsymbol{f}\boldsymbol{h}u(k) + e(k+1), \end{aligned} \tag{6}$$
$$\begin{aligned} y(k+2) &= \boldsymbol{f}\boldsymbol{G}\boldsymbol{x}(k+1) + \boldsymbol{f}\boldsymbol{h}u(k+1) + e(k+2) \\ &= \boldsymbol{f}\boldsymbol{G}^2\boldsymbol{x}(k) + \boldsymbol{f}\boldsymbol{G}\boldsymbol{h}u(k) + \boldsymbol{f}\boldsymbol{h}u(k+1) + e(k+2), \end{aligned} \tag{7}$$
$$\vdots$$
$$\begin{aligned} y(k+n-1) &= \boldsymbol{f}\boldsymbol{G}^{n-1}\boldsymbol{x}(k) + \boldsymbol{f}\boldsymbol{G}^{n-2}\boldsymbol{h}u(k) + \boldsymbol{f}\boldsymbol{G}^{n-3}\boldsymbol{h}u(k+1) + \cdots \\ &\quad + \boldsymbol{f}\boldsymbol{h}u(k+n-2) + e(k+n-1), \end{aligned} \tag{8}$$
$$\begin{aligned} y(k+n) &= \boldsymbol{f}\boldsymbol{G}^{n}\boldsymbol{x}(k) + \boldsymbol{f}\boldsymbol{G}^{n-1}\boldsymbol{h}u(k) + \boldsymbol{f}\boldsymbol{G}^{n-2}\boldsymbol{h}u(k+1) + \cdots \\ &\quad + \boldsymbol{f}\boldsymbol{h}u(k+n-1) + e(k+n). \end{aligned} \tag{9}$$

Define the parameter matrix $\boldsymbol{M}$ and the information vectors $\boldsymbol{\phi}_{\mathrm{y}}(k)$, $\boldsymbol{\phi}_{\mathrm{u}}(k)$ and $\boldsymbol{\phi}_{\mathrm{e}}(k)$ as:

$$\boldsymbol{M} := \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \boldsymbol{f}\boldsymbol{h} & 0 & \cdots & 0 & 0 \\ \boldsymbol{f}\boldsymbol{G}\boldsymbol{h} & \boldsymbol{f}\boldsymbol{h} & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ \boldsymbol{f}\boldsymbol{G}^{n-2}\boldsymbol{h} & \boldsymbol{f}\boldsymbol{G}^{n-3}\boldsymbol{h} & \cdots & \boldsymbol{f}\boldsymbol{h} & 0 \end{bmatrix} \in \mathbb{R}^{n\times n},$$
$$\boldsymbol{\phi}_{\mathrm{y}}(k) := [y(k-n), y(k-n+1), \cdots, y(k-1)]^{\mathrm{T}} \in \mathbb{R}^n,$$

$$\begin{aligned}
\boldsymbol{\phi}_{\rm u}(k) &:= [u(k-n), u(k-n+1), \cdots, u(k-1)]^{\rm T} \in \mathbb{R}^n, \\
\boldsymbol{\phi}_{\rm e}(k) &:= [e(k-n), e(k-n+1), \cdots, e(k-1)]^{\rm T} \in \mathbb{R}^n.
\end{aligned}$$

According to Equations (5)–(8), we have:

$$\begin{aligned}
\boldsymbol{\phi}_{\rm y}(k+n) &= \boldsymbol{T}\boldsymbol{x}(k) + \boldsymbol{M}\boldsymbol{\phi}_{\rm u}(k+n) + \boldsymbol{\phi}_{\rm e}(k+n) \\
&= \boldsymbol{x}(k) + \boldsymbol{M}\boldsymbol{\phi}_{\rm u}(k+n) + \boldsymbol{\phi}_{\rm e}(k+n). \qquad (10)
\end{aligned}$$

Equation (10) can be rewritten as:

$$\boldsymbol{x}(k) = \boldsymbol{\phi}_{\rm y}(k+n) - \boldsymbol{M}\boldsymbol{\phi}_{\rm u}(k+n) - \boldsymbol{\phi}_{\rm e}(k+n). \qquad (11)$$

Define the parameter vectors $\boldsymbol{\vartheta}$, $\boldsymbol{\vartheta}_{\rm g}$ and $\boldsymbol{\vartheta}_{\rm h}$ and the information vectors $\boldsymbol{\phi}_{\rm s}(k)$ and $\boldsymbol{\phi}_{\rm n}(k)$ as:

$$\begin{aligned}
\boldsymbol{\vartheta} &:= [\boldsymbol{\vartheta}_{\rm g}^{\rm T}, \boldsymbol{\vartheta}_{\rm h}^{\rm T}]^{\rm T} \in \mathbb{R}^{2n}, \\
\boldsymbol{\vartheta}_{\rm g} &:= [\boldsymbol{f}\boldsymbol{G}^n]^{\rm T} \in \mathbb{R}^n, \\
\boldsymbol{\vartheta}_{\rm h} &:= [-\boldsymbol{f}\boldsymbol{G}^n\boldsymbol{M} + [\boldsymbol{f}\boldsymbol{G}^{n-1}\boldsymbol{h}, \boldsymbol{f}\boldsymbol{G}^{n-2}\boldsymbol{h}, \cdots, \boldsymbol{f}\boldsymbol{h}]]^{\rm T} \in \mathbb{R}^n, \\
\boldsymbol{\phi}_{\rm s}(k) &:= \begin{bmatrix} \boldsymbol{\phi}_{\rm y}(k) \\ \boldsymbol{\phi}_{\rm u}(k) \end{bmatrix} \in \mathbb{R}^{2n}, \quad \boldsymbol{\phi}_{\rm n}(k) := \begin{bmatrix} -\boldsymbol{\phi}_{\rm e}(k) \\ \boldsymbol{0} \end{bmatrix} \in \mathbb{R}^{2n}.
\end{aligned}$$

Inserting (11) into (9) yields:

$$\begin{aligned}
y(k+n) &= \boldsymbol{f}\boldsymbol{G}^n[\boldsymbol{\phi}_{\rm y}(k+n) - \boldsymbol{M}\boldsymbol{\phi}_{\rm u}(k+n) - \boldsymbol{\phi}_{\rm e}(k+n)] \\
&\quad + [\boldsymbol{f}\boldsymbol{G}^{n-1}\boldsymbol{h}, \boldsymbol{f}\boldsymbol{G}^{n-2}\boldsymbol{h}, \cdots, \boldsymbol{f}\boldsymbol{h}]\boldsymbol{\phi}_{\rm u}(k+n) + e(k+n) \\
&= \boldsymbol{\phi}_{\rm y}^{\rm T}(k+n)\boldsymbol{\vartheta}_{\rm g} + \boldsymbol{\phi}_{\rm u}^{\rm T}(k+n)\boldsymbol{\vartheta}_{\rm h} - \boldsymbol{\phi}_{\rm e}^{\rm T}(k+n)\boldsymbol{\vartheta}_{\rm g} + e(k+n) \\
&= \boldsymbol{\phi}_{\rm s}^{\rm T}(k+n)\boldsymbol{\vartheta} + \boldsymbol{\phi}_{\rm n}^{\rm T}(k+n)\boldsymbol{\vartheta} + e(k+n). \qquad (12)
\end{aligned}$$

Define the intermediate variable:

$$w(k) := \boldsymbol{\phi}_{\rm n}^{\rm T}(k)\boldsymbol{\vartheta} + e(k). \qquad (13)$$

Replacing $k+n$ in Equation (12) with k gives:

$$\begin{aligned}
y(k) &= \boldsymbol{\phi}_{\rm s}^{\rm T}(k)\boldsymbol{\vartheta} + \boldsymbol{\phi}_{\rm n}^{\rm T}(k)\boldsymbol{\vartheta} + e(k) \qquad (14) \\
&= \boldsymbol{\phi}_{\rm s}^{\rm T}(k)\boldsymbol{\vartheta} + w(k). \qquad (15)
\end{aligned}$$

Equation (14) or (15) is the system identification model of the state-space system in Equation (1), where the information vector $\boldsymbol{\phi}_{\rm s}(k)$ is composed of the observed data. Define the parameter vector $\boldsymbol{\vartheta}_{\rm v}$ and the information vector $\boldsymbol{\phi}_{\rm v}(k)$ as:

$$\begin{aligned}
\boldsymbol{\vartheta}_{\rm v} &:= [e_1, e_2, \cdots, e_{n_e}] \in \mathbb{R}^{n_e}, \\
\boldsymbol{\phi}_{\rm v}(k) &:= [v(k-1), v(k-2), \cdots, v(k-n_e)]^{\rm T} \in \mathbb{R}^{n_e}.
\end{aligned}$$

From Equation (3), we have:

$$e(k) = \boldsymbol{\phi}_{\rm v}^{\rm T}(k)\boldsymbol{\vartheta}_{\rm v} + v(k). \qquad (16)$$

Thus, we obtain the noise identification model in Equation (16). The following derives the bias compensation-based identification algorithm based on the system model in Equation (15) and the noise model in Equation (16).

3. The Bias Compensation-Based Parameter and State Estimation Algorithm

The algorithm includes two parts: the parameter estimation algorithm and the state estimation algorithm. Two parts are implemented in an interactive way.

3.1. The Parameter Estimation Algorithm

According to the identification model in Equation (15), define the cost function:

$$J(\boldsymbol{\vartheta}) := \sum_{i=1}^{k}[y(i) - \boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(i)\boldsymbol{\vartheta}]^2$$

Using the least squares search and minimizing $J(\boldsymbol{\vartheta})$ give the least squares estimate $\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)$ of the parameter vector $\boldsymbol{\vartheta}$:

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) = \boldsymbol{P}(k)\sum_{i=1}^{k}\boldsymbol{\phi}_{\mathrm{s}}(i)y(i), \tag{17}$$

$$\boldsymbol{P}^{-1}(k) = \sum_{i=1}^{k}\boldsymbol{\phi}_{\mathrm{s}}(i)\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(i).$$

Inserting (14) into (17) and using (13), we have:

$$\begin{aligned}\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) &= \boldsymbol{P}(k)\sum_{i=1}^{k}\boldsymbol{\phi}_{\mathrm{s}}(i)[\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(i)\boldsymbol{\vartheta} + \boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(i)\boldsymbol{\vartheta} + e(i)] \\ &= \boldsymbol{\vartheta} + \boldsymbol{P}(k)\sum_{i=1}^{k}\boldsymbol{\phi}_{\mathrm{s}}(i)[\boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(i)\boldsymbol{\vartheta} + e(i)]. \end{aligned} \tag{18}$$

Obviously, the least squares estimate $\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)$ in Equation (18) is a biased estimate since $e(k)$ is correlated noise. Equation (18) can be rewritten as:

$$\boldsymbol{P}^{-1}(k)[\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) - \boldsymbol{\vartheta}] = \sum_{i=1}^{k}\boldsymbol{\phi}_{\mathrm{s}}(i)[\boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(i)\boldsymbol{\vartheta} + e(i)].$$

Dividing by k and taking limits on both sides give:

$$\lim_{k\to\infty}\frac{1}{k}[\boldsymbol{P}^{-1}(k)(\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) - \boldsymbol{\vartheta})] = \left[\lim_{k\to\infty}\frac{1}{k}\sum_{i=1}^{k}\boldsymbol{\phi}_{\mathrm{s}}(i)\boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(i)\right]\boldsymbol{\vartheta} + \lim_{k\to\infty}\frac{1}{k}\sum_{i=1}^{k}\boldsymbol{\phi}_{\mathrm{s}}(i)e(i). \tag{19}$$

Note that $e(k)$ in Equation (19) is the moving average noise, and $v(k)$ is white noise with zero mean and variance δ and is independent of the inputs. From Equation (3), we have:

$$\begin{aligned} R(0) &:= \mathrm{E}[e^2(k)] = (1 + e_1^2 + e_2^2 + \cdots + e_{n_e}^2)\delta = [1 + \boldsymbol{\vartheta}_{\mathrm{v}}^{\mathrm{T}}\boldsymbol{\vartheta}_{\mathrm{v}}]\delta, \\ R(i) &:= \mathrm{E}[e(k-i)e(k)] = (e_i + e_{i+1}e_1 + \cdots + e_{n_e}e_{n_e-i})\delta \\ &= \boldsymbol{\vartheta}_{\mathrm{v}}^{\mathrm{T}}(i:n_e)\begin{bmatrix} 1 \\ \boldsymbol{\vartheta}_{\mathrm{v}}(1:n_e-i)\end{bmatrix}\delta,\ i = 1, 2, \cdots, n_e, \end{aligned}$$

where $R(i)$ is the autocorrelation function of the noise $e(k)$ and $R(i) = 0$ when $i > n_e$. Define the autocorrelation function vectors $\boldsymbol{r}$ and $\boldsymbol{\zeta}$ and the autocorrelation function matrices $\boldsymbol{R}$, $\boldsymbol{\Lambda}$ and $\boldsymbol{Q}$ as:

$$\boldsymbol{r} := [R(n), R(n-1), \cdots, R(1), 0, 0\cdots, 0]^{\mathrm{T}} \in \mathbb{R}^{2n}, \tag{20}$$

$$\boldsymbol{\zeta} := [R(n)/\delta, R(n-1)/\delta, \cdots, R(1)/\delta, 0, 0\cdots, 0]^{\mathrm{T}} \in \mathbb{R}^{2n}, \tag{21}$$

$$\boldsymbol{R} := \mathrm{diag}[\boldsymbol{\Lambda}, \boldsymbol{0}] \in \mathbb{R}^{(2n)\times(2n)}, \tag{22}$$

$$\boldsymbol{\Lambda} := \begin{bmatrix} R(0) & R(1) & \cdots & R(n-1) \\ R(1) & R(0) & \cdots & R(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ R(n-1) & R(n-2) & \cdots & R(0) \end{bmatrix} \in \mathbb{R}^{n\times n}, \tag{23}$$

$$\boldsymbol{Q} := \boldsymbol{R}/\delta. \tag{24}$$

In fact, $\boldsymbol{\Lambda}$ is a Toeplitz matrix consisting of n autocorrelation functions. Equation (19) can be rewritten as:

$$\lim_{k\to\infty} \frac{1}{k}[\boldsymbol{P}^{-1}(k)(\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) - \boldsymbol{\vartheta})] = -(\boldsymbol{R}\boldsymbol{\vartheta} - \boldsymbol{r}).$$

or:

$$\lim_{k\to\infty} \hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) = \boldsymbol{\vartheta} - k\boldsymbol{P}(k)(\boldsymbol{R}\boldsymbol{\vartheta} - \boldsymbol{r}). \tag{25}$$

It can be seen from Equation (25) that the bias of the least squares estimate $\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)$ can be eliminated by adding a compensation term $\Delta\boldsymbol{\vartheta}(k) := k\boldsymbol{P}(k)(\boldsymbol{R}\boldsymbol{\vartheta} - \boldsymbol{r})$. That is, the estimate $\hat{\boldsymbol{\vartheta}}_{\mathrm{C}}(k) := \hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) + \Delta\boldsymbol{\vartheta}(k)$ is an unbiased estimate of the true parameter vector $\boldsymbol{\vartheta}$. Thus, the unbiased estimate $\hat{\boldsymbol{\vartheta}}_{\mathrm{C}}(k)$ can be computed by the following recursive expressions,

$$\begin{aligned} \hat{\boldsymbol{\vartheta}}_{\mathrm{C}}(k) &= \hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) + k\boldsymbol{P}(k)[\hat{\boldsymbol{R}}(k)\hat{\boldsymbol{\vartheta}}_{\mathrm{C}}(k) - \hat{\boldsymbol{r}}(k)] \\ &= \hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) + k\hat{\delta}(k)\boldsymbol{P}(k)[\hat{\boldsymbol{Q}}(k)\hat{\boldsymbol{\vartheta}}_{\mathrm{C}}(k) - \hat{\boldsymbol{\zeta}}(k)], \end{aligned} \tag{26}$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) = \hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k-1) + \boldsymbol{P}(k)\boldsymbol{\phi}_{\mathrm{s}}(k)[y(k) - \boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(k)\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k-1)], \tag{27}$$

$$\boldsymbol{P}^{-1}(k) = \boldsymbol{P}^{-1}(k-1) + \|\boldsymbol{\phi}_{\mathrm{s}}(k)\|^2,\ \boldsymbol{P}(0) = p_0\boldsymbol{I}_{2n}, \tag{28}$$

$$\boldsymbol{\phi}_{\mathrm{s}}(k) = \begin{bmatrix} \boldsymbol{\phi}_{\mathrm{y}}(k) \\ \boldsymbol{\phi}_{\mathrm{u}}(k) \end{bmatrix}, \quad \hat{\boldsymbol{\vartheta}}_{\mathrm{C}}(k) = \begin{bmatrix} \hat{\boldsymbol{\vartheta}}_{\mathrm{g}}(k) \\ \hat{\boldsymbol{\vartheta}}_{\mathrm{h}}(k) \end{bmatrix}. \tag{29}$$

Equation (26) shows that the unbiased estimate $\hat{\boldsymbol{\vartheta}}_{\mathrm{C}}(k)$ is related to the estimates of $\boldsymbol{R}$ and $\boldsymbol{r}$ (i.e., the noise variance δ and the noise parameter vector $\boldsymbol{\vartheta}_{\mathrm{v}}$). The following derives their estimates based on the interactive identification.

Let the least squares residual $\varepsilon_{\mathrm{LS}}(i) := y(i) - \boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(i)\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)$. Using Equation (14) and the relation $\sum_{i=1}^{k} \varepsilon_{\mathrm{LS}}(i)\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(i) = \boldsymbol{0}$ gives:

$$\begin{aligned} \sum_{i=1}^{k} \varepsilon_{\mathrm{LS}}^2(i) &= \sum_{i=1}^{k} \varepsilon_{\mathrm{LS}}(i)[\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(i)\boldsymbol{\vartheta} + \boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(i)\boldsymbol{\vartheta} + e(i) - \boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(i)\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)] \\ &= \sum_{i=1}^{k} [\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(i)\boldsymbol{\vartheta} + \boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(i)\boldsymbol{\vartheta} + e(i) - \boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(i)\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)][\boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(i)\boldsymbol{\vartheta} + e(i)] \\ &= \sum_{i=1}^{k} \boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(i)[\boldsymbol{\vartheta} - \hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)][\boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(i)\boldsymbol{\vartheta} + e(i)] + \sum_{i=1}^{k} [\boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(i)\boldsymbol{\vartheta} + e(i)]^2. \end{aligned}$$

Noting that $\boldsymbol{R}$ is a symmetric matrix, and using Equations (20)–(24), we have:

$$\begin{aligned} \lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} \varepsilon_{\mathrm{LS}}^2(i) &= -\boldsymbol{\vartheta}^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}[\boldsymbol{\vartheta} - \hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)] + \boldsymbol{r}^{\mathrm{T}}[\boldsymbol{\vartheta} - \hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)] + \boldsymbol{\vartheta}\boldsymbol{R}\boldsymbol{\vartheta} - 2\boldsymbol{r}^{\mathrm{T}}\boldsymbol{\vartheta} + R(0) \\ &= \boldsymbol{\vartheta}^{\mathrm{T}}\boldsymbol{R}\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) - \boldsymbol{r}^{\mathrm{T}}[\boldsymbol{\vartheta} + \hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)] + R(0) \\ &= \boldsymbol{\vartheta}^{\mathrm{T}}\delta\boldsymbol{Q}\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) - \delta\boldsymbol{\zeta}^{\mathrm{T}}[\boldsymbol{\vartheta} + \hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)] + \delta[1 + \boldsymbol{\vartheta}_{\mathrm{v}}^{\mathrm{T}}\boldsymbol{\vartheta}_{\mathrm{v}}]. \end{aligned}$$

Thus, we have:

$$\delta = \frac{\lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} \varepsilon_{\mathrm{LS}}^2(i)}{\boldsymbol{\vartheta}^{\mathrm{T}}\boldsymbol{Q}\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k) - \boldsymbol{\zeta}^{\mathrm{T}}[\boldsymbol{\vartheta} + \hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)] + 1 + \boldsymbol{\vartheta}_{\mathrm{v}}^{\mathrm{T}}\boldsymbol{\vartheta}_{\mathrm{v}}}.$$

Let $J(k) := \sum_{i=1}^{k} \varepsilon_{\rm LS}^2(i)$ be the cost function at time k, $\hat{\boldsymbol{Q}}(k)$ and $\hat{\boldsymbol{\zeta}}(k)$ be the estimates of $\boldsymbol{Q}$ and $\boldsymbol{\zeta}$ at instant k, respectively. Then, the estimate of the noise variance δ can be computed by:

$$\hat{\delta}(k) = \frac{\frac{1}{k}J(k)}{\hat{\boldsymbol{\vartheta}}_{\rm C}^{\rm T}(k)\hat{\boldsymbol{Q}}(k)\hat{\boldsymbol{\vartheta}}_{\rm LS}(k) - \hat{\boldsymbol{\zeta}}^{\rm T}(k)[\hat{\boldsymbol{\vartheta}}_{\rm C}(k) + \hat{\boldsymbol{\vartheta}}_{\rm LS}(k)] + 1 + \hat{\boldsymbol{\vartheta}}_{\rm v}^{\rm T}(k)\hat{\boldsymbol{\vartheta}}_{\rm v}(k)}, \tag{30}$$

$$J(k) = J(k-1) + \frac{[y(k) - \boldsymbol{\phi}_{\rm s}^{\rm T}(k)\hat{\boldsymbol{\vartheta}}_{\rm LS}(k-1)]^2}{1 + \boldsymbol{\phi}_{\rm s}^{\rm T}(k)\boldsymbol{P}(k-1)\boldsymbol{\phi}_{\rm s}(k)}, \tag{31}$$

$$\hat{\boldsymbol{\zeta}}(k) = [\hat{R}(n)/\hat{\delta}(k-1), \hat{R}(n-1)/\hat{\delta}(k-1), \cdots, \hat{R}(1)/\hat{\delta}(k-1), 0, 0 \cdots, 0]^{\rm T}, \tag{32}$$

$$\hat{\boldsymbol{Q}}(k) = \left[\begin{array}{cc} \left[\begin{array}{cccc} 1 + \hat{\boldsymbol{\vartheta}}_{\rm v}^{\rm T}(k)\hat{\boldsymbol{\vartheta}}_{\rm v}(k) & \hat{R}(1)/\hat{\delta}(k-1) & \cdots & \hat{R}(n-1)/\hat{\delta}(k-1) \\ \hat{R}(1)/\hat{\delta}(k-1) & 1 + \hat{\boldsymbol{\vartheta}}_{\rm v}^{\rm T}(k)\hat{\boldsymbol{\vartheta}}_{\rm v}(k) & \cdots & \hat{R}(n-2)/\hat{\delta}(k-1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}(n-1)/\hat{\delta}(k-1) & \hat{R}(n-2)/\hat{\delta}(k-1) & \cdots & 1 + \hat{\boldsymbol{\vartheta}}_{\rm v}^{\rm T}(k)\hat{\boldsymbol{\vartheta}}_{\rm v}(k) \end{array}\right] & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{array}\right], \tag{33}$$

$$\frac{\hat{R}(i)}{\hat{\delta}(k-1)} = \hat{\boldsymbol{\vartheta}}_{\rm v}^{\rm T}(i:n_e)\begin{bmatrix} 1 \\ \hat{\boldsymbol{\vartheta}}_{\rm v}(1:n_e-i) \end{bmatrix}, \; i = 1, 2, \cdots, n_e. \tag{34}$$

Note that Equations (30)–(34) involve the estimate $\hat{\boldsymbol{\vartheta}}_{\rm v}(k)$ of the noise parameter vector $\boldsymbol{\vartheta}_{\rm v}$, which can be computed by the noise model in Equation (16). Let $\hat{e}(k)$ and $\hat{v}(k)$ be the estimates of $e(k)$ and $v(k)$, respectively. Define the noise information vectors:

$$\hat{\boldsymbol{\phi}}_{\rm n}(k) := [-\hat{e}(k-n), -\hat{e}(k-n+1), \cdots, -\hat{e}(k-1), 0, 0, \cdots, 0]^{\rm T} \in \mathbb{R}^{2n}, \tag{35}$$

$$\hat{\boldsymbol{\phi}}_{\rm v}(k) := [\hat{v}(k-1), \hat{v}(k-2), \cdots, \hat{v}(k-n_e)]^{\rm T} \in \mathbb{R}^{n_e}. \tag{36}$$

From Equations (14) and (16), we have:

$$\hat{e}(k) = y(k) - \boldsymbol{\phi}_{\rm s}^{\rm T}(k)\hat{\boldsymbol{\vartheta}}_{\rm C}(k) - \hat{\boldsymbol{\phi}}_{\rm n}^{\rm T}(k)\hat{\boldsymbol{\vartheta}}_{\rm C}(k), \tag{37}$$

$$\hat{v}(k) = \hat{e}(k) - \hat{\boldsymbol{\phi}}_{\rm v}^{\rm T}(k)\hat{\boldsymbol{\vartheta}}_{\rm v}(k). \tag{38}$$

Using the least squares principle, the estimate $\hat{\boldsymbol{\vartheta}}_{\rm v}(k)$ of $\boldsymbol{\vartheta}_{\rm v}$ in Equation (16) can be computed by:

$$\hat{\boldsymbol{\vartheta}}_{\rm v}(k) = \hat{\boldsymbol{\vartheta}}_{\rm v}(k-1) + \boldsymbol{P}_{\rm v}(k)\hat{\boldsymbol{\phi}}_{\rm v}(k)[\hat{e}(k) - \hat{\boldsymbol{\phi}}_{\rm v}^{\rm T}(k)\hat{\boldsymbol{\vartheta}}_{\rm v}(k-1)], \tag{39}$$

$$\boldsymbol{P}_{\rm v}^{-1}(k) = \boldsymbol{P}_{\rm v}^{-1}(k-1) + \hat{\boldsymbol{\phi}}_{\rm v}(k)\hat{\boldsymbol{\phi}}_{\rm v}^{\rm T}(k). \tag{40}$$

Thus, Equations (26)–(40) form the bias compensation-based parameter estimation (BC-PE) algorithm for identifying the parameter vector $\boldsymbol{\vartheta}$. The BC-PE algorithm is an interactive estimation process: we can compute the estimates $\hat{\boldsymbol{\vartheta}}_{\rm v}(k)$ and $\hat{\delta}(k)$; with these obtained variable estimates, we derive the estimate of $\boldsymbol{R}$ and $\boldsymbol{r}$ and then update the unbiased estimate $\hat{\boldsymbol{\vartheta}}_{\rm C}(k)$ using Equations (26)–(29).

3.2. The State Estimation Algorithm

When the system parameter estimation and the noise estimation $\hat{e}(k)$ are obtained from the BC-PE algorithm, we can estimate the system state $\boldsymbol{x}(t)$ using Equation (11).

Post-multiplying Equation (4) by $\boldsymbol{h}$ and letting the elements of each row be equal, we have:

$$\boldsymbol{f}\boldsymbol{G}^{i-1}\boldsymbol{h} = h_i, \; i = 1, 2, \cdots, n.$$

Then, the parameter matrix $\boldsymbol{M}$ can be expressed as:

$$\boldsymbol{M} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ h_1 & 0 & \cdots & 0 & 0 \\ h_2 & h_1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ h_{n-1} & h_{n-2} & \cdots & h_1 & 0 \end{bmatrix}.$$

Post-multiplying Equation (4) by $\boldsymbol{G}$ and letting the elements of n-th row be equal, we have:

$$\boldsymbol{f}\boldsymbol{G}^n = [g_1, g_2, \cdots, g_n].$$

According to the definitions of $\boldsymbol{\vartheta}_{\mathrm{g}}$ and $\boldsymbol{\vartheta}_{\mathrm{h}}$, we have:

$$\begin{aligned}
\boldsymbol{\vartheta}_{\mathrm{g}} &= [\boldsymbol{f}\boldsymbol{G}^n]^{\mathrm{T}} = [g_1, g_2, \cdots, g_n]^{\mathrm{T}}, \\
\boldsymbol{\vartheta}_{\mathrm{h}} &= [-\boldsymbol{f}\boldsymbol{G}^n\boldsymbol{M} + [\boldsymbol{f}\boldsymbol{G}^{n-1}\boldsymbol{h}, \boldsymbol{f}\boldsymbol{G}^{n-2}\boldsymbol{h}, \cdots, \boldsymbol{f}\boldsymbol{h}]]^{\mathrm{T}} \\
&= \begin{bmatrix} -h_1g_2 - h_2g_3 - \cdots - h_{n-1}g_n + h_n \\ -h_1g_3 - h_2g_4 - \cdots - h_{n-2}g_n + h_{n-1} \\ \vdots \\ -h_1g_n + h_2 \\ h_1 \end{bmatrix} = \begin{bmatrix} -g_2 & -g_3 & \cdots & -g_n & 1 \\ -g_3 & -g_4 & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ -g_n & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_{n-1} \\ h_n \end{bmatrix}.
\end{aligned}$$

Once the estimate of the parameter vector $\boldsymbol{\vartheta}$ is computed by the BC-PE algorithm in Equations (26)–(40), we can extract the estimates $\hat{\boldsymbol{\vartheta}}_{\mathrm{g}}(k)$ and $\hat{\boldsymbol{\vartheta}}_{\mathrm{h}}(k)$ of the parameter vectors $\boldsymbol{\vartheta}_{\mathrm{g}}$ and $\boldsymbol{\vartheta}_{\mathrm{h}}$. Then, the estimates $\hat{g}_i(k)$ and $\hat{h}_i(k)$ can be computed by:

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{g}}(k) = [\hat{g}_1(k), \hat{g}_2(k), \cdots, \hat{g}_n(k)]^{\mathrm{T}}, \tag{41}$$

$$\begin{bmatrix} \hat{h}_1(k) \\ \hat{h}_2(k) \\ \vdots \\ \hat{h}_{n-1}(k) \\ \hat{h}_n(k) \end{bmatrix} = \begin{bmatrix} -\hat{g}_2(k) & -\hat{g}_3(k) & \cdots & -\hat{g}_n(k) & 1 \\ -\hat{g}_3(k) & -\hat{g}_4(k) & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ -\hat{g}_n(k) & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}^{-1} \hat{\boldsymbol{\vartheta}}_{\mathrm{h}}(k). \tag{42}$$

From Equation (11), we have:

$$\boldsymbol{x}(k-n) = \boldsymbol{\phi}_{\mathrm{y}}(k) - \boldsymbol{M}\boldsymbol{\phi}_{\mathrm{u}}(k) - \boldsymbol{\phi}_{\mathrm{e}}(k).$$

Replacing $\boldsymbol{M}$ and $\boldsymbol{\phi}_{\mathrm{e}}(k)$ with their corresponding estimates $\hat{\boldsymbol{M}}(k)$ and $\hat{\boldsymbol{\phi}}_{\mathrm{e}}(k)$, we have:

$$\hat{\boldsymbol{x}}(k-n) = \boldsymbol{\phi}_{\mathrm{y}}(k) - \hat{\boldsymbol{M}}(k)\boldsymbol{\phi}_{\mathrm{u}}(k) - \hat{\boldsymbol{\phi}}_{\mathrm{e}}(k), \tag{43}$$

$$\boldsymbol{\phi}_{\mathrm{y}}(k) = [y(k-n), y(k-n+1), \cdots, y(k-1)]^{\mathrm{T}}, \tag{44}$$

$$\boldsymbol{\phi}_{\mathrm{u}}(k) = [u(k-n), u(k-n+1), \cdots, u(k-1)]^{\mathrm{T}}, \tag{45}$$

$$\hat{\boldsymbol{\phi}}_{\mathrm{e}}(k) = [\hat{e}(k-n), \hat{e}(k-n+1), \cdots, \hat{e}(k-1)]^{\mathrm{T}}, \tag{46}$$

$$\hat{\boldsymbol{M}}(k) = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \hat{h}_1(k) & 0 & \cdots & 0 & 0 \\ \hat{h}_2(k) & \hat{h}_1(k) & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ \hat{h}_{n-1}(k) & \hat{h}_{n-2}(k) & \cdots & \hat{h}_1(k) & 0 \end{bmatrix}. \tag{47}$$

Equations (41)–(47) form the state estimation algorithm for the state-space system in Equation (1). Let $\boldsymbol{\theta} := [\boldsymbol{g}^{\mathrm{T}}, \boldsymbol{h}^{\mathrm{T}}, \boldsymbol{\vartheta}_{\mathrm{v}}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2n+n_e}$, $\boldsymbol{g} := [g_1, g_2, \cdots, g_n] \in \mathbb{R}^n$. By interactively implementing the parameter estimation algorithm in Equations (26)–(40) and the state estimation algorithm in Equations (41)–(47), the estimation of the system parameter vector $\boldsymbol{\theta}$ and the state $\boldsymbol{x}(k)$ can be obtained.

The steps for implementing the bias compensation-based parameter and state estimation (BC-PSE) algorithm in Equations (26) and (47) for state-space systems with colored noise are listed as follows.

1. Let $k = 1$, and set the initial values $\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(0) = \hat{\boldsymbol{\vartheta}}_{\mathrm{C}}(0) = \mathbf{1}_{2n}/p_0$, $\hat{\boldsymbol{\vartheta}}_{\mathrm{v}}(0) = \mathbf{1}_{n_e}/p_0$, $\boldsymbol{P}(0) = p_0\boldsymbol{I}_{2n}$, $J(0) = 0$, $p_0 = 10^6$.
2. Collect the input-output data $u(k)$ and $y(k)$. Construct $\boldsymbol{\phi}_{\mathrm{y}}(k)$ using Equation (44), $\boldsymbol{\phi}_{\mathrm{u}}(k)$ using Equation (45) and $\boldsymbol{\phi}_{\mathrm{s}}(k)$ using Equation (29).
3. Compute $\boldsymbol{P}(k)$ using Equation (28) and $J(k)$ using Equation (31). Update the parameter estimate $\hat{\boldsymbol{\vartheta}}_{\mathrm{LS}}(k)$ using Equation (27).
4. Construct $\hat{\boldsymbol{\phi}}_{\mathrm{n}}(k)$ using Equation (35) and $\hat{\boldsymbol{\phi}}_{\mathrm{v}}(k)$ using Equation (36). Compute $\hat{e}(k)$ using Equation (37) and $\hat{v}(k)$ using Equation (38), and compute $\boldsymbol{P}_{\mathrm{v}}(k)$ using Equation (40). Update the noise parameter estimate $\hat{\boldsymbol{\vartheta}}_{\mathrm{v}}(k)$ using Equation (39).
5. Compute $\hat{\boldsymbol{\zeta}}(k)$, $\hat{\boldsymbol{Q}}(k)$ and $\hat{R}(i)/\hat{\delta}(k-1)$ using Equations (32)–(34). Compute $\hat{\delta}(k)$ using Equation (30) and $\hat{\boldsymbol{\vartheta}}_{\mathrm{C}}(k)$ using Equation (26).
6. Read $\hat{\boldsymbol{\vartheta}}_{\mathrm{g}}(k)$ and $\hat{\boldsymbol{\vartheta}}_{\mathrm{h}}(k)$ using Equation (29). Acquire the parameter estimates $\hat{g}_i(k)$ and $\hat{h}_i(k)$ using Equations (41) and (42).
7. Construct $\hat{\boldsymbol{\phi}}_{\mathrm{e}}(k)$ using Equation (46) and $\hat{\boldsymbol{M}}(k)$ using Equation (47).
8. Compute the state estimate $\hat{\boldsymbol{x}}(k)$ using Equation (43).
9. Let $\hat{\boldsymbol{\theta}}(k) := \begin{bmatrix} \hat{\boldsymbol{g}}(k) \\ \hat{\boldsymbol{h}}(k) \\ \hat{\boldsymbol{\vartheta}}_{\mathrm{v}}(k) \end{bmatrix}$. If $k < L$, increase k by one, and go to Step 2; otherwise, stop, and obtain the parameter estimation vector $\hat{\boldsymbol{\theta}}(L)$.

4. Examples

Example 1. *Consider the following state-space system with moving average noise,*

$$\begin{aligned}
\boldsymbol{x}(k+1) &= \begin{bmatrix} 0 & 1 \\ g_1 & g_2 \end{bmatrix} \boldsymbol{x}(k) + \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} u(k) = \begin{bmatrix} 0 & 1 \\ -0.90 & 0.80 \end{bmatrix} \boldsymbol{x}(k) + \begin{bmatrix} 1.10 \\ -1.60 \end{bmatrix} u(k), \\
y(k) &= [1, 0]\boldsymbol{x}(k) + e(k), \\
e(k) &= (1 + e_1 z^{-1} + e_2 z^{-2})v(k) = (1 + 0.20z^{-1} - 0.60z^{-1})v(k).
\end{aligned}$$

The parameter vector to be estimated is:

$$\boldsymbol{\theta} = [g_1, g_2, h_1, h_2, e_1, e_2]^{\mathrm{T}} = [-0.90, 0.80, 1.10, -1.60, 0.20, -0.60]^{\mathrm{T}}.$$

In simulation, the input $\{u(k)\}$ is set as a persistent excitation sequence and $\{v(k)\}$ as a zero-mean noise sequence with variance $\delta = 0.25$. Take the data length $L = 1200$. Use the first 1000 data and apply the BC-PSE algorithm to get the estimates of the parameter vector $\boldsymbol{\theta}$ and the system state $\boldsymbol{x}(k)$. The parameter estimates and their errors under different noise variances are displayed in Table 2 and Figure 1, where the estimation error $\tau := \|\hat{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}\|/\|\boldsymbol{\theta}\| \times 100\%$. The parameter estimate $\hat{\boldsymbol{\theta}}(k)$ versus k is plotted in Figure 2, and the state estimate $\hat{\boldsymbol{x}}(k)$ computed by the BC-PSE algorithm is illustrated in Figures 3 and 4.

The remaining 200 data from $k = 1001$–1200 are taken to test the effectiveness of the estimated model. As a comparison, the curves of the estimated output $\hat{y}(k)$ and the predicted output $y(k)$ are depicted in Figure 5.

Table 2. The bias compensation-based parameter estimates and errors for Example 1.

δ	k	g_1	g_2	h_1	h_2	e_1	e_2	$\tau(\%)$
1.00	20	−0.90672	0.58805	2.47895	−2.33308	0.12542	−1.41696	74.94801
	50	−0.91146	0.77864	1.56458	−1.45439	0.06347	−0.68669	21.66474
	100	−0.89769	0.78062	1.45065	−1.56252	0.11477	−0.61555	15.33989
	200	−0.89947	0.79122	1.18977	−1.58852	0.18174	−0.62250	4.02568
	500	−0.90289	0.79326	1.19050	−1.56703	0.17911	−0.62993	4.35605
	1000	−0.89502	0.79742	1.14729	−1.63721	0.18966	−0.61945	2.71347
0.25	20	−0.89902	0.68391	1.87036	−1.86695	0.37006	−0.70475	35.74537
	50	−0.90830	0.79067	1.34730	−1.49092	0.18690	−0.51892	11.92791
	100	−0.90003	0.79044	1.28652	−1.56142	0.20501	−0.50891	8.91773
	200	−0.90083	0.79596	1.15147	−1.58293	0.22360	−0.53086	3.84174
	500	−0.90199	0.79695	1.14809	−1.57856	0.20617	−0.56528	2.67739
	1000	−0.89776	0.79869	1.12493	−1.61634	0.20773	−0.57641	1.63983
True values		−0.90000	0.80000	1.10000	−1.60000	0.20000	−0.60000	0.00000

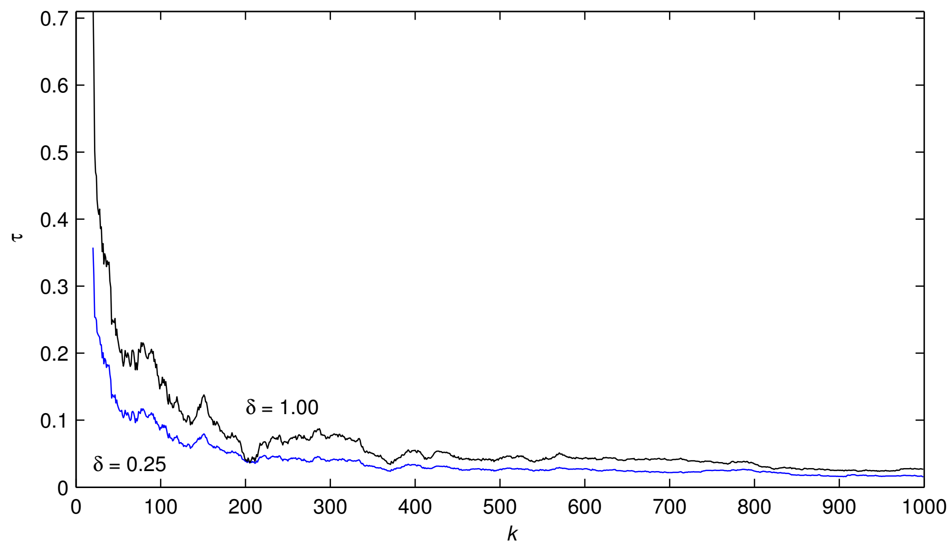

Figure 1. The estimation errors τ versus k with different noise variances for Example 1.

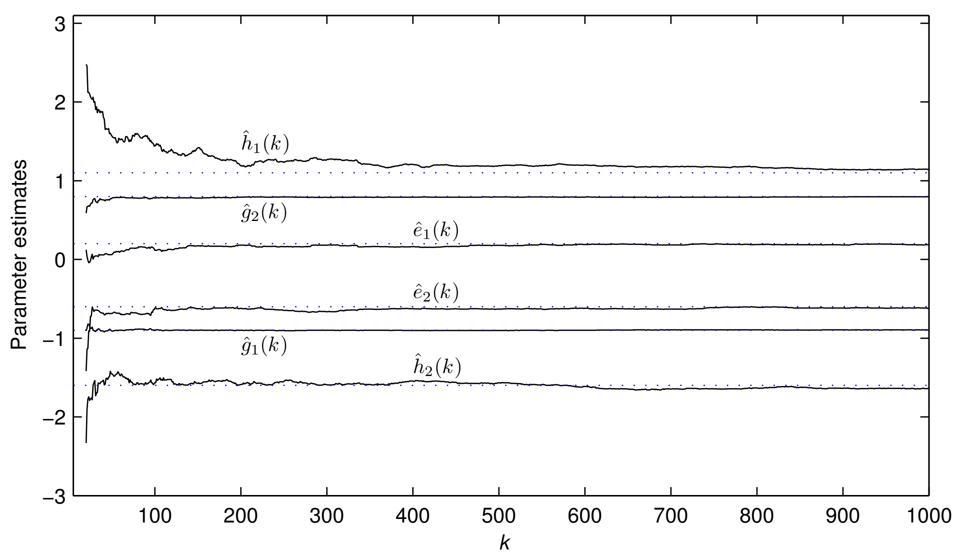

Figure 2. The parameter estimates $\hat{\boldsymbol{\theta}}(k)$ versus k for Example 1.

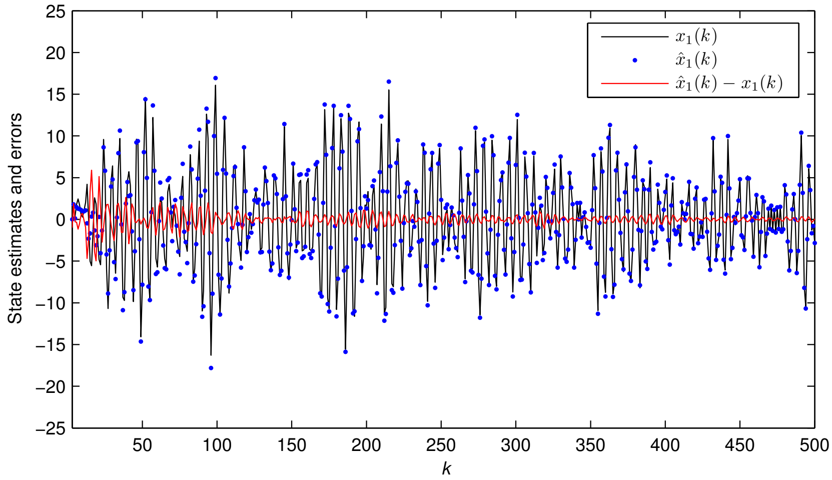

Figure 3. The state $x_1(k)$ and the estimated state $\hat{x}_1(k)$ versus k for Example 1 ($\delta = 0.25$).

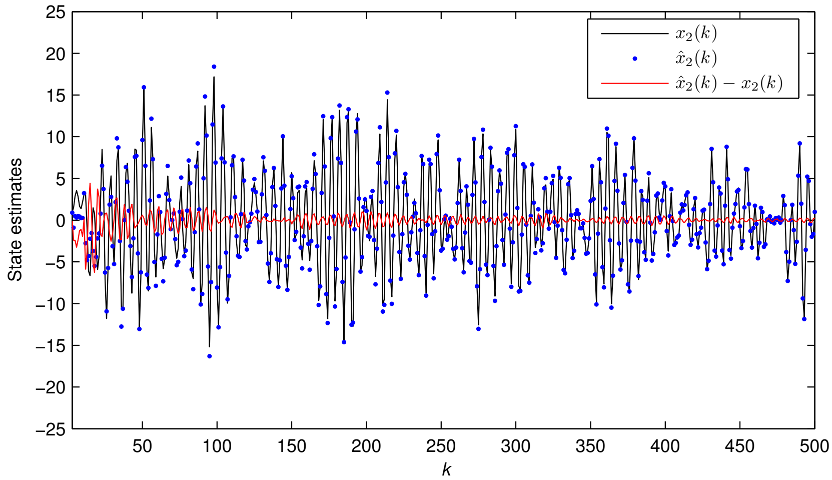

Figure 4. The state $x_2(k)$ and the estimated state $\hat{x}_2(k)$ versus k for Example 1 ($\delta = 0.25$).

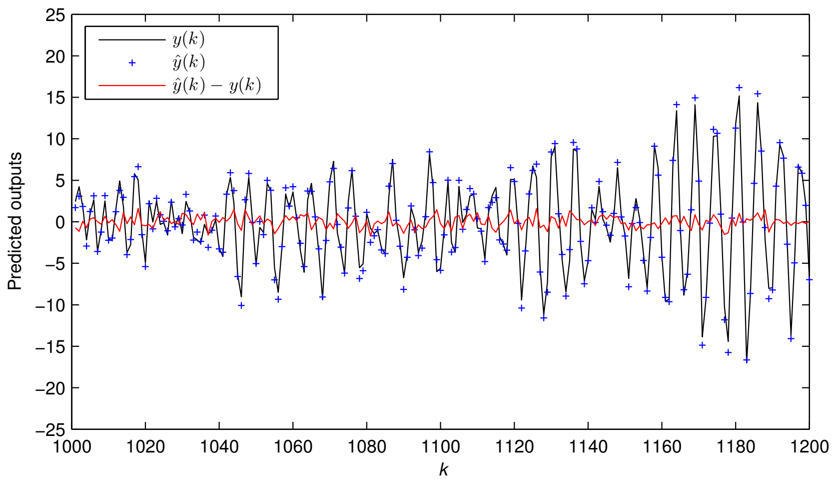

Figure 5. The actual output $y(k)$ and the estimated output $\hat{y}(k)$ versus k for Example 1.

Example 2. *Consider a third-order observability canonical state-space system,*

$$\boldsymbol{x}(k+1) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -0.80 & 0.50 & 0.60 \end{bmatrix} \boldsymbol{x}(k) + \begin{bmatrix} 2.20 \\ 0.30 \\ 0.20 \end{bmatrix} u(k),$$
$$y(k) = [1, 0, 0]\boldsymbol{x}(k) + e(k),$$
$$e(k) = (1 + e_1 z^{-1} + e_2 z^{-2})v(k) = (1 + 0.50z^{-1} + 0.40z^{-1})v(k).$$

The parameter vector to be estimated is:

$$\boldsymbol{\theta} = [g_1, g_2, g_3, h_1, h_2, h_3, e_1, e_2]^{\mathrm{T}} = [-0.80, 0.50, 0.60, 2.20, 0.30, 0.20, 0.50, 0.40]^{\mathrm{T}}.$$

The simulation conditions are the same as those in Example 1. Table 3 and Figure 6 compare the bias compensation-based parameter estimates and errors under noise variances $\delta = 0.25$ and $\delta = 1.00$. Figures 7 and 8 depict the curves of the parameter estimates $\hat{\boldsymbol{\theta}}(k)$ versus the instant k.

Table 3. The bias compensation-based parameter estimates and errors for Example 2.

δ	k	g_1	g_2	g_3	h_1	h_2	h_3	e_1	e_2	τ (%)
1.00	20	−0.62874	0.23404	0.69918	2.71399	−0.51430	0.03521	0.55758	0.54550	40.52480
	50	−0.63162	0.17455	0.82387	2.36896	−0.10338	0.17046	0.57261	0.36257	24.04086
	100	−0.72920	0.37705	0.69853	2.10786	0.00205	0.03490	0.53153	0.47341	15.57075
	200	−0.77990	0.43541	0.64826	2.07686	0.11157	0.16642	0.52363	0.43391	9.54558
	500	−0.80210	0.49447	0.60384	2.19878	0.24877	0.16126	0.56242	0.43637	3.76452
	1000	−0.80600	0.50437	0.59777	2.23514	0.29924	0.19105	0.52608	0.42701	2.04986
0.25	20	−0.70399	0.32858	0.67496	2.51439	−0.09217	0.18308	0.56489	0.42596	21.34339
	50	−0.69558	0.29717	0.73981	2.28977	0.11809	0.20666	0.54171	0.38144	13.16282
	100	−0.77049	0.44740	0.64636	2.15712	0.16039	0.13532	0.48821	0.46680	7.35863
	200	−0.79424	0.47428	0.62275	2.14390	0.21064	0.19565	0.49299	0.43757	4.56623
	500	−0.80223	0.49921	0.60237	2.20136	0.27695	0.18522	0.53537	0.42783	2.05074
	1000	−0.80385	0.50317	0.59907	2.21845	0.30107	0.19783	0.51498	0.41939	1.21121
True values		−0.80000	0.50000	0.60000	2.20000	0.30000	0.20000	0.50000	0.40000	0.00000

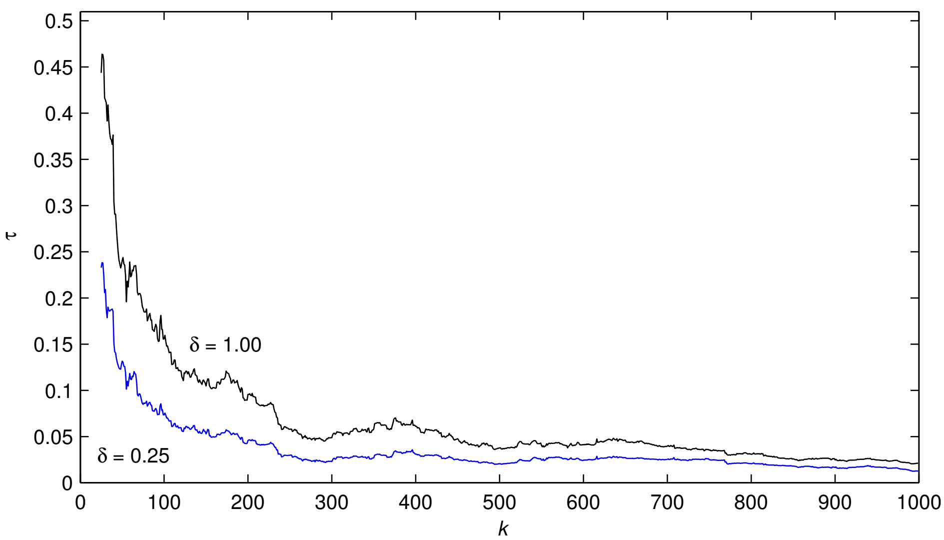

Figure 6. The estimation errors τ versus k with different noise variances for Example 2.

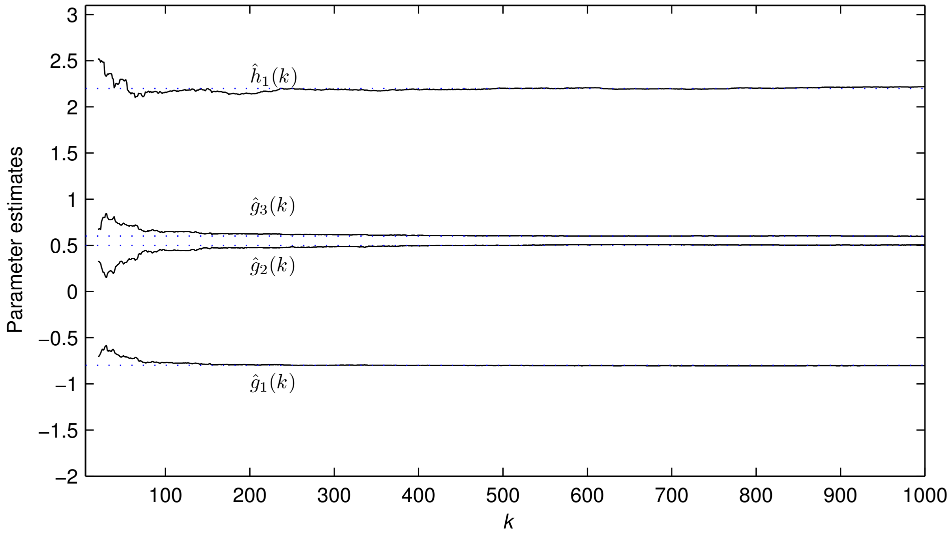

Figure 7. The parameter estimates $\hat{g}_1(k)$, $\hat{g}_2(k)$, $\hat{g}_3(k)$ and $\hat{h}_1(k)$ versus k for Example 2.

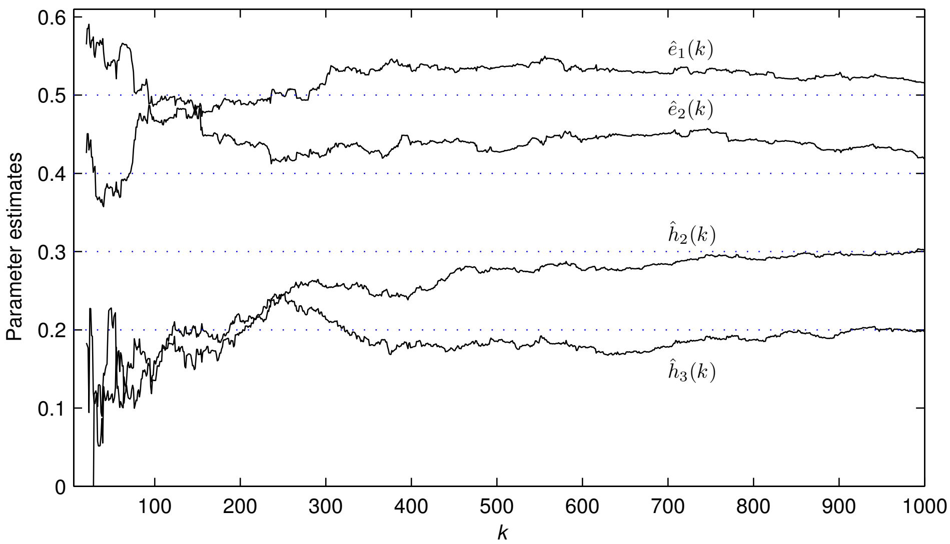

Figure 8. The parameter estimates $\hat{h}_2(k)$, $\hat{h}_3(k)$, $\hat{e}_1(k)$ and $\hat{e}_2(k)$ versus k for Example 2.

Figures 9–11 plot the dynamics of three state estimates $\hat{\boldsymbol{x}}_i(k),\ i = 1, 2, 3$. Figure 12 describes the output comparison of $\hat{y}(k)$ with $y(k)$.

From Tables 2 and 3 and Figures 1–12, we can draw some conclusions as follows.

- The parameter estimation errors given by the BC-PSE algorithm become small with k increasing; see Tables 2 and 3 and Figures 1 and 6.
- A low noise variance leads to higher accuracy of parameter estimates; see Figures 1 and 6.
- As the data length k increases, the parameter estimates approach their true values; see Figures 2, 7 and 8.
- The estimated states are close to the actual states of the system model; see Figures 3, 4 and 9–11.
- The predicted outputs match well with the actual outputs; see Figures 5 and 12.

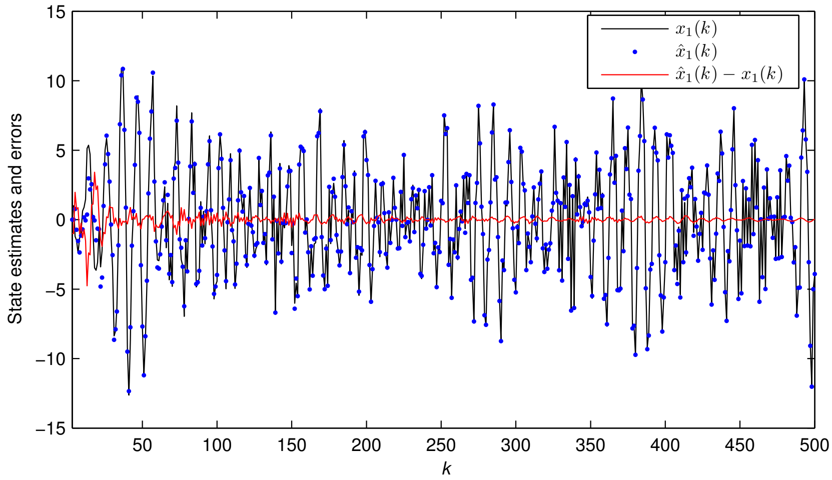

Figure 9. The true state $x_1(k)$ and the estimated state $\hat{x}_1(k)$ versus k for Example 2 ($\delta = 0.25$).

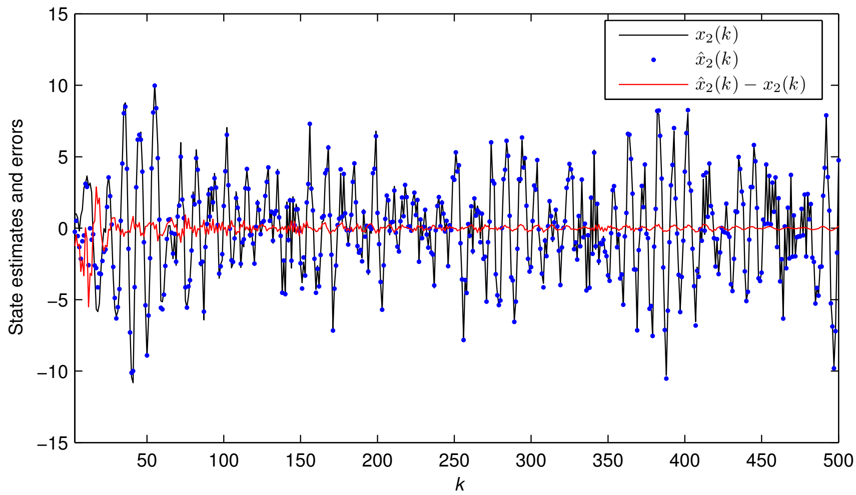

Figure 10. The true state $x_2(k)$ and the estimated state $\hat{x}_2(k)$ versus k for Example 2 ($\delta = 0.25$).

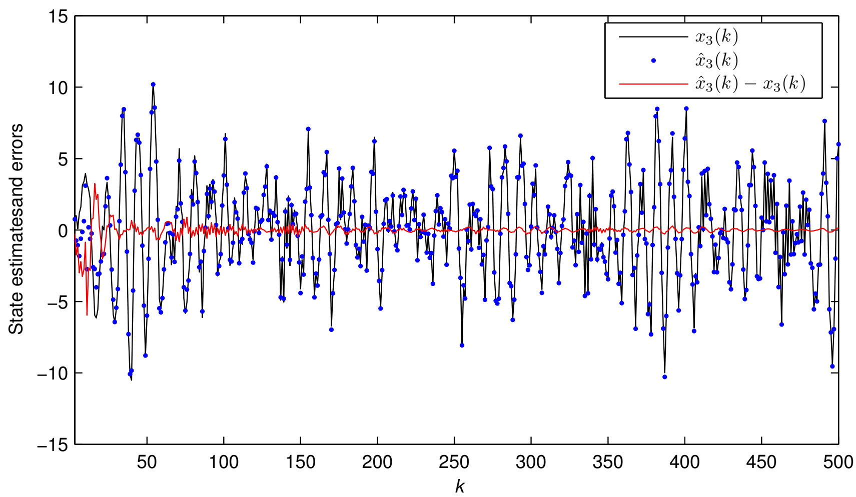

Figure 11. The true state $x_3(k)$ and the estimated state $\hat{x}_3(k)$ versus k for Example 2 ($\delta = 0.25$).

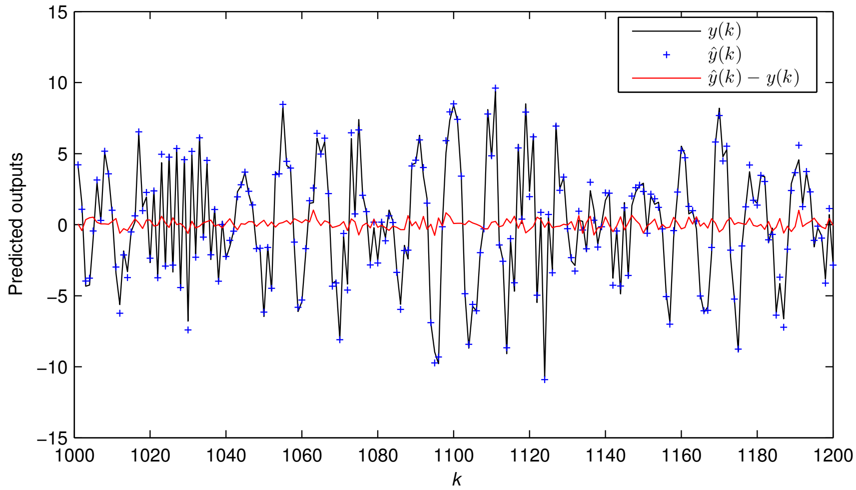

Figure 12. The actual output $y(k)$ and the estimated output $\hat{y}(k)$ versus k for Example 2.

5. Conclusions

This paper discusses the identification of the observability canonical state-space system with colored noise via our proposed bias compensation-based parameter and state estimation algorithm. The numerical results indicate that the algorithm can effectively estimate the system states and parameters. The advantage of this algorithm is that the parameter estimates are unbiased. The algorithm can be combined with other recursive algorithms, such as the multi-innovation algorithm, to study the identification of nonlinear state space systems [39], dual-rate systems [40], signal modeling [41] and time series analysis [42,43].

Author Contributions: Joint work.

Funding: This work was supported by the Science and Technology Project of Henan Province (China, 182102210536, 182102210538), the Key Research Project of Henan Higher Education Institutions (China, 18A120003, 18A130001, 18B520036), the National Science Foundation of China (China, 61503122) and Nanhu Scholars Program for Young Scholars of XYNU(China).

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

References

1. Na, J.; Yang, J.; Wu, X.; Guo, Y. Robust adaptive parameter estimation of sinusoidal signals. *Automatica* **2015**, *53*, 376–384. [CrossRef]
2. Kalafatis, A.D.; Wang, L.; Cluett, W.R. Identification of time-varying pH processes using sinusoidal signals. *Automatica* **2005**, *41*, 685–691. [CrossRef]
3. Na, J.; Yang, J.; Ren, X.M.; Guo, Y. Robust adaptive estimation of nonlinear system with time-varying parameters. *Int. J. Adapt. Control Process.* **2015**, *29*, 1055–1072. [CrossRef]
4. Liu, S.Y.; Xu, L.; Ding, F. Iterative parameter estimation algorithms for dual-frequency signal models. *Algorithms* **2017**, *10*, 118. [CrossRef]
5. Na, J.; Mahyuddin, M.N.; Herrmann, G.; Ren, X.; Barber, P. Robust adaptive finite-time parameter estimation and control for robotic systems. *Int. J. Robust Nonlinear Control* **2015**, *25*, 3045–3071. [CrossRef]
6. Huang, W.; Ding, F. Coupled least squares identification algorithms for multivariate output-error systems. *Algorithms* **2017**, *10*, 12. [CrossRef]
7. Goos, J.; Pintelon, R. Continuous-time identification of periodically parameter-varying state space models. *Automatica* **2016**, *71*, 254–263. [CrossRef]
8. AlMutawa, J. Identification of errors-in-variables state space models with observation outliers based on minimum covariance determinant. *J. Process Control* **2009**, *19*, 879–887. [CrossRef]
9. Yuan, Y.; Zhang, H.; Wu, Y.; Zhu, T.; Ding, H. Bayesian learning-based model predictive vibration control for thin-walled workpiece machining processes. *IEEE/ASME Trans. Mechatron.* **2017**, *22*, 509–520. [CrossRef]

10. Ding, F. Combined state and least squares parameter estimation algorithms for dynamic systems. *Appl. Math. Model.* **2014**, *38*, 403–412. [CrossRef]
11. Ma, J.X.; Xiong, W.L.; Chen, J.; Ding, F. Hierarchical identification for multivariate Hammerstein systems by using the modified Kalman filter. *IET Control Theory Appl.* **2017**, *11*, 857–869. [CrossRef]
12. Fatehi, A.; Huang, B. Kalman filtering approach to multi-rate information fusion in the presence of irregular sampling rate and variable measurement delay. *J. Process Control* **2017**, *53*, 15–25. [CrossRef]
13. Zhao, S.; Shmaliy, Y.S.; Liu, F. Fast Kalman-like optimal unbiased FIR filtering with applications. *IEEE Trans. Signal Process.* **2016**, *64*, 2284–2297. [CrossRef]
14. Zhou, Z.P.; Liu, X.F. State and fault estimation of sandwich systems with hysteresis. *Int. J. Robust Nonlinear Control* **2018**, *28*, 3974–3986. [CrossRef]
15. Zhao, S.; Huang, B.; Liu, F. Linear optimal unbiased filter for time-variant systems without apriori information on initial condition. *IEEE Trans. Autom. Control* **2017**, *62*, 882–887. [CrossRef]
16. Zhao, S.; Shmaliy, Y.S.; Liu, F. On the iterative computation of error matrix in unbiased FIR filtering. *IEEE Signal Process. Lett.* **2017**, *24*, 555–558. [CrossRef]
17. Erazo, K.; Nagarajaiah, S. An offline approach for output-only Bayesian identification of stochastic nonlinear systems using unscented Kalman filtering. *J. Sound Vib.* **2017**, *397*, 222–240. [CrossRef]
18. Verhaegen, M.; Verdult, V. *Filtering and System Identification: A Least Squares Approach*; Cambridge University Press: Cambridge, UK, 2007.
19. Ljung, L. *System Identification: Theory for the User*, 2nd ed.; Prentice Hall: Englewood Cliffs, NJ, USA, 1999.
20. Yu, C.P.; Ljung, L.; Verhaegen, M. Identification of structured state-space models. *Automatica* **2018**, *90*, 54–61. [CrossRef]
21. Naitali, A.; Giri, F. Persistent excitation by deterministic signals for subspace parametric identification of MISO Hammerstein systems. *IEEE Trans. Autom. Control* **2016**, *61*, 258–263. [CrossRef]
22. Ase, H.; Katayama, T. A subspace-based identification of Wiener-Hammerstein benchmark model. *Control Eng. Pract.* **2015**, *44*, 126–137. [CrossRef]
23. 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]
24. 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]
25. Pavelkova, L.; Karny, M. State and parameter estimation of state-space model with entry-wise correlated uniform noise. *Int. J. Adapt. Control Signal Process.* **2014**, *28*, 1189–1205. [CrossRef]
26. Ma, J.X.; Wu, O.Y.; Huang, B.; Ding, F. Expectation maximization estimation for a class of input nonlinear state space systems by using the Kalman smoother. *Signal Process.* **2018**, *145*, 295–303. [CrossRef]
27. Li, J.H.; Zheng, W.X.; Gu, J.P.; Hua, L. A recursive identification algorithm for Wiener nonlinear systems with linear state-space subsystem. *Circuits Syst. Signal Process.* **2018**, *37*, 2374–2393. [CrossRef]
28. Wang, H.W.; Liu, T. Recursive state-space model identification of non-uniformly sampled systems using singular value decomposition. *Chin. J. Chem. Eng.* **2014**, *22*, 1268–1273. [CrossRef]
29. Ding, J.L. Data filtering based recursive and iterative least squares algorithms for parameter estimation of multi-input output systems. *Algorithms* **2016**, *9*, 49. [CrossRef]
30. Yu, C.P.; You, K.Y.; Xie, L.H. Quantized identification of ARMA systems with colored measurement noise. *Automatica* **2016**, *66*, 101–108. [CrossRef]
31. Jafari, M.; Salimifard, M.; Dehghani, M. Identification of multivariable nonlinear systems in the presence of colored noises using iterative hierarchical least squares algorithm. *ISA Trans.* **2014**, *53*, 1243–1252. [CrossRef] [PubMed]
32. Sagara, S.; Wada, K. On-line modified least-squares parameter estimation of linear discrete dynamic systems. *Int. J. Control* **1977**, *25*, 329–343. [CrossRef]
33. Mejari, M.; Piga, D.; Bemporad, A. A bias-correction method for closed-loop identification of linear parameter-varying systems. *Automatica* **2018**, *87*, 128–141. [CrossRef]
34. Ding, J.; Ding, F. Bias compensation based parameter estimation for output error moving average systems. *Int. J. Adapt. Control Signal Process.* **2011**, *25*, 1100–1111. [CrossRef]
35. Diversi, R. Bias-eliminating least-squares identification of errors-in-variables models with mutually correlated noises. *Int. J. Adapt. Control Signal Process.* **2013**, *27*, 915–924. [CrossRef]

36. Zhang, Y. Unbiased identification of a class of multi-input single-output systems with correlated disturbances using bias compensation methods. *Math. Comput. Model.* **2011**, *53*, 1810–1819. [CrossRef]
37. Zheng, W.X. A bias correction method for identification of linear dynamic errors-in-variables models. *IEEE Trans. Autom. Control* **2002**, *47*, 1142–1147. [CrossRef]
38. Wang, X.H.; Ding, F.; Alsaedi, A.; Hayat, T. Filtering based parameter estimation for observer canonical state space systems with colored noise. *J. Frankl. Inst.* **2017**, *354*, 593–609. [CrossRef]
39. Pan, W.; Yuan, Y.; Goncalves, J.; Stan, G.B. A sparse Bayesian approach to the identification of nonlinear state-space systems. *IEEE Trans. Autom. Control* **2016**, *61*, 182–187. [CrossRef]
40. Yang, X.Q.; Yang, X.B. Local identification of LPV dual-rate system with random measurement delays. *IEEE Trans. Ind. Electron.* **2018**, *65*, 1499–1507. [CrossRef]
41. 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]
42. 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] [PubMed]
43. Chen, G.Y.; Gan, M.; Chen, G.L. Generalized exponential autoregressive models for nonlinear time series: Stationarity, estimation and applications. *Inf. Sci.* **2018**, *438*, 46–57. [CrossRef]