



ARTICLE

State Estimation Moving Window Gradient Iterative Algorithm for Bilinear Systems Using the Continuous Mixed p -norm Technique

Wentao Liu, Junxia Ma and Weili Xiong*

Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi, 214122, China

*Corresponding Author: Weili Xiong. Email: greenpre@163.com

Received: 30 November 2021 Accepted: 28 March 2022

ABSTRACT

This paper studies the parameter estimation problems of the nonlinear systems described by the bilinear state space models in the presence of disturbances. A bilinear state observer is designed for deriving identification algorithms to estimate the state variables using the input-output data. Based on the bilinear state observer, a novel gradient iterative algorithm is derived for estimating the parameters of the bilinear systems by means of the continuous mixed p -norm cost function. The gain at each iterative step adapts to the data quality so that the algorithm has good robustness to the noise disturbance. Furthermore, to improve the performance of the proposed algorithm, a dynamic moving window is designed which can update the dynamical data by removing the oldest data and adding the newest measurement data. A numerical example of identification of bilinear systems is presented to validate the theoretical analysis.

KEYWORDS

Bilinear state space model; parameter estimation; moving window; continuous mixed p -norm

1 Introduction

Although many physical dynamic behaviors are characterized as linear systems in the neighborhood of a single operating point, when they exhibit strong nonlinearities or must be described over the entire operating range, linear models may not yield appropriate results [1–3]. Owing to that, the study of system identification and parameter estimation for nonlinear systems has drawn considerable attention of academic researchers [4–6]. Various models are exploited to describe actual nonlinear systems with relatively simple structures [7–11] such as Hammerstein systems, Bilinear systems, and Wiener systems.

The parameter estimation of the system models is important for control system analysis and design. The parameters of the models can be estimated by using some identification methods [12–15] such as the hierarchical algorithms [16–19]. In this paper, we confine our discussion to bilinear systems, which are simplicity in model structures and capable of approaching arbitrary nonlinear systems with much higher accuracy than traditional linear approximations theoretically [20]. Bruni et al. designed a bilinear time-series model-based self-tuning controller for a multi-machine power system to enhance



the region of stability of the system, and return the states to their stable equilibrium [21]. Yeo et al. developed the bilinear model predictive control algorithm for paper plant systems to control the grade change operations in paper production [22]. Wang et al. simulated the bridge nonlinear boundary as bilinear translational, and applied the nonlinear least squares optimization algorithm to identify the nonlinear translational and rotational boundary parameters of the bridge [23].

Many researches have been studied in the parameter estimation of bilinear state space systems. A stochastic gradient algorithm and a gradient-based iterative algorithm have been proposed for the parameter identification of bilinear systems by using the auxiliary model. The gradient-based iterative algorithm uses fixed batch data to update the parameters, so that the parameter estimation accuracy can be greatly improved compared with the stochastic gradient algorithm. Li et al. combined the maximum likelihood theory with the data filtering technique for bilinear systems with colored noises. The state observer is vital in the field of state estimation of bilinear systems [24]. Tsai considered a H-infinity fuzzy observer for bilinear systems by means of a linear matrix inequality approach [25]. Phan et al. presented a full-order bilinear state observer for the bilinear system, which optimized the observer gain by interaction matrices [26]. Zhang et al. considered the state estimation problem of bilinear systems and proposed a state filtering method for the single-input single-output bilinear systems and multiple-input multiple-output bilinear systems by minimizing the covariance matrix of the state estimation errors [27]. Recently, some state and parameter estimation methods have been proposed for linear and bilinear state space systems in the presence of exogenous noises [28–30].

However, in industrial applications, the dynamical processes often work on various noise environments, and some works used nonlinear filtering technique such as median filtering and least mean p -norm to overcome the problems. Zayyani proposed a mixed-norm adaptive filter algorithm for systems identification based on minimization of the logarithmic continuous mixed p -norm [31]. Moreover, an improved mixed p -norm algorithm had been studied to combat non-Gaussian interference and its computational complexity as well as the mean convergence had also been analyzed [32,33].

Inspired by the above researches, we study the parameter estimation algorithms for bilinear state space models with the noise disturbances. Based on the iterative search and the state estimator, a bilinear state continues mixed p -norm gradient iteration (BSO-CMPN-GI) is proposed. The proposed algorithms control the proportions of the error norms and offer an extra degree of freedom within the adaptation. By taking into consideration each p -norm of errors for $1 \leq p \leq 2$, the proposed algorithms combine the benefits of the variable error norms and thus are more robust against noise interference. Furthermore, the moving data window theory is introduced to improve the effectiveness of the proposed algorithms. The contributions of this paper are summarized as follows:

- A state observer is designed to obtain the system states variables consisting of the product terms of state variables and control variables.
- A bilinear state observer-based continues mixed p -norm gradient iteration (BSO-CMPN-GI) algorithm is presented for the bilinear state space system under the noise interference to estimate the unknown system parameters.
- A bilinear state observer-based moving window CMPN-GI (BSO-MW-CMPN-GI) algorithm is derived to update collected data and thus maintain high data utilization.

The outlines of this paper are organized as follows. Section 2 introduces some definitions and proposes the identification model of the bilinear state space system and introduces a bilinear state observer for the state estimation. Sections 3 and 4 derive a BSO-CMPN-GI and a MW-CMPN-GI

algorithms based on the bilinear state estimator, respectively. An example to illustrate the effectiveness of the proposed algorithms in Section 5. Finally, Section 6 gives some concluding remarks.

2 System Description and Identification Model

Consider the following single-input single-output bilinear state space model:

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{x}_t u_t + \mathbf{g}u_t + \mathbf{w}_t, \tag{1}$$

$$y_t = \mathbf{c}\mathbf{x}_t + v_t, \tag{2}$$

where $u_t \in \mathbb{R}^n$ is the system input variables, $y_t \in \mathbb{R}^n$ is the measurement output variables, \mathbf{x}_t is the state vector, \mathbf{w}_t is an uncorrelated process noise with zero mean, and v_t is an uncorrelated measurement noise with zero mean. $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times n}$, $\mathbf{g} \in \mathbb{R}^n$ and $\mathbf{c} \in \mathbb{R}^{1 \times n}$ are system parameter matrix and parameter vector. Consider the observer canonical form of bilinear systems, the system parameters can be expressed as

$$\mathbf{A} := \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ -a_{n-1} & 0 & \cdots & 0 & 1 \\ -a_n & 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}, \mathbf{B} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix} \in \mathbb{R}^{n \times n}, \mathbf{g} := \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_{n-1} \\ g_n \end{bmatrix} \in \mathbb{R}^n, \tag{3}$$

$$\mathbf{c} := [1, 0, \dots, 0, 0] \in \mathbb{R}^{1 \times n},$$

where $\mathbf{b}_j := [b_{j1}, b_{j2}, \dots, b_{jn}] \in \mathbb{R}^{1 \times n}$, $j = 1, 2, \dots, n$.

Assumption 1: The dimension n of the system state vector is known, $u_t = 0$, $\mathbf{x}_t = 0$, $y_t = 0$, $\mathbf{w}_t = 0$ and $v_t = 0$ for $t \leq 0$.

Assumption 2: The bilinear system in Eqs. (1) and (2) is observable and controllable.

Eq. (1) can be written as

$$\begin{cases} x_{1,t+1} = -a_1 x_{1,t} + x_{2,t} + \mathbf{b}_1 \mathbf{x}_t u_t + g_1 u_t + w_{1,t}, \\ x_{2,t+1} = -a_2 x_{1,t} + x_{3,t} + \mathbf{b}_2 \mathbf{x}_t u_t + g_2 u_t + w_{2,t}, \\ \vdots \\ x_{n-1,t+1} = -a_{n-1} x_{1,t} + x_{n,t} + \mathbf{b}_{n-1} \mathbf{x}_t u_t + g_{n-1} u_t + w_{n-1,t}, \\ x_{n,t+1} = -a_n x_{1,t} + \mathbf{b}_n \mathbf{x}_t u_t + g_n u_t + w_{n,t}. \end{cases} \tag{4}$$

Adding both sides of the above equations has

$$\begin{aligned} x_{1,t+1} &= -a_1 x_{1,t} - a_2 x_{1,t-1} - \cdots - a_n x_{1,t-n+1} + \mathbf{b}_1 \mathbf{x}_t u_t \\ &\quad + \mathbf{b}_2 \mathbf{x}_{t-1} u_{t-1} + \cdots + \mathbf{b}_n \mathbf{x}_{t-n+1} u_{t-n+1} + g_1 u_t \\ &\quad + g_2 u_{t-1} + \cdots + g_n u_{t-n+1} + w_{1,t-t} + w_{2,t-1} + \cdots + w_{n,t-n+1} \\ &= - \sum_{j=1}^n a_j x_{1,t-j+1} + \sum_{j=1}^n \mathbf{b}_j \mathbf{x}_{t-j+1} u_{t-j+1} + \sum_{j=1}^n g_j u_{t-j+1} + \sum_{j=1}^n w_{j,t-j+1}. \end{aligned} \tag{5}$$

Substituting Eq. (5) into Eq. (2) obtains the regression form of the bilinear state model

$$y_t = - \sum_{j=1}^n a_j x_{1,t-j} + \sum_{j=1}^n \mathbf{b}_j \mathbf{x}_{t-j} u_{t-j} + \sum_{j=1}^n g_j u_{t-j} + \sum_{j=1}^n w_{j,t-j} + v_t. \tag{6}$$

Define the information vector $\boldsymbol{\psi}_t$ and parameter vector $\boldsymbol{\vartheta}$ as

$$\boldsymbol{\psi}_t := [\mathbf{d}_t, \mathbf{k}_t, \mathbf{h}_t]^T \in \mathbb{R}^{n^2+2n}, \quad (7)$$

$$\mathbf{d}_t := [-x_{1,t-1}, -x_{1,t-2}, \dots, -x_{1,t-n}]^T \in \mathbb{R}^n, \quad (8)$$

$$\mathbf{k}_t := [\mathbf{x}_{t-1}^T \mathbf{u}_{t-1}, \mathbf{x}_{t-2}^T \mathbf{u}_{t-2}, \dots, \mathbf{x}_{t-n}^T \mathbf{u}_{t-n}]^T \in \mathbb{R}^{n^2}, \quad (9)$$

$$\mathbf{h}_t := [u_{t-1}, u_{t-2}, \dots, u_{t-n}]^T \in \mathbb{R}^n, \quad (10)$$

$$\boldsymbol{\vartheta} := [\mathbf{a}, \mathbf{B}, \mathbf{g}]^T \in \mathbb{R}^{n^2+2n}, \quad (11)$$

$$\mathbf{a} := [a_1, a_2, \dots, a_n]^T \in \mathbb{R}^n, \quad (12)$$

$$\mathbf{B} := [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n]^T \in \mathbb{R}^{n^2}, \quad (13)$$

$$\mathbf{g} := [g_1, g_2, \dots, g_n]^T \in \mathbb{R}^n. \quad (14)$$

Redefine the noise term as

$$e_t := \sum_{j=1}^n w_{j,t-j} + v_t. \quad (15)$$

Eq. (6) can be rewritten as

$$y_t = \boldsymbol{\psi}_t^T \boldsymbol{\vartheta} + e_t. \quad (16)$$

The proposed parameter estimation algorithms in this paper are based on this identification model in (16). Many identification methods are derived based on the identification models of the systems [34–43] and these methods can be used to estimate the parameters of other linear systems and nonlinear systems [44–50] and can be applied to other fields such as chemical process control systems. The purpose of this paper is to obtain the estimates of the unknown parameters in the bilinear state space model by means of the measurement data $\{u_t, y_t\}$. However, the state space model contains not only unknown parameters, but also unknown state variables. That means the information vector $\boldsymbol{\psi}_t$ in Eq. (16) contains the unknown state vector \mathbf{x}_{t-i} . Given this, we adopt the state observer for estimating the unknown states variables [51]. According to the bilinear state space model, the state observer is designed as

$$\hat{\mathbf{x}}_{t+1} = \hat{\mathbf{A}}\hat{\mathbf{x}}_t + \hat{\mathbf{B}}\hat{\mathbf{x}}_t u_t + \hat{\mathbf{g}}u_t, \quad (17)$$

where

$$\hat{\mathbf{A}} = \begin{bmatrix} -\hat{a}_1 & 1 & 0 & \cdots & 0 \\ -\hat{a}_2 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ -\hat{a}_{n-1} & 0 & \cdots & 0 & 1 \\ -\hat{a}_n & 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}, \hat{\mathbf{B}} = \begin{bmatrix} \hat{\mathbf{b}}_1 \\ \hat{\mathbf{b}}_2 \\ \vdots \\ \hat{\mathbf{b}}_{n-1} \\ \hat{\mathbf{b}}_n \end{bmatrix} \in \mathbb{R}^{n \times n}, \hat{\mathbf{g}} = \begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \\ \vdots \\ \hat{g}_{n-1} \\ \hat{g}_n \end{bmatrix} \in \mathbb{R}^n.$$

Remark 1: Differently from the identification of bilinear-in-parameter systems which involve the product terms of the parameter vectors and the information matrices, this paper studies the identification problem of bilinear state-space systems with the product terms of state variables and control variables, which makes it difficult for the parameter and state estimation.

3 Bilinear State Observer Based Continuous Mixed p -norm Gradient Iterative Identification Algorithm

Set the data length be L , based on the input and output data $\{u_t, y_t, 1 \leq t \leq L\}$, define the stacked input vector \mathbf{Y}_L and the stacked information matrix Ψ_L as

$$\mathbf{Y}_L := \begin{bmatrix} y_L \\ y_{L-1} \\ \vdots \\ y_t \\ \vdots \\ y_1 \end{bmatrix} \in \mathbb{R}^L, \quad \Psi_L := \begin{bmatrix} \psi_L^T \\ \psi_{L-1}^T \\ \vdots \\ \psi_t^T \\ \vdots \\ \psi_1^T \end{bmatrix} \in \mathbb{R}^{L \times (n^2+2n)}.$$

To suppress the effect of noise interference, define the continuous mixed p -norm cost function

$$J_1(\vartheta) := \int_1^2 \lambda(p) |\mathbf{V}_L|^p dp, \tag{18}$$

where $\lambda(p)$ is the probability density-like weighting function which is constrained by $\int_1^2 \lambda(p) dp = 1$, and $\mathbf{V}_L := \mathbf{Y}_L - \Psi_L \vartheta$.

Using the negative gradient search method to minimize the cost function $J_1(\vartheta)$ with respect to ϑ gives

$$\text{grad}[J_1(\vartheta)] = \frac{\partial J_1(\vartheta)}{\partial \vartheta} = \begin{bmatrix} \frac{\partial J_1(\vartheta)}{\partial \mathbf{a}} \\ \frac{\partial J_1(\vartheta)}{\partial \mathbf{B}} \\ \frac{\partial J_1(\vartheta)}{\partial \mathbf{g}} \end{bmatrix} \tag{19}$$

where

$$\begin{aligned} \frac{\partial J_1(\vartheta)}{\partial \mathbf{a}} &= \int_1^2 p \lambda(p) |\mathbf{V}_L|^{p-1} \text{sgn}(\mathbf{V}_L) \frac{\partial(\mathbf{V}_L)}{\partial \mathbf{a}} dp \\ &= - \int_1^2 p \lambda(p) |\mathbf{V}_L|^{p-1} \text{sgn}(\mathbf{V}_L) dp [\mathbf{d}_L, \mathbf{d}_{L-1}, \dots, \mathbf{d}_t, \dots, \mathbf{d}_1]^T, \\ \frac{\partial J_1(\vartheta)}{\partial \mathbf{B}} &= - \int_1^2 p \lambda(p) |\mathbf{V}_L|^{p-1} \text{sgn}(\mathbf{V}_L) dp [\mathbf{k}_L, \mathbf{k}_{L-1}, \dots, \mathbf{k}_t, \dots, \mathbf{k}_1]^T, \\ \frac{\partial J_1(\vartheta)}{\partial \mathbf{g}} &= - \int_1^2 p \lambda(p) |\mathbf{V}_L|^{p-1} \text{sgn}(\mathbf{V}_L) dp [\mathbf{h}_L, \mathbf{h}_{L-1}, \dots, \mathbf{h}_t, \dots, \mathbf{h}_1]^T. \end{aligned}$$

Therefore, Eq. (19) can be expressed as

$$\begin{aligned} \text{grad}[J_1(\vartheta)] &= - \int_1^2 p \lambda(p) |\mathbf{V}_L|^{p-1} \text{sgn}(\mathbf{V}_L) \Psi_L dp \\ &= -\Gamma_L \text{sgn}(\mathbf{V}_L) \Psi_L, \end{aligned}$$

where

$$\Gamma_L = \int_1^2 p\lambda(p)|V_L|^{p-1}dp. \quad (20)$$

Let $s = 1, 2, 3, \dots$ be the iteration variables, $\hat{\vartheta}^s$ be the parameter estimation vector at iteration s . Using the negative gradient search and minimizing $J_1(\vartheta)$ yield

$$\begin{aligned} \hat{\vartheta}^s &= \hat{\vartheta}^{s-1} - \mu^s \text{grad}[J_1(\hat{\vartheta}^{s-1})] \\ &= \hat{\vartheta}^{s-1} + \mu^s \Gamma_L \text{sgn}(V_L) \Psi_L, \end{aligned} \quad (21)$$

where μ^s is the iteration step size [52–57], which satisfies

$$0 < \mu^s \leq \frac{2}{\lambda_{\max}[\Psi_L^T \Psi_L]}. \quad (22)$$

Eqs. (21) and (22) cannot figure out the parameter estimation vector $\hat{\vartheta}^s$, because the information matrix Ψ_L contains the unknown states x_{t-j} . The solution is to replace the unknown states x_{t-j} in the information matrix with their estimates obtained by the state observer at the previous iteration. Define the estimated information matrix as

$$\begin{aligned} \hat{\Psi}_L^s &:= [(\hat{\psi}_L^s)^T, (\hat{\psi}_{L-1}^s)^T, \dots, (\hat{\psi}_1^s)^T]^T \in \mathbb{R}^{L \times (n^2+2n)}, \\ \hat{\psi}_t^s &= [-\hat{x}_{1,t-1}^{s-1}, \dots, -\hat{x}_{1,t-n}^{s-1}, (\hat{x}_{t-1}^{s-1})^T u_{t-1}, \dots, (\hat{x}_{t-n}^{s-1})^T u_{t-n}, u_{t-1}, \dots, u_{t-n}]^T \in \mathbb{R}^{n^2+2n}, \end{aligned}$$

where \hat{x}_{t-j}^{s-1} is given by

$$\hat{x}_{t-j}^{s-1} = \hat{A}^{s-1} \hat{x}_{t-j-1}^{s-1} + \hat{B}^{s-1} \hat{x}_{t-j-1}^{s-1} u_{t-j-1} + \hat{g}^{s-1} u_{t-j-1}, \quad j = 1, 2, \dots, L.$$

Replace the information matrix Ψ_L in Eqs. (21) and (22) with its estimated value $\hat{\Psi}_L^s$, the estimate \hat{V}_L^s at iteration s is $Y_L - \hat{\Psi}_L^s \hat{\vartheta}^{s-1}$. Replacing V_L in (20) with its estimates gives

$$\hat{\Gamma}_L^s = \int_1^2 p\lambda(p)|\hat{V}_L^s|^{p-1}dp. \quad (23)$$

To obtain a closed form formula for $\hat{\Gamma}_L^s$, a uniform weighting function $\lambda(p) = 1$ is assumed. Then, the bilinear state observe-based continuous mixed p -norm gradient iterative (BSO-CMPN-GI) identification algorithm (24)–(30) for the bilinear system is summarized in the following:

$$\hat{\vartheta}^s = \hat{\vartheta}^{s-1} + \mu^s \hat{\Gamma}_L^s \text{sgn}(\hat{V}_L^s) \hat{\Psi}_L^s, \quad (24)$$

$$0 < \mu^s \leq \frac{2}{\lambda_{\max}[(\hat{\Psi}_L^s)^T \hat{\Psi}_L^s]}, \quad (25)$$

$$\hat{\Gamma}_L^s = \frac{(2|\hat{V}_L^s| - 1) \ln(|\hat{V}_L^s|) - |\hat{V}_L^s| + 1}{\ln^2(|\hat{V}_L^s|)}, \quad (26)$$

$$\hat{V}_L^s = Y_L - \hat{\Psi}_L^s \hat{\vartheta}^{s-1}, \quad (27)$$

$$\hat{\Psi}_L^s = [(\hat{\psi}_L^s)^T, (\hat{\psi}_{L-1}^s)^T, \dots, (\hat{\psi}_1^s)^T]^T, \quad (28)$$

$$\hat{\psi}_t^s = [-\hat{x}_{1,t-1}^{s-1}, \dots, -\hat{x}_{1,t-n}^{s-1}, (\hat{x}_{t-1}^{s-1})^T u_{t-1}, \dots, (\hat{x}_{t-n}^{s-1})^T u_{t-n}, u_{t-1}, \dots, u_{t-n}]^T, \quad (29)$$

$$\hat{x}_{t-j}^s = \hat{A}^s \hat{x}_{t-j-1}^s + \hat{B}^s \hat{x}_{t-j-1}^s u_{t-j-1} + \hat{g}^s u_{t-j-1}. \quad (30)$$

The flowchart of the BSO-CMPN-GI algorithm in Eqs. (24)–(30) is shown in Fig. 1.

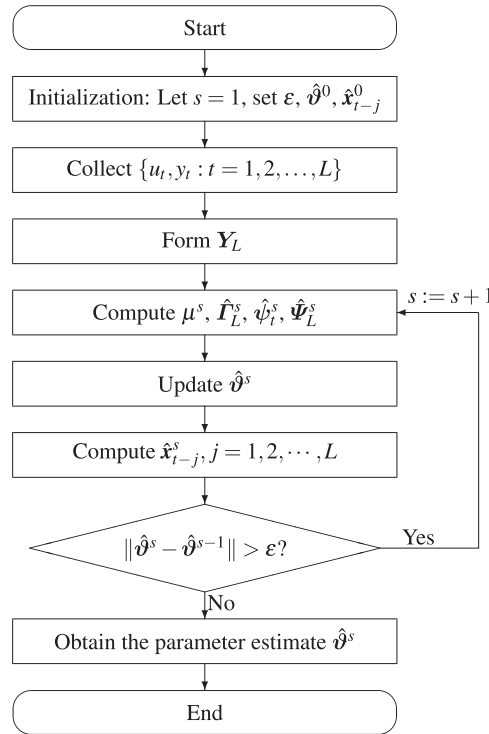


Figure 1: The flowchart of the BSO-CMPN-GI algorithm

Remark 2: The parameter p in BSO-CMPN-GI is adapted by continuous p -norm without resorting to a priori knowledge of the noise.

Remark 3: The BSO-CMPN-GI algorithm makes full use of the measurement data in each iteration of the calculation process, but it requires a batch of data to be collected in advance, and thus is implemented offline. Therefore, an on-line identification algorithm derived from the BSO-CMPN-GI algorithm will be introduced by exploiting the past and current measurement data to estimate the unknown parameters in Section 4.

4 Bilinear State Observe Moving Window Continuous Mixed p -norm Gradient Iterative Algorithm

In this section, we introduce the moving window method to derive an on-line identification algorithm and enhance the performance of the BSO-CMPN-GI algorithm. The length of the moving window is set as a fixed value m . Define the stacked output vector $Y_{t,m}$ and the stacked information matrix $\Psi_{t,m}$ as

$$Y_{t,m} := \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-m+1} \end{bmatrix} \in \mathbb{R}^h, \quad \Psi_{t,m} := \begin{bmatrix} \psi_t^T \\ \psi_{t-1}^T \\ \vdots \\ \psi_{t-m+1}^T \end{bmatrix} \in \mathbb{R}^{h \times (n^2 + 2n)}.$$

Consider the measurements from $t - m + 1$ to t and define the cost function

$$J_2(\vartheta) := \int_1^2 \lambda_t(p) |V_{t,m}|^p dp, \quad (31)$$

where

$$V_{t,m} = Y_{t,m} - \Psi_{t,m} \vartheta \quad (32)$$

$$= \begin{bmatrix} e_t \\ e_{t-1} \\ \vdots \\ e_{t-m+1} \end{bmatrix} = \begin{bmatrix} y_t - \mathbf{d}_t^T \mathbf{a} - \mathbf{k}_t^T \mathbf{B} - \mathbf{h}_t^T \mathbf{g} \\ y_{t-1} - \mathbf{d}_{t-1}^T \mathbf{a} - \mathbf{k}_{t-1}^T \mathbf{B} - \mathbf{h}_{t-1}^T \mathbf{g} \\ \vdots \\ y_{t-m+1} - \mathbf{d}_{t-m+1}^T \mathbf{a} - \mathbf{k}_{t-m+1}^T \mathbf{B} - \mathbf{h}_{t-m+1}^T \mathbf{g} \end{bmatrix}, \quad (33)$$

Taking the gradient of $J_2(\vartheta)$ gives

$$\text{grad}[J_2(\vartheta)] = \frac{\partial J_2(\vartheta)}{\partial \vartheta} = \begin{bmatrix} \frac{\partial J_2(\vartheta)}{\partial \mathbf{a}} \\ \frac{\partial J_2(\vartheta)}{\partial \mathbf{B}} \\ \frac{\partial J_2(\vartheta)}{\partial \mathbf{g}} \end{bmatrix}, \quad (34)$$

where

$$\begin{aligned} \frac{\partial J_2(\vartheta)}{\partial \mathbf{a}} &= \int_1^2 p \lambda_t(p) |V_{t,m}|^{p-1} \text{sgn}(V_{t,m}) \frac{\partial(V_{t,m})}{\partial \mathbf{a}} dp \\ &= - \int_1^2 p \lambda_t(p) |V_{t,m}|^{p-1} \text{sgn}(V_{t,m}) dp [\mathbf{d}_t, \mathbf{d}_{t-1}, \dots, \mathbf{d}_{t-m+1}]^T, \\ \frac{\partial J_2(\vartheta)}{\partial \mathbf{B}} &= - \int_1^2 p \lambda_t(p) |V_{t,m}|^{p-1} \text{sgn}(V_{t,m}) dp [\mathbf{k}_t, \mathbf{k}_{t-1}, \dots, \mathbf{k}_{t-m+1}]^T, \\ \frac{\partial J_2(\vartheta)}{\partial \mathbf{g}} &= - \int_1^2 p \lambda_t(p) |V_{t,m}|^{p-1} \text{sgn}(V_{t,m}) dp [\mathbf{h}_t, \mathbf{h}_{t-1}, \dots, \mathbf{h}_{t-m+1}]^T. \end{aligned}$$

Therefore, Eq. (4) can be expressed as

$$\begin{aligned} \text{grad}[J_2(\vartheta)] &= - \int_1^2 p \lambda_t(p) |V_{t,m}|^{p-1} \text{sgn}(V_{t,m}) \Psi_{t,m} dp \\ &= - \Gamma_{t,m} \text{sgn}(V_{t,m}) \Psi_{t,m}, \end{aligned}$$

where

$$\Gamma_{t,m} = \int_1^2 p |V_{t,m}|^{p-1} dp. \quad (35)$$

Let $s = 1, 2, 3, \dots$ be iteration variables. Similar to the derivation of the BSO-CMPN-GI algorithm, using the negative gradient search method and minimizing $J_2(\vartheta)$ get

$$\begin{aligned} \hat{\vartheta}_t^s &= \hat{\vartheta}_t^{s-1} - \mu_t^s \text{grad}[J_2(\hat{\vartheta}_t^{s-1})] \\ &= \hat{\vartheta}_t^{s-1} + \mu_t^s \Gamma_{t,m} \text{sgn}(\mathbf{V}_{t,m}) \Psi_{t,m}, \end{aligned} \quad (36)$$

where

$$0 < \mu_t^s \leq \frac{2}{\lambda_{\max}[\Psi_{t,m}^T \Psi_{t,m}]}. \quad (37)$$

As pointed earlier in Section 3, the unknown states x_{t-j} in the information matrix $\Psi_{t,m}$ are replaced with their estimates. The estimated information matrix $\hat{\Psi}_{t,m}^s$ is redefined as

$$\begin{aligned} \hat{\Psi}_{t,m}^s &:= [(\hat{\psi}_t^s)^T, (\hat{\psi}_{t-1}^s)^T, \dots, (\hat{\psi}_{t-m+1}^s)^T] \in \mathbb{R}^L \times (n^2 + 2n) \\ \hat{\psi}_t^s &= [-\hat{x}_{1,t-1}^{s-1}, \dots, -\hat{x}_{1,t-n}^{s-1}, (\hat{x}_{t-1}^{s-1})^T u_{t-1}, \dots, (\hat{x}_{t-n}^{s-1})^T u_{t-n}, u_{t-1}, \dots, u_{t-n}]^T \in \mathbb{R}^{(n^2+2n)}. \end{aligned}$$

where x_{t-j}^{s-1} is estimated by the state observer:

$$\hat{x}_{t-j}^{s-1} = \hat{\mathbf{A}}_t^{s-1} \hat{x}_{t-j-1}^{s-1} + \hat{\mathbf{B}}_t^{s-1} \hat{x}_{t-j-1}^{s-1} u_{t-j-1} + \hat{\mathbf{g}}_t^{s-1} u_{t-j-1}. \quad (38)$$

Based on the above derivation, a bilinear state observe-based moving window continuous mixed p -norm gradient iterative (BSO-MW-CMPN-GI) algorithm proposed for the bilinear state space system is summarized as follows:

$$\hat{\vartheta}_t^s = \hat{\vartheta}_t^{s-1} + \mu_t^s \Gamma_{t,m}^s \text{sgn}(\mathbf{V}_{t,m}^s) \hat{\Psi}_{t,m}^s, \quad (39)$$

$$0 < \mu_t^s \leq \frac{2}{\lambda_{\max}[(\hat{\Psi}_{t,m}^s)^T \hat{\Psi}_{t,m}^s]}, \quad (40)$$

$$\hat{\Gamma}_{t,m}^s = \frac{(2|\hat{\mathbf{V}}_{t,m}^s| - 1) \ln(|\hat{\mathbf{V}}_{t,m}^s|) - |\hat{\mathbf{V}}_{t,m}^s| + 1}{\ln^2(|\hat{\mathbf{V}}_{t,m}^s|)}, \quad (41)$$

$$\hat{\mathbf{V}}_{t,m}^s = \mathbf{Y}_{t,m} - \hat{\Psi}_{t,m}^s \hat{\vartheta}_t^{s-1}, \quad (42)$$

$$\hat{\Psi}_{t,m}^s := [(\hat{\psi}_t^s)^T, (\hat{\psi}_{t-1}^s)^T, \dots, (\hat{\psi}_{t-m+1}^s)^T], \quad (43)$$

$$\hat{\psi}_t^s = [-\hat{x}_{1,t-1}^{s-1}, \dots, -\hat{x}_{1,t-n}^{s-1}, (\hat{x}_{t-1}^{s-1})^T u_{t-1}, \dots, (\hat{x}_{t-n}^{s-1})^T u_{t-n}, u_{t-1}, \dots, u_{t-n}]^T, \quad (44)$$

$$\hat{x}_{t-j}^s = \hat{\mathbf{A}}_t^s \hat{x}_{t-j-1}^s + \hat{\mathbf{B}}_t^s \hat{x}_{t-j-1}^s u_{t-j-1} + \hat{\mathbf{g}}_t^s u_{t-j-1}. \quad (45)$$

The flowchart of the BSO-MW-CMPN-GI algorithm in Eqs. (39)–(45) is shown in Fig. 2.

Remark 4: The parameter estimate given by the BSO-MW-CMPN-GI algorithm depends only on the iterative counter s , but also time t . As the sampling time t increases, the BSO-MW-CMPN-GI algorithm can utilize a batch of data to calculate the parameter estimate simultaneously.

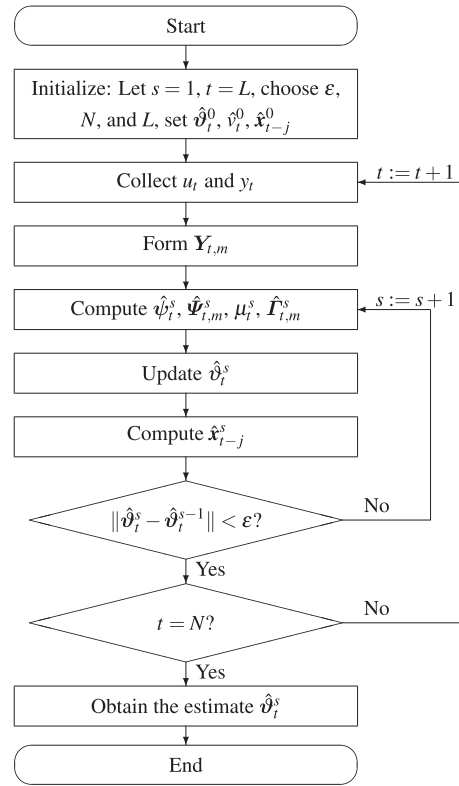


Figure 2: The flowchart of the BSO-MW-CMPN-GI algorithm for computing $\hat{\vartheta}_t^s$

5 Simulation

Case 1: About parameter estimation

Consider a second-order bilinear state space system:

$$\mathbf{x}_{t+1} = \begin{bmatrix} -0.80 & 1 \\ -0.45 & 0 \end{bmatrix} \mathbf{x}_t + \begin{bmatrix} 0.04 & 0.15 \\ 0.25 & 0.18 \end{bmatrix} \mathbf{x}_t u_t + \begin{bmatrix} 1.27 \\ 1.16 \end{bmatrix} u_t + \mathbf{w}_t,$$

$$y_t = [1, 0] \mathbf{x}_t + v_t.$$

The parameter vector to be estimated is

$$\vartheta = [0.80, 0.45, 0.04, 0.15, 0.25, 0.18, 1.27, 1.16]^T$$

In simulation, the input $\{u_t\}$ is taken as an uncorrelated uniform disturbance random signal sequence with zero mean and unit variance, $\{w_t\}$ is taken as an uncorrelated process white noise vector sequence with zero mean and variance $Q = [\sigma_w^2, 0; 0, \sigma_w^2]$, $\sigma_w^2 = 0.02^2$, $\{v_t\}$ is taken as a white noise vector sequence with zero mean and variance $\sigma_v^2 = 1.00^2$ and $\sigma_v^2 = 1.50^2$, respectively. Apply the BSO-CMPN-GI algorithm in Eqs. (24)–(30) and the BSO-MW-CMPN-GI algorithm in Eqs. (39)–(45) with the data length $L = 500$ to estimate the states x_t and parameters ϑ of this bilinear system, respectively. For comparison with the different algorithm, we introduce the bilinear state observe-based gradient iterative (BSO-GI) algorithm. In the BSO-MW-CMPN-GI algorithm, $s = 15$ represents the iteration variable, and $m = 60$ represents the recursive variable. The parameter estimates and errors $\delta := \|\hat{\vartheta}_t - \vartheta\|/\|\vartheta\|$ vs. s of the BSO-GI algorithm, the BSO-CMPN-GI algorithm and

the BSO-MW-CMPN-GI algorithm with different noise variances are shown in Tables 1 and 2. The parameter estimation errors δ vs. s of the BSO-GI algorithm, the BSO-CMPN-GI algorithm and the BSO-MW-CMPN-GI algorithm with different noise variances are shown in Figs. 3–6, respectively. The estimated output (EO) and the actual output (AO) of the four algorithms are shown in Figs. 7 and 8. From Tables 1 and 2 and Figs. 3–8, the following conclusions can be drawn:

- The parameter estimation errors given by the BSO-GI algorithm, the BSO-CMPN-GI algorithm and the BSO-MW-CMPN-GI algorithm become smaller as the iteration s increases. It thus to say the proposed algorithms are effective for bilinear systems.
- The state estimates are close to their true values with t increasing.
- Under the same data length, a lower noise variance leads to higher parameter estimation accuracy by the BSO-GI algorithm, the BSO-CMPN-GI algorithm and the BSO-MW-CMPN-GI algorithm.
- The BSO-CMPN-GI algorithm and the BSO-MW-CMPN-GI algorithm possess higher parameter estimation accuracy at the same noise variance compared with the BSO-GI algorithm.
- When comparing the BSO-CMPN-GI algorithm and the BSO-MW-CMPN-GI algorithm, the parameter estimation errors of the BSO-MW-CMPN-GI algorithm become smaller with m increasing, and approach to zero if m is large enough.

Table 1: The parameter estimates and errors with $\sigma^2 = 1.00^2$

Algorithms	s/m	a_1	a_2	b_{11}	b_{12}	b_{21}	b_{22}	g_1	g_2	δ (%)
BSO-GI	1	0.61876	0.61273	0.69457	0.63482	0.38220	0.36491	0.45135	-0.21511	92.24568
	20	0.70587	0.40500	0.10239	0.07684	0.28265	0.07919	1.29140	1.00820	11.82990
	50	0.81087	0.41739	0.06110	0.11101	0.25343	0.11478	1.28666	1.22632	5.55331
	300	0.81427	0.41734	0.05518	0.11715	0.24064	0.12820	1.28683	1.23679	5.42356
	500	0.81427	0.41734	0.05518	0.11715	0.24064	0.12820	1.28683	1.23679	5.42356
BSO-CMPN-GI	1	0.66102	0.66319	0.73450	0.67753	0.43753	0.40706	0.52491	-0.13652	89.61731
	20	0.74316	0.36344	0.08656	0.06822	0.27021	0.10784	1.29569	1.04323	10.03720
	50	0.79850	0.42837	0.05658	0.11532	0.21864	0.15164	1.28695	1.22166	4.46772
	300	0.79902	0.42732	0.05460	0.11635	0.21291	0.15495	1.28663	1.22366	4.57158
	500	0.79813	0.42874	0.05359	0.11596	0.21357	0.15497	1.28626	1.22368	4.53760
BSO-MW-CMPN-GI	1	0.48656	0.46967	0.10661	0.08819	0.26243	0.09292	1.31132	0.84548	23.43355
	5	0.79794	0.41400	0.04145	0.11807	0.20383	0.16096	1.31538	1.24927	6.15390
	20	0.79722	0.42250	0.04978	0.10488	0.21888	0.14161	1.30957	1.21724	5.09596
	50	0.80133	0.43472	0.03665	0.12151	0.21644	0.13795	1.29251	1.19960	3.92121
	60	0.80246	0.43821	0.03819	0.12520	0.21148	0.15026	1.28898	1.19617	3.49939
True values		0.80000	0.45000	0.04000	0.15000	0.25000	0.18000	1.27000	1.16000	

Table 2: The parameter estimates and errors with $\sigma^2 = 1.50^2$

Algorithms	s/k	a_1	a_2	b_{11}	b_{12}	b_{21}	b_{22}	g_1	g_2	δ (%)
BSO-GI	1	0.61450	0.61472	0.69926	0.63402	0.38763	0.37738	0.45898	-0.20242	91.78360
	20	0.71081	0.38354	0.11737	0.04826	0.28331	0.04731	1.29900	1.03783	12.69672
	50	0.81667	0.39619	0.06896	0.09614	0.24797	0.09247	1.29665	1.26099	8.06372
	300	0.82052	0.39570	0.05825	0.10732	0.22565	0.11513	1.29725	1.27380	7.82752
	500	0.82052	0.39570	0.05825	0.10732	0.22565	0.11513	1.29725	1.27380	7.82752

(Continued)

Table 2 (continued)

Algorithms	s/k	a_1	a_2	b_{11}	b_{12}	b_{21}	b_{22}	g_1	g_2	δ (%)
BSO-CMPN-GI	1	0.66058	0.67114	0.74431	0.68430	0.45736	0.43678	0.54482	-0.08676	88.05004
	20	0.70979	0.38841	0.10781	0.05499	0.24957	0.08174	1.29908	1.04184	11.29602
	50	0.79822	0.40981	0.06272	0.09945	0.19873	0.13965	1.29824	1.25268	6.83866
	300	0.79968	0.41030	0.05632	0.10549	0.18500	0.15328	1.29669	1.25761	6.96239
	500	0.79968	0.41030	0.05632	0.10549	0.18500	0.15328	1.29669	1.25761	6.96239
BSO-MW-CMPN-GI	1	0.60409	0.32275	0.09324	0.05963	0.24551	0.09576	1.33478	0.82048	22.14057
	5	0.79684	0.38475	0.03655	0.11005	0.17120	0.15953	1.33667	1.29024	9.30110
	20	0.79656	0.40420	0.04760	0.08972	0.19317	0.13151	1.32589	1.24236	7.36956
	50	0.80177	0.42162	0.02914	0.11338	0.18645	0.12773	1.30345	1.21922	5.90409
	60	0.80523	0.42796	0.03189	0.11936	0.17734	0.14765	1.29610	1.21551	5.44134
True values		0.80000	0.45000	0.04000	0.15000	0.25000	0.18000	1.27000	1.16000	

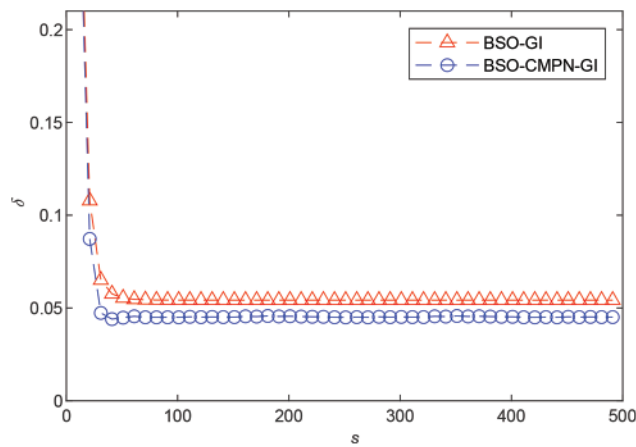


Figure 3: The estimation errors δ vs. t with $\sigma^2 = 1.00^2$

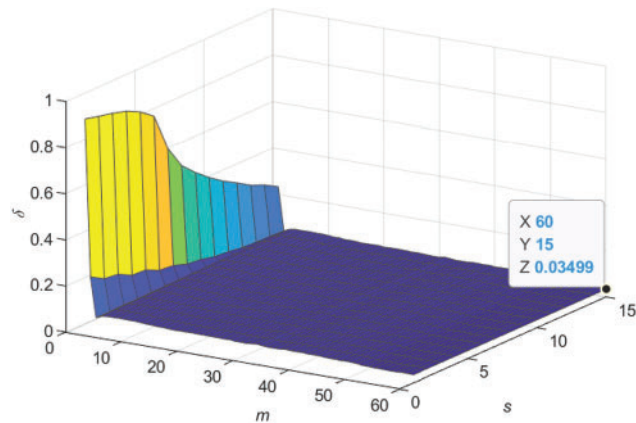


Figure 4: The BSO-MW-CMPN-GI estimation errors δ vs. m and s ($m = 60$ and $s = 15$) with $\sigma^2 = 1.00^2$

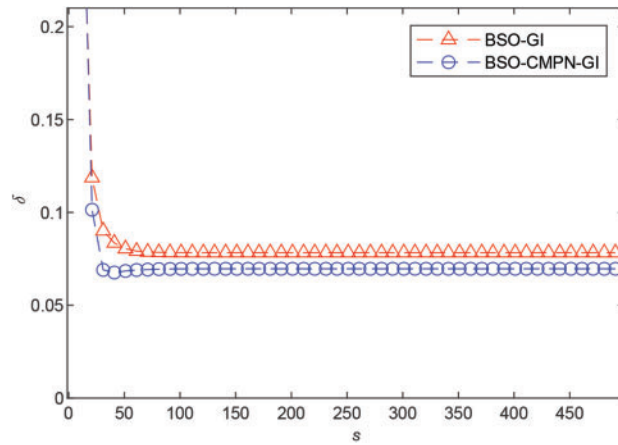


Figure 5: The estimation errors δ vs. t with $\sigma^2 = 1.50^2$

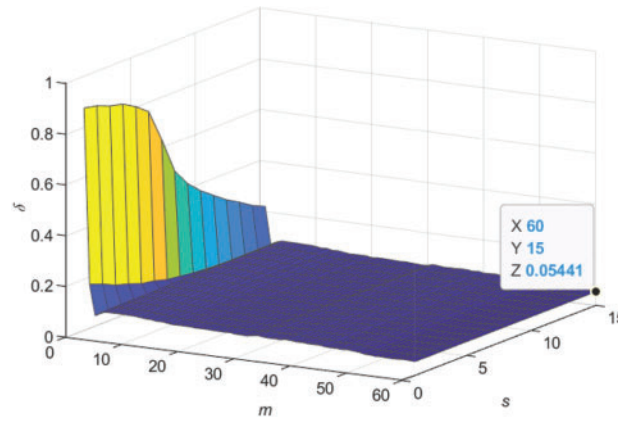


Figure 6: The BSO-MW-CMPN-GI estimation errors δ vs. m and s ($m = 60$ and $s = 15$) with $\sigma^2 = 1.50^2$

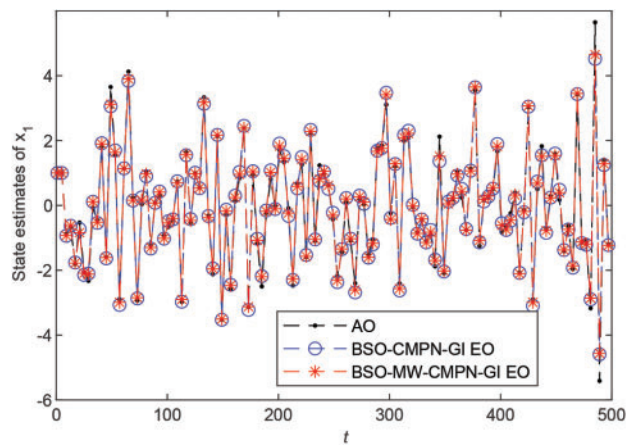


Figure 7: State $x_{1,t}$ and the estimated state $\hat{x}_{1,t}$ vs. t

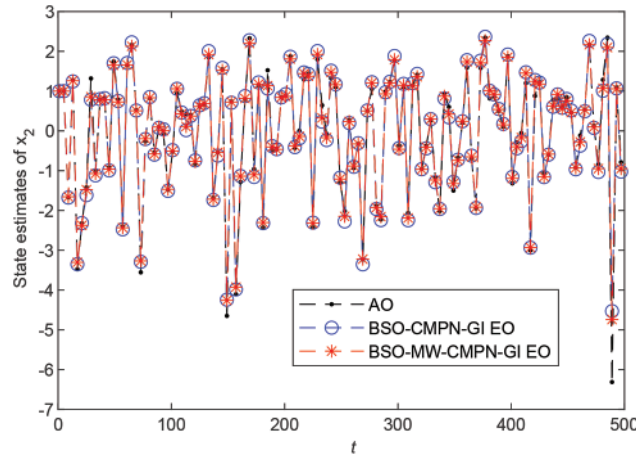


Figure 8: State $x_{2,t}$ and the estimated state $\hat{x}_{2,t}$ vs. t

Case 2: About model validation

Applying the BSO-CMPN-GI algorithm and the BSO-MW-CMPN-GI algorithm to construct the estimated model for the model validation, respectively. Take the data from $t = 1001$ to $t = 2000$ to calculate the root mean square errors (RMSEs) of the predicted output \hat{y}_t . Using the BSO-CMPN-GI and BSO-MW-CMPN-GI estimates in Table 1 with the noise variance $\sigma_v^2 = 1.00^2$. The actual output (AO) y_t and the predicted output (PO) \hat{y}_t are plotted in Fig. 9. The RMSEs of the BSO-GI, BSO-CMPN-GI and BSO-MW-CMPN-GI are

$$\delta_1 = \sqrt{\frac{1}{L_t} \sum_{j=1001}^{2000} [y_j - \hat{y}_{\text{BSO-GI}_j}]^2} = 1.0223,$$

$$\delta_2 = \sqrt{\frac{1}{L_t} \sum_{j=1001}^{2000} [y_j - \hat{y}_{\text{BSO-CMPN-GI}_j}]^2} = 1.0187,$$

$$\delta_3 = \sqrt{\frac{1}{L_t} \sum_{j=1001}^{2000} [y_j - \hat{y}_{\text{BSO-MW-CMPN-GI}_j}]^2} = 1.0127,$$

where $\hat{y}_{\text{BSO-GI}_j}$, $\hat{y}_{\text{BSO-CMPN-GI}_j}$ and $\hat{y}_{\text{BSO-MW-CMPN-GI}_j}$ represent the predicted output of the BSO-GI algorithm, the BSO-CMPN-GI algorithm and the BSO-MW-CMPN-GI algorithm, respectively.

From Fig. 9, we can see that the predicted outputs of the BSO-CMPN-GI and the BSO-MW-CMPN-GI are very close to the true outputs, and the RMSEs of the two algorithms are very close to the noise standard deviation $\sigma_v^2 = 1.00^2$. In other words, the estimated model can capture the dynamics of the system.

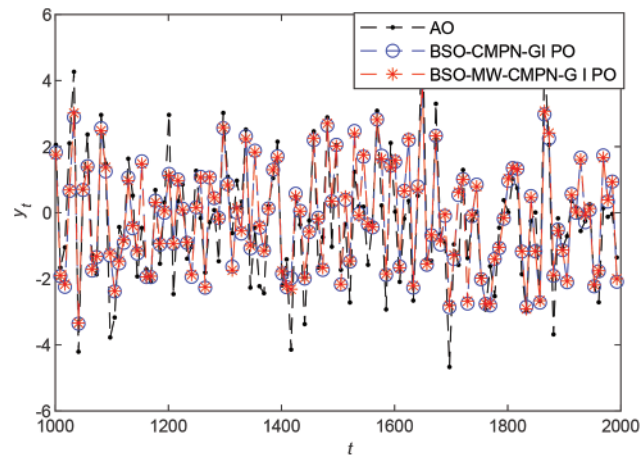


Figure 9: The system actual outputs and the predicted outputs vs. t with $\sigma_v^2 = 1.00^2$

6 Conclusions

This paper studies the parameter identification problems of the nonlinear systems described by the bilinear state space models with the noise disturbances. A bilinear state observe-based continuous mixed p -norm gradient iterative algorithm and a bilinear state observe-based moving window continuous mixed p -norm gradient iterative algorithm are proposed to estimate the parameters of the bilinear system. The proposed optimal algorithms are robustness for stochastic white noise by means of the continuously mixed p -norms theory. In order to improve the performance of the proposed algorithms, the moving window identification theory is introduced into the proposed algorithm. Although the proposed algorithms are effective for identifying the bilinear system, it also has some limitations. For example, the disturbance noise is confined to the white noise. The proposed model parameter estimation methods in the paper can combine some mathematical strategies [58–60] and other estimation algorithms [61–65] to study the parameter identification problems of linear and nonlinear systems with different disturbances [66–68] and can be applied to other fields [69–74] such as engineering application systems.

Funding Statement: This research was funded by the National Natural Science Foundation of China (No. 61773182) and the 111 Project (B12018).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Xu, L. (2022). Separable multi-innovation newton iterative modeling algorithm for multi-frequency signals based on the sliding measurement window. *Circuits Systems and Signal Processing*, 41(2), 805–830. DOI 10.1007/s00034-021-01801-x.
2. Xu, L. (2022). Separable newton recursive estimation method through system responses based on dynamically discrete measurements with increasing data length. *International Journal of Control Automation and Systems*, 20(2), 432–443. DOI 10.1007/s12555-020-0619-y.

3. Zhang, X., Liu, Q., Hayat, T. (2020). Recursive identification of bilinear time-delay systems through the redundant rule. *Journal of the Franklin Institute*, 357(1), 726–747. DOI 10.1016/j.jfranklin.2019.11.003.
4. Gan, M., Chen, G., Chen, L., Chen, C. (2020). Term selection for a class of separable nonlinear models. *IEEE Transactions on Neural Networks and Learning Systems*, 31(2), 445–451. DOI 10.1109/TNNLS.5962385.
5. Ding, F. (2013). Coupled-least-squares identification for multivariable systems. *IET Control Theory and Applications*, 7(1), 68–79. DOI 10.1049/iet-cta.2012.0171.
6. Ding, F., Liu, X. G., Chu, J. (2013). Gradient-based and least-squares-based iterative algorithms for hammerstein systems using the hierarchical identification principle. *IET Control Theory and Applications*, 7(2), 176–184. DOI 10.1049/iet-cta.2012.0313.
7. Zhao, Y. B., Liu, G. P., Rees, D. (2008). A predictive control-based approach to networked hammerstein systems: Design and stability analysis. *IEEE Transactions on Cybernetics*, 38(3), 700–708. DOI 10.1109/TSMCB.2008.918572.
8. Mahata, K., Schoukens, J., de Cock, A. (2016). Information matrix and D-optimal design with Gaussian inputs for wiener model identification. *Automatica*, 69, 65–77. DOI 10.1016/j.automatica.2016.02.026.
9. Wang, Y. J., Tang, S. H., Gu, X. B. (2022). Parameter estimation for nonlinear volterra systems by using the multi-innovation identification theory and tensor decomposition. *Journal of the Franklin Institute*, 359(2), 1782–1802. DOI 10.1016/j.jfranklin.2021.11.015.
10. Xu, L., Song, G. L. (2020). A recursive parameter estimation algorithm for modeling signals with multi-frequencies. *Circuits Systems and Signal Processing*, 39(8), 4198–4224. DOI 10.1007/s00034-020-01356-3.
11. Wang, Y. J., Yang, L. (2021). An efficient recursive identification algorithm for multilinear systems based on tensor decomposition. *International Journal of Robust and Nonlinear Control*, 31(11), 7920–7936. DOI 10.1002/rnc.5718.
12. Ding, F., Zhang, X., Xu, L. (2019). The innovation algorithms for multivariable state-space models. *International Journal of Adaptive Control and Signal Processing*, 33(11), 1601–1608. DOI 10.1002/acs.3053.
13. Ding, F., Liu, G., Liu, X. P. (2011). Parameter estimation with scarce measurements. *Automatica*, 47(11), 1646–1655. DOI 10.1016/j.automatica.2011.05.007.
14. Liu, Y. J., Shi, Y. (2014). An efficient hierarchical identification method for general dual-rate sampled-data systems. *Automatica*, 50(3), 962–970. DOI 10.1016/j.automatica.2013.12.025.
15. Zhang, X. (2022). Optimal adaptive filtering algorithm by using the fractional-order derivative. *IEEE Signal Processing Letters*, 29, 399–403. DOI 10.1109/LSP.2021.3136504.
16. Ding, J., Liu, X. P., Liu, G. (2011). Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data. *IEEE Transactions on Automatic Control*, 56(11), 2677–2683. DOI 10.1109/TAC.2011.2158137.
17. Ding, F., Liu, Y. J., Bao, B. (2012). Gradient based and least squares based iterative estimation algorithms for multi-input multi-output systems. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 226(1), 43–55. DOI 10.1177/0959651811409491.
18. Xu, L., Chen, F. Y., Hayat, T. (2021). Hierarchical recursive signal modeling for multi-frequency signals based on discrete measured data. *International Journal of Adaptive Control and Signal Processing*, 35(5), 676–693. DOI 10.1002/acs.3221.
19. Wang, Y. J. (2016). Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model. *Automatica*, 71, 308–313. DOI 10.1016/j.automatica.2016.05.024.
20. Mohler, R. R., Kolodziej, W. J. (2007). An overview of bilinear system theory and applications. *IEEE Transactions on Systems Man and Cybernetics*, 10(10), 683–688. DOI 10.1109/TSMC.1980.4308378.
21. Bruni, C., Dipillo, G., Koch, G. (2003). Bilinear systems: An appealing class of “nearly linear” systems in theory and applications. *IEEE Transactions on Automatic Control*, 19(4), 334–348. DOI 10.1109/TAC.1974.1100617.

22. Yeo, Y. K., Choo, Y. U. (2006). Bilinear model predictive control of grade change operations in paper production plants. *Korean Journal of Chemical Engineering*, 23(2), 167–170. DOI 10.1007/BF02705710.
23. Wang, Z. C., Zha, G. P., Ren, W. X. (2018). Nonlinear boundary parameter identification of bridges based on temperature-induced strains. *Structural Engineering and Mechanics*, 68(5), 563–573. DOI 10.12989/sem.2018.68.5.563.
24. Li, M., Liu, X. (2020). Maximum likelihood least squares based iterative estimation for a class of bilinear systems using the data filtering technique. *International Journal of Control Automation and Systems*, 18(6), 1581–1592. DOI 10.1007/s12555-019-0191-5.
25. Tsai, S. H. (2012). A global exponential fuzzy observer design for time-delay takagi-sugeno uncertain discrete fuzzy bilinear systems with disturbance. *IEEE Transactions on Fuzzy Systems*, 20(6), 1063–1075. DOI 10.1109/TFUZZ.2012.2192936.
26. Phan, M. Q., Vicario, F., Longman, R. W., Betti, R. (2015). Optimal bilinear observers for bilinear state-space models by interaction matrices. *International Journal of Control*, 88(8), 1504–1522. DOI 10.1080/00207179.2015.1007530.
27. Zhang, X., Yang, E. (2019). State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors. *International Journal of Adaptive Control and Signal Processing*, 33(7), 1157–1173. DOI 10.1002/acs.3027.
28. Zhang, X. (2020). Adaptive parameter estimation for a general dynamical system with unknown states. *International Journal of Robust and Nonlinear Control*, 30(4), 1351–1372. DOI 10.1002/rnc.4819.
29. Zhang, X. (2020). Recursive parameter estimation methods and convergence analysis for a special class of nonlinear systems. *International Journal of Robust and Nonlinear Control*, 30(4), 1373–1393. DOI 10.1002/rnc.4824.
30. Zhang, X. (2020). Recursive parameter estimation and its convergence for bilinear systems. *IET Control Theory and Applications*, 14(5), 677–688. DOI 10.1049/iet-cta.2019.0413.
31. Zayyani, H. (2014). Continuous mixed p -norm adaptive algorithm for system identification. *IEEE Signal Process Letters*, 21(9), 1108–1110. DOI 10.1109/LSP.2014.2325495.
32. Shi, L., Zhao, H., Zakharov, Y. (2019). Generalized variable step size continuous mixed p -norm adaptive filtering algorithm. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 66(6), 1078–1082. DOI 10.1109/TCSI-I.2018.2873254.
33. Shi, L., Zhao, H. (2021). Generalized variable step-size diffusion continuous mixed p -norm algorithm. *Circuits Systems and Signal Processing*, 40(7), 3609–3620. DOI 10.1007/s00034-020-01640-2.
34. Ding, F., Chen, T. (2004). Combined parameter and output estimation of dual-rate systems using an auxiliary model. *Automatica*, 40(9), 1739–1748. DOI 10.1016/j.automatica.2004.05.001.
35. Ding, F., Chen, T. (2005). Parameter estimation of dual-rate stochastic systems by using an output error method. *IEEE Transactions on Automatic Control*, 50(9), 1436–1441. DOI 10.1109/TAC.2005.854654.
36. Ding, F., Shi, Y., Chen, T. (2017). Auxiliary model-based least-squares identification methods for hammerstein output-error systems. *Systems & Control Letters*, 56(7), 373–380. DOI 10.1016/j.sysconle.2006.10.026.
37. Ji, Y., Kang, Z., Zhang, C. (2021). Two-stage gradient-based recursive estimation for nonlinear models by using the data filtering. *International Journal of Control Automation and Systems*, 19(8), 2706–2715. DOI 10.1007/s12555-019-1060-y.
38. Wang, J. W., Ji, Y., Zhang, C. (2021). Iterative parameter and order identification for fractional-order nonlinear finite impulse response systems using the key term separation. *International Journal of Adaptive Control and Signal Processing*, 35(8), 1562–1577. DOI 10.1002/acs.3257.
39. Li, M. H., Liu, X. M. (2021). Iterative identification methods for a class of bilinear systems by using the particle filtering technique. *International Journal of Adaptive Control and Signal Processing*, 35(10), 2056–2074. DOI 10.1002/acs.3308.

40. Zhou, Y. H. (2020). Modeling nonlinear processes using the radial basis function-based state-dependent autoregressive models. *IEEE Signal Processing Letters*, 27, 1600–1604. DOI 10.1109/LSP.2020.3021925.
41. Zhou, Y. H., Zhang, X. (2022). Partially-coupled nonlinear parameter optimization algorithm for a class of multivariate hybrid models. *Applied Mathematics and Computation*, 414, 126663. DOI 10.1016/j.amc.2021.126663.
42. Zhou, Y. H., Zhang, X. (2021). Hierarchical estimation approach for RBF-AR models with regression weights based on the increasing data length. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 68(12), 3597–3601. DOI 10.1109/TCSII.2021.3076112.
43. Li, M. H., Liu, X. M. (2021). Maximum likelihood hierarchical least squares-based iterative identification for dual-rate stochastic systems. *International Journal of Adaptive Control and Signal Processing*, 35(2), 240–261. DOI 10.1002/acs.3203.
44. Ding, F., Shi, Y., Chen, T. (2006). Performance analysis of estimation algorithms of non-stationary ARMA processes. *IEEE Transactions on Signal Processing*, 54(3), 1041–1053. DOI 10.1109/TSP.2005.862845.
45. Wan, L. J. (2019). Decomposition-and gradient-based iterative identification algorithms for multivariable systems using the multi-innovation theory. *Circuits Systems and Signal Processing*, 38(7), 2971–2991. DOI 10.1007/s00034-018-1014-2.
46. Wang, X. H. (2022). Modified particle filtering-based robust estimation for a networked control system corrupted by impulsive noise. *International Journal of Robust and Nonlinear Control*, 32(2), 830–850. DOI 10.1002/rnc.5850.
47. Hou, J., Chen, F. W., Li, P. H., Zhu, Z. Q. (2021). Gray-box parsimonious subspace identification of hammerstein-type systems. *IEEE Transactions on Industrial Electronics*, 68(10), 9941–9951. DOI 10.1109/TIE.2020.3026286.
48. Ding, F., Liu, G., Liu, X. P. (2011). Partially coupled stochastic gradient identification methods for non-uniformly sampled systems. *IEEE Transactions on Automatic Control*, 55(8), 1976–1981. DOI 10.1109/TAC.2010.2050713.
49. Zhao, Z., Zhou, Y., Wang, X., Wang, Z., Bai, Y. (2022). Water quality evolution mechanism modeling and health risk assessment based on stochastic hybrid dynamic systems. *Expert Systems with Applications*, 193, 116404. DOI 10.1016/j.eswa.2021.116404.
50. Zhao, Z., Zhou, Y., Wang, X., Wang, Z., Bai, Y. (2022). Microbiological predictive modeling and risk analysis based on the one-step kinetic integrated wiener process. *Innovative Food Science & Emerging Technologies*, 75, 102912. DOI 10.1016/j.ifset.2021.102912.
51. Zhang, X. (2020). Hierarchical parameter and state estimation for bilinear systems. *International Journal of Systems Science*, 51(2), 275–290. DOI 10.1080/00207721.2019.1704093.
52. Pan, J., Jiang, X., Ding, W. (2017). A filtering based multi-innovation extended stochastic gradient algorithm for multivariable control systems. *International Journal of Control Automation and Systems*, 15(3), 1189–1197. DOI 10.1007/s12555-016-0081-z.
53. Ding, F., Lv, L., Pan, J., Wan, X., Jin, X. B. (2020). Two-stage gradient-based iterative estimation methods for controlled autoregressive systems using the measurement data. *International Journal of Control Automation and Systems*, 18(4), 886–896. DOI 10.1007/s12555-019-0140-3.
54. Ding, F., Wang, F. F., Wu, M. H. (2017). Decomposition based least squares iterative identification algorithm for multivariate pseudo-linear ARMA systems using the data filtering. *Journal of the Franklin Institute*, 354(3), 1321–1339. DOI 10.1016/j.jfranklin.2016.11.030.
55. Pan, J., Ma, H., Liu, Q. Y. (2020). Recursive coupled projection algorithms for multivariable output-error-like systems with coloured noises. *IET Signal Processing*, 14(7), 455–466. DOI 10.1049/iet-spr.2019.0481.
56. Xu, L., Zhu, Q. M. (2021). Decomposition strategy-based hierarchical least mean square algorithm for control systems from the impulse responses. *International Journal of Systems Science*, 52(9), 1806–1821. DOI 10.1080/00207721.2020.1871107.

57. Zhang, X., Xu, L., Hayat, T. (2018). Combined state and parameter estimation for a bilinear state space system with moving average noise. *Journal of the Franklin Institute*, 355(6), 3079–3103. DOI 10.1016/j.jfranklin.2018.01.011.
58. Li, X. Y., Wu, B. Y. (2022). A kernel regression approach for identification of first order differential equations based on functional data. *Applied Mathematics Letters*, 127, 107832. DOI 10.1016/j.aml.2021.107832.
59. Geng, F. Z., Wu, X. Y. (2021). Reproducing kernel functions based univariate spline interpolation. *Applied Mathematics Letters*, 122, 107525. DOI 10.1016/j.aml.2021.107525.
60. Li, X. Y., Wu, B. Y. (2021). Superconvergent kernel functions approaches for the second kind fredholm integral equations. *Applied Numerical Mathematics*, 167, 202–210. DOI 10.1016/j.apnum.2021.05.004.
61. Pan, J., Li, W., Zhang, H. P. (2018). Control algorithms of magnetic suspension systems based on the improved double exponential reaching law of sliding mode control. *International Journal of Control Automation and Systems*, 16(6), 2878–2887. DOI 10.1007/s12555-017-0616-y.
62. Ma, H., Pan, J., Ding, W. (2019). Partially-coupled least squares based iterative parameter estimation for multi-variable output-error-like autoregressive moving average systems. *IET Control Theory and Applications*, 13(8), 3040–3051. DOI 10.1049/iet-cta.2019.0112.
63. Ding, F., Liu, X. P., Yang, H. Z. (2008). Parameter identification and intersample output estimation for dual-rate systems. *IEEE Transactions on Systems, Man, and Cybernetics–Part A: Systems and Humans*, 38(4), 966–975. DOI 10.1109/TSMCA.2008.923030.
64. Ding, F., Liu, X. P., Liu, G. (2010). Multiinnovation least squares identification for linear and pseudo-linear regression models. *IEEE Transactions on Systems, Man, and Cybernetics–Part B: Cybernetics*, 40(3), 767–778. DOI 10.1109/TSMCB.2009.2028871.
65. Xu, L., Sheng, J. (2020). Separable multi-innovation stochastic gradient estimation algorithm for the nonlinear dynamic responses of systems. *International Journal of Adaptive Control and Signal Processing*, 34(7), 937–954. DOI 10.1002/acs.3113.
66. Wang, Y. J., Wu, M. H. (2018). Recursive parameter estimation algorithm for multivariate output-error systems. *Journal of the Franklin Institute*, 355(12), 5163–5181. DOI 10.1016/j.jfranklin.2018.04.013.
67. Xu, L., Yang, E. F. (2021). Auxiliary model multiinnovation stochastic gradient parameter estimation methods for nonlinear sandwich systems. *International Journal of Robust and Nonlinear Control*, 31(1), 148–165. DOI 10.1002/rnc.5266.
68. Liu, S. Y., Hayat, T. (2019). Hierarchical principle-based iterative parameter estimation algorithm for dual-frequency signals. *Circuits Systems and Signal Processing*, 38(7), 3251–3268. DOI 10.1007/s00034-018-1015-1.
69. Zhao, G., Cao, T., Wang, Y., Zhou, H., Zhang, C. (2021). Optimal sizing of isolated microgrid containing photovoltaic/photothermal/wind/diesel/battery. *International Journal of Photoenergy*, 2021, 5566597. DOI 10.1155/2021/5566597.
70. Wang, X., Zhao, M., Zhou, Y., Wan, Z., Xu, W. (2021). Design and analysis for multi-disc coreless axial-flux permanent-magnet synchronous machine. *IEEE Transactions on Applied Superconductivity*, 31(8), 1–4. DOI 10.1109/TASC.2021.3091078.
71. Wang, X., Wan, Z., Tang, L., Xu, W., Zhao, M. (2021). Electromagnetic performance analysis of an axial flux hybrid excitation motor for HEV drives. *IEEE Transactions on Applied Superconductivity*, 31(8), 1–5. DOI 10.1109/TASC.2021.3101785.
72. Li, M., Xu, G., Lai, Q., Chen, J. (2022). A chaotic strategy-based quadratic opposition-based learning adaptive variable-speed whale optimization algorithm. *Mathematics and Computers in Simulation*, 193, 71–99. DOI 10.1016/j.matcom.2021.10.003.

73. Shu, J., He, J. C., Li, L. (2022). MSIS: Multispectral instance segmentation method for power equipment. *Computational Intelligence and Neuroscience*, 2022. DOI 10.1155/2022/2864717.
74. Peng, H., He, W., Zhang, Y., Li, X., Ding, Y. (2022). Covert non-orthogonal multiple access communication assisted by multi-antenna jamming author links open overlay. *Physical Communication*, 2022. DOI 10.1016/j.phycom.2022.101598.