Received: 12 November 2018 | Accepted: 13 May 2019
DOI: 10.1002/acs.3027

RESEARCH ARTICLE

WILEY

State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors

Xiao Zhang[1] | Feng Ding[1,2,3] | Erfu Yang[4]

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

[2]College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao, China

[3]School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan, China

[4]Space Mechatronic Systems Technology Laboratory, University of Strathclyde, Scotland, UK

Correspondence
Feng Ding, Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China.
Email: fding@jiangnan.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 61873111; 111 Project, Grant/Award Number: B12018; Postgraduate Research and Practice Innovation Program of Jiangsu Province, Grant/Award Number: KYCX18_1854

Summary
This paper considers the state estimation problem of bilinear systems in the presence of disturbances. The standard Kalman filter is recognized as the best state estimator for linear systems, but it is not applicable for bilinear systems. It is well known that the extended Kalman filter (EKF) is proposed based on the Taylor expansion to linearize the nonlinear model. In this paper, we show that the EKF method is not suitable for bilinear systems because the linearization method for bilinear systems cannot describe the behavior of the considered system. Therefore, this paper proposes a state filtering method for the single-input–single-output bilinear systems by minimizing the covariance matrix of the state estimation errors. Moreover, the state estimation algorithm is extended to multiple-input–multiple-output bilinear systems. The performance analysis indicates that the state estimates can track the true states. Finally, the numerical examples illustrate the specific performance of the proposed method.

KEYWORDS
bilinear state estimator, Kalman filter, signal processing, state estimation

1 | INTRODUCTION

For decades, there are many activities in utilizing linear models,[1-3] bilinear models, and nonlinear models[4,5] and many identification methods have been developed. Specifically, much attractive attention has been paid to bilinear systems as they are simple nonlinear systems and represent the intermediary structure between linear models and nonlinear models.[6] Some methods based on the approximation idea have been used in parameter estimation for bilinear systems.[7] Dai and Sinha utilized the block functions for parameter estimation of the bilinear system.[8] Hizir et al identified the bilinear systems through equivalent linear models.[9] Nowadays, the Carleman linearization is an approach to reach the approximation, and the bilinear model is proven to be an effective approximator for some nonlinear systems, which can solve the nonlinear system state filtering problems in signal processing and control.[10]

Parameter estimation methods and state filtering can be applied to many areas,[11,12] such as information fusion and fault diagnosis,[13,14] system modelling,[15,16] and signal processing.[17-19] The Kalman filter (KF) is considered as one of the most common state filtering methods for the linear state-space systems with Gaussian noise since the 1960s.[20] However, the Kalman filtering method cannot be applied to bilinear systems. In the literature, the particle filters can be used for state estimation of nonlinear systems with non-Gaussian noise. For nonlinear systems with Gaussian noise, the typical filters such as the extended KF (EKF), the unscented KF, and the Guass-Hermite quadrature filter can be applied to state filtering. Moreover, Zhao et al proposed a Kalman-like optimal unbiased finite impulse response filter.[21,22] Johnston and Krishnamurthy presented an iterative algorithm for the state estimation of bilinear systems based on the expectation maximum.[23] Germani et al proposed a decomposition-based robust filtering algorithm to estimate the states of the time-varying bilinear systems with unknown inputs.[24] Kulikov and Kulikova presented a continuous-discrete unscented state filter for nonlinear stochastic models in radar tracking.[25]

Regardless of the filtering technique for state estimation of bilinear systems, the state observer is vital in the field of control. Hara and Furuta considered a minimal order state estimator for bilinear systems, whose estimation error was independent of inputs.[26] Tsai developed a linear matrix inequality approach to design a robust H-infinity fuzzy observer for a class of time-delay Takagi-Sugeno uncertain discrete bilinear systems.[27] Gomez-Exposito et al presented a three-stage state estimation method for the energy management system based on the explicit nonlinear transformation and the linear weighted least squares solution.[28] Phan et al formulated a full-order bilinear state observer and optimized the observer gain by interaction matrices.[29]

Since the KF is the optimal linear filter for state estimation, it is not applicable for nonlinear state estimation. In this paper, a bilinear state estimator is formulated to solve the state estimation problem for the single-input–single-output bilinear system on the basis of the extremum principle. Then, the proposed state estimation method is extended to obtain the unknown states of the multiple-input–multiple-output bilinear system. The basic idea is to minimize the covariance matrix of the state estimation errors and to obtain the optimal gain vector based on the delta operator. The main contributions of this paper are listed as follows:

- present a bilinear state estimator for a bilinear state-space system disturbed by the process noise and the measurement noise;
- apply the delta operator to minimize the state estimation error covariance matrix and compute the optimal gain vector;
- analyze the convergence of the proposed algorithm and demonstrate the performance of the state filter through a numerical example.

The outline of this paper is as follows. Section 2 describes the state filtering problem of bilinear systems. According to the extremum principle, the derivation of the bilinear state estimator for the single-input–single-output bilinear system is presented in Section 3. Then, the extension of the state filtering method for the multiple-input–multiple-output bilinear system is proposed to estimate the unknown states in Section 4. In Section 5, the performance of the proposed algorithms is shown based on an illustrative example. Finally, some conclusions are given in Section 6.

2 | PROBLEM STATEMENT

In this paper, the following symbols are exploited in such a way that the expression "$A =: X$" or "$X := A$" stands for "A is defined as X", the superscript T denotes the matrix/vector transpose, the symbol $\boldsymbol{I}$ ($\boldsymbol{I}_n$) stands for an identity matrix of appropriate sizes ($n \times n$), and $\mathbf{1}_n$ marks an n-dimensional column vector whose elements are all unity.

Consider a bilinear state-space model[30-32]

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{x}_k u_k + \boldsymbol{f} u_k + \boldsymbol{w}_k, \tag{1}$$

$$y_k = \boldsymbol{c}\boldsymbol{x}_k + d u_k + v_k, \tag{2}$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ is the system state vector, $u_k \in \mathbb{R}$ is the system input, $y_k \in \mathbb{R}$ is the system output, $v_k \in \mathbb{R}$ is an uncorrelated random noise with zero mean, $\boldsymbol{w}_k \in \mathbb{R}^n$ is an uncorrelated process noise vector with zero mean, and $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times n}$, $\boldsymbol{f} \in \mathbb{R}^n$, $\boldsymbol{c} \in \mathbb{R}^{1\times n}$, and $d \in \mathbb{R}$ are the known parameters of the system.

Assume that $\boldsymbol{w}_k$ and v_k are uncorrelated and satisfy

$(A1)$ $\mathrm{E}[\boldsymbol{w}_k] = \boldsymbol{0}, \mathrm{E}[v_k] = 0, \mathrm{E}[\boldsymbol{w}_k v_i] = \boldsymbol{0},$

(*A*2) $\mathrm{E}[\boldsymbol{w}_k\boldsymbol{w}_s^{\mathrm{T}}]=\boldsymbol{0}, \quad \mathrm{E}[v_kv_s]=0, \quad k\neq s,$
(*A*3) $\mathrm{E}[\boldsymbol{w}_k\boldsymbol{w}_k^{\mathrm{T}}]=\boldsymbol{R}_w\in\mathbb{R}^{n\times n}, \quad \mathrm{E}[v_k^2]=R_v\in\mathbb{R}.$

Remark 1. It is well known that the standard KF is applied for linear system state filtering, which is not applicable for nonlinear systems. In order to overcome this difficulty, one uses the EKF to linearize a nonlinear system function by computing the partial derivative of the state and measurement equations to obtain the estimated states. Since the bilinear systems are a special class of nonlinear systems, it is curious whether we can utilize the linearization method to solve the state estimation problem for bilinear systems. The answer is no. The details are as follows.

The bilinear state-space model in (1)-(2) can be expressed as the following nonlinear model:

$$\boldsymbol{x}_{k+1}=\boldsymbol{g}(\boldsymbol{x}_k,u_k)+\boldsymbol{w}_k, \tag{3}$$

$$y_k=\boldsymbol{c}\boldsymbol{x}_k+du_k+v_k, \tag{4}$$

$$\boldsymbol{g}(\boldsymbol{x}_k,u_k)=\boldsymbol{A}\boldsymbol{x}_k+\boldsymbol{B}\boldsymbol{x}_ku_k+\boldsymbol{f}u_k. \tag{5}$$

Then we use the linearization method for (3) and get a linear approximate model

$$\begin{aligned}\boldsymbol{x}_{k+1}&=\frac{\partial\boldsymbol{g}(\boldsymbol{0},0)}{\partial\boldsymbol{x}_k}\boldsymbol{x}_k+\frac{\partial\boldsymbol{g}(\boldsymbol{0},0)}{\partial u_k}u_k+\boldsymbol{w}_k\\&=\boldsymbol{A}\boldsymbol{x}_k+\boldsymbol{f}u_k+\boldsymbol{w}_k.\end{aligned} \tag{6}$$

Remark 2. This linearization method is utilized in the iterative linear-quadratic-Gaussian method for locally-optimal control and estimation of nonlinear stochastic systems.[33] However, the linearization of the bilinear system leads to a linear model and lost a bilinear term so it cannot describe the bilinear behavior of the original bilinear system. This motivates us to study new state estimation methods to solve the state filtering problem for bilinear systems. The main objectives of this paper lie in the following:

- to design a bilinear state estimator for the single-input–single-output bilinear system by minimizing the covariance matrix of the state estimation errors based on the delta operator;
- to extend the proposed bilinear state filtering algorithm to multiple-input–multiple-output bilinear systems;
- to demonstrate the effectiveness of the proposed methods through the convergence analysis and the numerical example.

3 | THE BILINEAR STATE ESTIMATOR

In this section, we derive a bilinear state estimator based on the extremum principle for state estimation and choose a suitable gain vector so that the state estimation error is minimal, which is similar to the requirement of the KF for the linear case.

3.1 | The derivation of the bilinear state estimator

Similar to the state observer, we construct a bilinear state estimator

$$\hat{\boldsymbol{x}}_{k+1}=\boldsymbol{A}\hat{\boldsymbol{x}}_k+\boldsymbol{B}\hat{\boldsymbol{x}}_ku_k+\boldsymbol{f}u_k+\boldsymbol{L}_k(y_k-\boldsymbol{c}\hat{\boldsymbol{x}}_k-du_k), \tag{7}$$

where $\hat{\boldsymbol{x}}_k$ is the state estimation vector of $\boldsymbol{x}_k$ and $\boldsymbol{L}_k$ is the gain vector to be determined. Define the state estimation error

$$\tilde{\boldsymbol{x}}_k:=\boldsymbol{x}_k-\hat{\boldsymbol{x}}_k. \tag{8}$$

Then, we have

$$\tilde{\boldsymbol{x}}_{k+1}=(\boldsymbol{A}-\boldsymbol{L}_k\boldsymbol{c})\tilde{\boldsymbol{x}}_k+\boldsymbol{B}\tilde{\boldsymbol{x}}_ku_k+\boldsymbol{w}_k-\boldsymbol{L}_kv_k. \tag{9}$$

Then, we have

$$\mathrm{E}[\tilde{\boldsymbol{x}}_{k+1}] = (\boldsymbol{A} - \boldsymbol{L}_k\boldsymbol{c})\mathrm{E}[\tilde{\boldsymbol{x}}_k] + \boldsymbol{B}\mathrm{E}[\tilde{\boldsymbol{x}}_k]u_k.$$

Define the state estimation error covariance matrix $\boldsymbol{P}_k$ as

$$\boldsymbol{P}_k = \mathrm{E}\left[\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\right]. \tag{10}$$

Thus, $\boldsymbol{P}_{k+1}$ can be expressed as

$$\begin{aligned}\boldsymbol{P}_{k+1} &= \mathrm{E}\left[\tilde{\boldsymbol{x}}_{k+1}\tilde{\boldsymbol{x}}_{k+1}^{\mathrm{T}}\right]\\ &= (\boldsymbol{A} - \boldsymbol{L}_k\boldsymbol{c})\mathrm{E}\left[\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\right]\left(\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}}\right) + (\boldsymbol{A} - \boldsymbol{L}_k\boldsymbol{c})\mathrm{E}\left[\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k^{T}\right]\boldsymbol{B}^{\mathrm{T}}u_k\\ &\quad + \boldsymbol{B}u_k\mathrm{E}\left[\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\right]\left(\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}}\right) + \boldsymbol{B}u_k\mathrm{E}\left[\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k^{T}\right]\boldsymbol{B}^{\mathrm{T}}u_k + \boldsymbol{R}_w + \boldsymbol{L}_kR_v\boldsymbol{L}_k^{T}.\end{aligned} \tag{11}$$

The aim is to choose an optimal gain vector $\boldsymbol{L}_k$ to minimize the state estimation error covariance matrix $\boldsymbol{P}_{k+1}$. From (11), we find that $\boldsymbol{P}_{k+1}$ is complicated, and it is difficult to compute the partial derivative of $\boldsymbol{P}_{k+1}$ with respect to $\boldsymbol{L}_k$ on the condition that the order $n \geqslant 2$. Therefore, we first consider the scalar case that the system order $n = 1$, then we have

$$\begin{aligned}P_{k+1} &= \mathrm{E}\left[\tilde{x}_{k+1}^2\right]\\ &= (A - L_kc)^2P_k + 2(A - L_kc)Bu_kP_k + R_w + L_k^2R_v + B^2u_k^2P_k.\end{aligned} \tag{12}$$

Calculating the partial derivative of P_{k+1} by L_k yields

$$\frac{\partial P_{k+1}}{\partial L_k} = 2[A - L_kc]cP_k - 2cBP_ku_k + 2L_kR_v = 0, \tag{13}$$

which gives

$$L_k = \frac{AcP_k + cBP_ku_k}{c^2P_k + R_v}. \tag{14}$$

Suppose that $n \geqslant 2$. Equation (11) gives

$$\begin{aligned}\boldsymbol{P}_{k+1} &= (\boldsymbol{A} - \boldsymbol{L}_k\boldsymbol{c})\boldsymbol{P}_k\left(\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}u_k\right) + \boldsymbol{R}_w\\ &\quad + \boldsymbol{B}u_k\boldsymbol{P}_k\left(\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}u_k\right) + \boldsymbol{L}_kR_v\boldsymbol{L}_k^{\mathrm{T}}.\end{aligned} \tag{15}$$

Remark 3. It is difficult in the matrix case to compute the partial derivative of (15). Li and Todorov[33] computed the partial derivative of the trace of the covariance matrix with respect to the gain vector for the purpose of computing the optimal filter gain, which is different from the method in this paper. The basic idea of the derivation in this paper is to minimize the magnitude of the estimation errors by introducing the delta operator to minimize the state estimation error covariance matrix for the purpose of computing the optimal gain vector $\boldsymbol{L}_k$.

Assume that $\boldsymbol{L}_k$ is the optimal gain vector to minimize the state estimation error covariance matrix $\boldsymbol{P}_{k+1}$, that is, $\boldsymbol{P}_{k+1}$ is the minimal state estimation covariance matrix. Obviously, if there exists the departure $\delta\boldsymbol{L}_k$ from the filtering gain vector to the optimal gain vector $\boldsymbol{L}_k$, the estimation error covariance matrix obtained from (15) will deviate from the minimal $\boldsymbol{P}_{k+1}$ and reaches $\boldsymbol{P}_{k+1} + \delta\boldsymbol{P}_{k+1}$, where $\delta\boldsymbol{P}_{k+1}$ is the nonnegative definite matrix. From (15), we find that $\boldsymbol{L}_k + \delta\boldsymbol{L}_k$ and $\boldsymbol{P}_{k+1} + \delta\boldsymbol{P}_{k+1}$ satisfy

$$\begin{aligned}\boldsymbol{P}_{k+1} + \delta\boldsymbol{P}_{k+1} &= [\boldsymbol{A} - (\boldsymbol{L}_k + \delta\boldsymbol{L}_k)\boldsymbol{c}]\boldsymbol{P}_k[\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{c}^{\mathrm{T}}(\boldsymbol{L}_k + \delta\boldsymbol{L}_k)^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}u_k]\\ &\quad + \boldsymbol{B}u_k\boldsymbol{P}_k[\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{c}^{\mathrm{T}}(\boldsymbol{L}_k + \delta\boldsymbol{L}_k)^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}u_k]\\ &\quad + \boldsymbol{R}_w + (\boldsymbol{L}_k + \delta\boldsymbol{L}_k)R_v(\boldsymbol{L}_k + \delta\boldsymbol{L}_k)^{\mathrm{T}},\end{aligned} \tag{16}$$

where $\boldsymbol{P}_{k+1}$ and $\boldsymbol{L}_k$ satisfy (15). Substituting (15) into (16) gives

$$\begin{aligned}\delta\boldsymbol{P}_{k+1} &= (\boldsymbol{A}-\boldsymbol{L}_k\boldsymbol{c}-\delta\boldsymbol{L}_k\boldsymbol{c})\boldsymbol{P}_k\left(\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}}u_k-\boldsymbol{c}^{\mathrm{T}}\delta\boldsymbol{L}_k^{\mathrm{T}}\right)+\boldsymbol{B}u_k\boldsymbol{P}_k\left(\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}}u_k\right)\\ &\quad-\boldsymbol{B}u_k\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}\delta\boldsymbol{L}_k^{\mathrm{T}}+\boldsymbol{R}_w+\boldsymbol{L}_kR_v\boldsymbol{L}_k^{\mathrm{T}}+\boldsymbol{L}_kR_v\delta\boldsymbol{L}_k^{\mathrm{T}}+\delta\boldsymbol{L}_kR_v\boldsymbol{L}_k^{\mathrm{T}}+\delta\boldsymbol{L}_kR_v\delta\boldsymbol{L}_k^{\mathrm{T}}-\boldsymbol{P}_{k+1}\\ &= -\delta\boldsymbol{L}_k\left(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}}+\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{B}^{\mathrm{T}}u_k-R_v\boldsymbol{L}_k^{\mathrm{T}}\right)\\ &\quad-\left(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}}+\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{B}^{\mathrm{T}}u_k-R_v\boldsymbol{L}_k^{\mathrm{T}}\right)^{\mathrm{T}}\delta\boldsymbol{L}_k^{\mathrm{T}}+\delta\boldsymbol{L}_k(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}+R_v)\delta\boldsymbol{L}_k^{\mathrm{T}}\\ &= \boldsymbol{W}_k+\boldsymbol{W}_k^{\mathrm{T}}+\delta\boldsymbol{L}_k(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}+R_v)\delta\boldsymbol{L}_k^{\mathrm{T}},\end{aligned}\tag{17}$$

where $\boldsymbol{W}_k := -\delta\boldsymbol{L}_k(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}}+\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{B}^{\mathrm{T}}u_k-R_v\boldsymbol{L}_k^{\mathrm{T}})$. If we take

$$\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}}+\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{B}^{\mathrm{T}}u_k-R_v\boldsymbol{L}_k^{\mathrm{T}}=\boldsymbol{0},$$

then we can obtain

$$\boldsymbol{L}_k=(\boldsymbol{A}+\boldsymbol{B}u_k)\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}+R_v)^{-1}.\tag{18}$$

Thus, we have $\boldsymbol{W}_k=\boldsymbol{0}$ and

$$\delta\boldsymbol{P}_{k+1}=\delta\boldsymbol{L}_k(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}+R_v)\delta\boldsymbol{L}_k^{\mathrm{T}}.\tag{19}$$

Remark 4. From (19), we can see that $\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}+R_v$ is nonnegative because R_v is nonnegative, $\boldsymbol{P}_k$ is nonnegative. If $\delta\boldsymbol{L}_k\neq\boldsymbol{0}$, then $\delta\boldsymbol{P}_{k+1}$ is the nonnegative definite matrix. This explains that the nonnegative deviation $\delta\boldsymbol{P}_{k+1}$ of the minimal covariance matrix $\boldsymbol{P}_{k+1}$ is generated when any departure $\delta\boldsymbol{L}_k$ affects the optimal gain vector $\boldsymbol{L}_k$. Therefore, $\boldsymbol{L}_k=(\boldsymbol{A}+\boldsymbol{B}u_k)\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}+R_v)^{-1}$ in (18) is the optimal gain vector, which makes the state estimation error covariance matrix minimal.

To summarize, the bilinear state estimator is as follows:

$$\hat{\boldsymbol{x}}_{k+1}=\boldsymbol{A}\hat{\boldsymbol{x}}_k+\boldsymbol{B}\hat{\boldsymbol{x}}_ku_k+\boldsymbol{f}u_k+\boldsymbol{L}_k(y_k-\boldsymbol{c}\hat{\boldsymbol{x}}_k-du_k),\tag{20}$$

$$\boldsymbol{L}_k=\boldsymbol{A}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}+R_v)^{-1}+\boldsymbol{B}u_k\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}+R_v)^{-1},\tag{21}$$

$$\boldsymbol{P}_{k+1}=(\boldsymbol{A}-\boldsymbol{L}_k\boldsymbol{c}+\boldsymbol{B}u_k)\boldsymbol{P}_k\left(\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}}u_k\right)+\boldsymbol{R}_w+\boldsymbol{L}_kR_v\boldsymbol{L}_k^{\mathrm{T}}.\tag{22}$$

Let $n=1$, Equation (21) reduces to

$$L_k=\frac{AcP_k+cBP_ku_k}{c^2P_k+R_v},\tag{23}$$

which means Equation (14).

3.2 | Theoretical analysis of the bilinear state estimator

From (9), we have

$$\tilde{\boldsymbol{x}}_{k+1}=(\boldsymbol{G}_k-\boldsymbol{L}_k\boldsymbol{c})\tilde{\boldsymbol{x}}_k+\boldsymbol{\eta}_k,\tag{24}$$

where

$$\boldsymbol{G}_k := \boldsymbol{A}+\boldsymbol{B}u_k,\tag{25}$$

$$\boldsymbol{\eta}_k := \boldsymbol{w}_k-\boldsymbol{L}_kv_k.\tag{26}$$

In order to analyze the error dynamics, we give some lemmas.

Definition 1. If there exist real numbers $0<\varrho,\varsigma<\infty$ and $0<\phi<1$ such that

$$\mathrm{E}\{\|\tilde{\boldsymbol{x}}_k\|^2\}\leqslant\varrho\|\tilde{\boldsymbol{x}}_0\|^2\phi^k+\varsigma,\quad k=1,2,3,\ \dots,\tag{27}$$

then the stochastic process $\tilde{\boldsymbol{x}}_k$ is called exponentially bounded in mean square.

Definition 2. If the stochastic process satisfies

$$\sup_{k\geqslant 0} \|\tilde{\boldsymbol{x}}_k\| < \infty, \text{ a.s., } k = 1, 2, 3, \ \dots, \tag{28}$$

then the stochastic process is said to be bounded with probability one.

Lemma 1. *For the state estimation error in (9), if there exist a stochastic process* $V_k(\tilde{\boldsymbol{x}}_k)$ *and real numbers* $\chi_1, \chi_2, \mu > 0$ *and* $0 < \beta < 1$ *such that*

$$\chi_1 \|\tilde{\boldsymbol{x}}_k\|^2 \leqslant V_k(\tilde{\boldsymbol{x}}_k) \leqslant \chi_2 \|\tilde{\boldsymbol{x}}_k\|^2 \tag{29}$$

and

$$\mathrm{E}\{V_{k+1}(\tilde{\boldsymbol{x}}_{k+1})|\tilde{\boldsymbol{x}}_k\} - V_k(\tilde{\boldsymbol{x}}_k) \leqslant \mu - \beta V_k(\tilde{\boldsymbol{x}}_k) \tag{30}$$

are fulfilled for every solution of (24), then the stochastic process is exponentially bounded in mean square, ie, we have

$$\mathrm{E}\{\|\tilde{\boldsymbol{x}}_k\|^2\} \leqslant \frac{\chi_2}{\chi_1}\|\tilde{\boldsymbol{x}}_0\|^2(1-\beta)^k + \frac{\mu}{\chi_1}\sum_{i=1}^{k-1}(1-\beta)^i \tag{31}$$

$$\leqslant \frac{\chi_2}{\chi_1}\|\tilde{\boldsymbol{x}}_0\|^2(1-\beta)^k + \frac{\mu}{\chi_1\beta}, \quad k = 1, 2, 3, \ \dots. \tag{32}$$

Moreover, the stochastic process $V_k(\tilde{x}_k)$ *is bounded with probability one.*[34,35]

Lemma 2. *For the system in (1)-(2) and the bilinear state estimator in (20)-(22), assume the conditions that* $\boldsymbol{G}_k$ *is nonsingular and there exist positive real numbers* $0 < g, c_1 < \infty$ *and* $p_1, p_2 > 0$ *such that* $\|\boldsymbol{G}_k\| \leqslant g$, $\|\boldsymbol{c}\| \leqslant c_1$, $p_1\boldsymbol{I}_n \leqslant \boldsymbol{P}_k \leqslant p_2\boldsymbol{I}_n$ *hold. Let* $0 < \beta := 1 - 1/[1 + \frac{q_1}{(g-gp_2c_1^2/r)^2}] < 1$ *and* $\boldsymbol{\Xi}_k = \boldsymbol{P}_k^{-1}$. *Then, we have*

$$(\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})^{\mathrm{T}}\boldsymbol{\Xi}_{k+1}(\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c}) \leqslant (1-\beta)\boldsymbol{\Xi}_k, \quad k = 1, 2, 3, \ \dots, \tag{33}$$

with $\boldsymbol{L}_k$ *given by (21).*

Proof. From (21) and (22), we have

$$\boldsymbol{P}_{k+1} = \boldsymbol{G}_k\boldsymbol{P}_k\boldsymbol{G}_k^{\mathrm{T}} + \boldsymbol{R}_w + \boldsymbol{L}_kR_v\boldsymbol{L}_k^{\mathrm{T}} \tag{34}$$

$$= (\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})\boldsymbol{P}_k(\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})^{\mathrm{T}} + \boldsymbol{R}_w + (\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}}. \tag{35}$$

Inserting (21) into (35) obtains

$$(\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})\boldsymbol{P}_k = \boldsymbol{G}_k\boldsymbol{P}_k - \boldsymbol{G}_k\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}} + R_v)^{-1}\boldsymbol{c}\boldsymbol{P}_k. \tag{36}$$

Multiplying both sides of (36) by $\boldsymbol{G}_k^{-1}$ gives

$$\boldsymbol{G}_k^{-1}(\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})\boldsymbol{P}_k = \boldsymbol{P}_k - \boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}}(\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}} + R_v)^{-1}\boldsymbol{c}\boldsymbol{P}_k, \tag{37}$$

which is a symmetric matrix. Then, we apply the matrix inversion lemma to obtain

$$\boldsymbol{G}_k^{-1}(\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})\boldsymbol{P}_k = \left(\boldsymbol{P}_k^{-1} + \boldsymbol{c}^{\mathrm{T}}R_v^{-1}\boldsymbol{c}\right)^{-1} > 0 \tag{38}$$

because $\boldsymbol{P}_k^{-1}$ is positive definite. Moreover, since $\boldsymbol{P}_k, R_v > 0$, from (21), we conclude that

$$\boldsymbol{P}_{k+1} \geqslant (\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})\boldsymbol{P}_k(\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})^{\mathrm{T}} + \boldsymbol{R}_w \geqslant \boldsymbol{R}_w.$$

Inequality (38) implies that the inverse of the matrix $\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c}$ exists and then we may obtain

$$\boldsymbol{P}_{k+1} \geqslant (\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})[\boldsymbol{P}_k + (\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})^{-1}\boldsymbol{R}_w(\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})^{-\mathrm{T}}](\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})^{\mathrm{T}}. \tag{39}$$

From (21), the conditions in Lemma 2, and $\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}} \geqslant 0$, we have $\|\boldsymbol{L}_k\| \leqslant \frac{gp_2c_1}{r}$ and obtain

$$\boldsymbol{P}_{k+1} \geqslant (\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})\left[\boldsymbol{P}_k + \frac{q_1}{\left(g - gp_2c_1^2/r\right)^2}\boldsymbol{I}_n\right](\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})^{\mathrm{T}}. \tag{40}$$

Using the assumptions that $\boldsymbol{P}_k \geqslant p_1\boldsymbol{I}_n$ and $\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c}$ is nonsingular, by taking the inverse of both sides of (40), and premultiplying and postmultiplying by $(\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})^{\mathrm{T}}$ and $\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c}$, we get

$$\begin{aligned}(\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c})^{\mathrm{T}}\Xi_{k+1}(\boldsymbol{G}_k - \boldsymbol{L}_k\boldsymbol{c}) &\leqslant \left[\boldsymbol{P}_k + \frac{q_1}{\left(g - gp_2c_1^2/r\right)^2}\boldsymbol{I}_n\right]^{-1} \\ &= \left[1 + \frac{q_1}{\left(g - gp_2c_1^2/r\right)^2}\right]^{-1}\Xi_k \\ &= (1-\beta)\Xi_k. \end{aligned} \tag{41}$$

The proof is completed. □

Lemma 3. *For the system in (1)-(2) and the state estimator in (20)-(22), assume that the initial estimation error satisfies* $\|\tilde{\boldsymbol{x}}_0\| \leqslant \varepsilon$ *and the covariance matrices of the noise terms are bounded via* $q_1\boldsymbol{I}_n \leqslant \boldsymbol{R}_w \leqslant \alpha\boldsymbol{I}_n$, $r \leqslant R_v \leqslant \alpha$ *for some* $q_1, r, \alpha, \varepsilon > 0$. *Let* $\Xi_k = \boldsymbol{P}_k^{-1}$ *and* $\kappa_n := n/p_1 + (gp_2c_1)^2/(p_1r^2) > 0$ *independent of* α. *Then, we have*

$$\boldsymbol{E}\left\{\boldsymbol{\eta}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{\eta}_k\right\} \leqslant \kappa_n\alpha, \quad k = 1, 2, 3, \ldots, \tag{42}$$

with $\boldsymbol{\eta}_k$ *given by (26).*

Proof. Since the noises $\boldsymbol{w}_k$ and v_k are uncorrelated white noise, we establish the following relation:

$$\begin{aligned}\boldsymbol{\eta}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{\eta}_k &= (\boldsymbol{w}_k - \boldsymbol{L}_kv_k)^{\mathrm{T}}\Xi_{k+1}(\boldsymbol{w}_k - \boldsymbol{L}_kv_k) \\ &= \boldsymbol{w}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{w}_k - \boldsymbol{w}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{L}_kv_k - v_k\boldsymbol{L}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{w}_k + v_k\boldsymbol{L}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{L}_kv_k. \end{aligned} \tag{43}$$

Taking the expectation yields

$$\mathrm{E}\left\{\boldsymbol{\eta}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{\eta}_k\right\} = \mathrm{E}\left\{\boldsymbol{w}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{w}_k\right\} + \mathrm{E}\left\{\boldsymbol{L}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{L}_kv_k^2\right\}. \tag{44}$$

From (21), the conditions in Theorem 1, and $\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}} \geqslant 0$, we have $\|\boldsymbol{L}_k\| \leqslant gp_2c_1/r$. Using (44) and $\boldsymbol{c}\boldsymbol{P}_k\boldsymbol{c}^{\mathrm{T}} \geqslant 0$, we obtain

$$\mathrm{E}\left\{\boldsymbol{\eta}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{\eta}_k\right\} \leqslant \frac{1}{p_1}\mathrm{E}\left\{\boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{w}_k\right\} + \frac{(gp_2c_1)^2}{p_1r^2}\mathrm{E}\left\{v_k^2\right\}. \tag{45}$$

Using $\boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{w}_k = \mathrm{tr}[\boldsymbol{w}_k\boldsymbol{w}_k^{\mathrm{T}}]$ and combining the conditions (A1) to (A3), we have

$$\mathrm{E}\left\{\boldsymbol{\eta}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{\eta}_k\right\} \leqslant \frac{1}{p_1}\mathrm{tr}[\boldsymbol{R}_w] + \frac{(gp_2c_1)^2}{p_1r^2}R_v. \tag{46}$$

Then, using the conditions in Lemma 3 and the definition of κ_n, we have

$$\mathrm{E}\left\{\boldsymbol{\eta}_k^{\mathrm{T}}\Xi_{k+1}\boldsymbol{\eta}_k\right\} \leqslant \kappa_n\alpha.$$

The proof of Lemma 3 is finished. □

Theorem 1. *For the bilinear system in (1)-(2) and the bilinear state estimator in (20)-(22), assume that the conditions in Lemma 2 and Lemma 3 hold. Let* $\chi_1 = 1/p_1$, $\chi_2 = 1/p_2$, *and* $\mu = \kappa_n\alpha$. *For* $k = 1, 2, 3, \ldots,$ *the following inequalities hold:*

$$\chi_1\|\tilde{\boldsymbol{x}}_k\|^2 \leqslant V_k(\tilde{\boldsymbol{x}}_k) \leqslant \chi_2\|\tilde{\boldsymbol{x}}_k\|^2, \tag{47}$$

and

$$\mathrm{E}\{V_{k+1}(\tilde{\boldsymbol{x}}_{k+1})|\tilde{\boldsymbol{x}}_k\} - V_k(\tilde{\boldsymbol{x}}_k) \leqslant \mu - \beta V_k(\tilde{\boldsymbol{x}}_k). \tag{48}$$

Then, the state estimation error $\tilde{\boldsymbol{x}}_k$ *given by (8) is exponentially bounded in mean square and bounded with probability one.*

Proof. Choose a nonnegative function $V_k(\tilde{\boldsymbol{x}}_k) = \tilde{\boldsymbol{x}}_k^{\mathrm{T}} \Xi_k \tilde{\boldsymbol{x}}_k$. From the conditions in Lemma 2, we have

$$\frac{1}{p_2}\|\tilde{\boldsymbol{x}}_k\|^2 \leqslant V_k(\tilde{\boldsymbol{x}}_k) \leqslant \frac{1}{p_1}\|\tilde{\boldsymbol{x}}_k\|^2.$$

For the purpose of satisfying the requirements of Lemma 1, we need an upper bound on $\mathrm{E}\{V_{k+1}(\tilde{\boldsymbol{x}}_{k+1})|\tilde{\boldsymbol{x}}_k\}$. From (24), we have

$$V_{k+1}(\tilde{\boldsymbol{x}}_{k+1}) = [\boldsymbol{G}_k - \boldsymbol{L}_k \boldsymbol{c}\tilde{\boldsymbol{x}}_k + \boldsymbol{\eta}_k]^{\mathrm{T}} \Xi_{k+1} [\boldsymbol{G}_k - \boldsymbol{L}_k \boldsymbol{c}\tilde{\boldsymbol{x}}_k + \boldsymbol{\eta}_k]. \tag{49}$$

Applying Lemma 2 and combining $V_k(\tilde{\boldsymbol{x}}_k) = \tilde{\boldsymbol{x}}_k^{\mathrm{T}} \Xi_k \tilde{\boldsymbol{x}}_k$ obtain

$$\begin{aligned} V_{k+1}(\tilde{\boldsymbol{x}}_{k+1}) \leqslant{}& (1-\beta)V_k(\tilde{\boldsymbol{x}}_k) + \boldsymbol{\eta}_k^{\mathrm{T}} \Xi_{k+1} \boldsymbol{\eta}_k \\ &+ \tilde{\boldsymbol{x}}_k^{\mathrm{T}} (\boldsymbol{G}_k - \boldsymbol{L}_k \boldsymbol{c})^{\mathrm{T}} \Xi_{k+1} \boldsymbol{\eta}_k + \boldsymbol{\eta}_k^{\mathrm{T}} \Xi_{k+1} (\boldsymbol{G}_k - \boldsymbol{L}_k \boldsymbol{c}) \tilde{\boldsymbol{x}}_k. \end{aligned} \tag{50}$$

Because the terms Ξ_{k+1}, $\boldsymbol{G}_k$, $\boldsymbol{L}_k\boldsymbol{c}$, $\boldsymbol{c}$, and $\tilde{\boldsymbol{x}}_k$ are independent of $\boldsymbol{\eta}_k$. Taking the conditional expectation $\mathrm{E}\{V_{k+1}(\tilde{\boldsymbol{x}}_{k+1})|\tilde{\boldsymbol{x}}_k\}$ and applying Lemma 3 yield

$$\begin{aligned} \mathrm{E}\{V_{k+1}(\tilde{\boldsymbol{x}}_{k+1})|\tilde{\boldsymbol{x}}_k\} - V_k(\tilde{\boldsymbol{x}}_k) &\leqslant \kappa_n \alpha - \beta V_k(\tilde{\boldsymbol{x}}_k) \\ &= \mu - \beta V_k(\tilde{\boldsymbol{x}}_k), \end{aligned}$$

for $\|\tilde{\boldsymbol{x}}_k\| \leqslant \varepsilon$. According to Lemma 1, we conclude that the stochastic process $V_k(\tilde{\boldsymbol{x}}_k)$ is exponentially bounded in mean square and the stochastic process $V_k(\tilde{\boldsymbol{x}}_k)$ is bounded with probability one. However, we must notice that the supermartingale inequality

$$\mathrm{E}\{V_{k+1}(\tilde{\boldsymbol{x}}_{k+1})|\tilde{\boldsymbol{x}}_k\} - V_k(\tilde{\boldsymbol{x}}_k) \leqslant \kappa_n \alpha - \beta V_k(\tilde{\boldsymbol{x}}_k) \leqslant 0 \tag{51}$$

is fulfilled to guarantee the boundedness of the state estimation error for $\varepsilon' \leqslant \|\tilde{\boldsymbol{x}}_k\| \leqslant \varepsilon$. Choosing $\alpha \leqslant \frac{\beta \varepsilon'^2}{\kappa_n p_1}$, we have

$$\kappa_n \alpha \leqslant \beta \frac{1}{p_1} \varepsilon'^2 \leqslant \beta \frac{1}{p_1} \|\tilde{\boldsymbol{x}}_k\|^2 \leqslant \beta V_k(\tilde{\boldsymbol{x}}_k),$$

then Inequality (51) holds. Thus, we conclude that the state estimation error $\tilde{\boldsymbol{x}}_k$ remains bounded if the certain conditions are satisfied. The proof of Theorem 1 is finished.

In general, covariance matrix $\boldsymbol{R}_w$ of the process noise vector $\boldsymbol{w}_k$ and the variance R_v of the observation noise v_k in (21)-(22) are unknown. Therefore, the unknown $\boldsymbol{R}_w$ and R_v in (21)-(22) may be replaced with their estimates $\hat{\boldsymbol{R}}_{w,k}$ and $\hat{R}_{v,k}$, ie,

$$\hat{\boldsymbol{R}}_{w,k} = \frac{1}{k}\sum_{j=1}^{k} (\hat{\boldsymbol{x}}_{j+1} - \boldsymbol{A}\hat{\boldsymbol{x}}_j - \boldsymbol{B}\hat{\boldsymbol{x}}_j u_j - \boldsymbol{f} u_j)(\hat{\boldsymbol{x}}_{j+1} - \boldsymbol{A}\hat{\boldsymbol{x}}_j - \boldsymbol{B}\hat{\boldsymbol{x}}_j u_j - \boldsymbol{f} u_j)^{\mathrm{T}} \in \mathbb{R}^{n\times n}, \tag{52}$$

$$\hat{R}_{v,k} = \frac{1}{k}\sum_{j=1}^{k} (y_j - \boldsymbol{c}\hat{\boldsymbol{x}}_j - d u_j)^2 \in \mathbb{R}. \tag{53}$$

Replacing R_v and $\boldsymbol{R}_w$ with their estimates $\hat{R}_{v,k}$ and $\hat{\boldsymbol{R}}_{w,k}$ obtains the bilinear state estimator. The steps of computing the state estimate $\hat{\boldsymbol{x}}_k$ in (20)-(22) and (52)-(53) are listed in the following.

1. Let $k = 1$, set the initial values $\hat{\boldsymbol{x}}_1 = \mathbf{1}_n$, $\boldsymbol{P}_1 = \boldsymbol{I}_n$, $u_{k-i} = 0$, and $y_{k-i} = 0$, for $i = 1, 2, \dots, n$, and the system parameters $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{f}, \boldsymbol{c}$, and d.
2. Collect the input-output data u_k and y_k.
3. Compute the gain vector $\boldsymbol{L}_k$ by (21) and the covariance matrix $\boldsymbol{P}_{k+1}$ by (22).
4. Update the state estimates $\hat{\boldsymbol{x}}_{k+1}$ by (20).
5. Compute the covariance matrix $\hat{\boldsymbol{R}}_{w,k}$ by (52) and the variance $\hat{R}_{v,k}$ by (53).
6. Increase k by 1 and go to Step 2 and continue the recursive calculation.

□

4 | EXTENSION FOR THE MULTIVARIATE BILINEAR SYSTEM

In this section, we consider the following multiple-input–multiple-output bilinear system:

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \sum_{i=1}^{m} \boldsymbol{B}_i \boldsymbol{x}_k u_{i,k} + \boldsymbol{f}\boldsymbol{u}_k + \boldsymbol{w}_k, \tag{54}$$

$$\boldsymbol{y}_k = \boldsymbol{c}\boldsymbol{x}_k + \boldsymbol{d}\boldsymbol{u}_k + \boldsymbol{v}_k, \tag{55}$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ is the system state vector, $\boldsymbol{u}_k \in \mathbb{R}^m$ is the system input vector, $\boldsymbol{y}_k \in \mathbb{R}^l$ is the system output vector, $\boldsymbol{v}_k \in \mathbb{R}^l$ is an uncorrelated random noise vector with zero mean, $\boldsymbol{w}_k \in \mathbb{R}^n$ is an uncorrelated process noise vector with zero mean, and $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{B}_i \in \mathbb{R}^{n\times n}$, $\boldsymbol{f} \in \mathbb{R}^{n\times m}$, $\boldsymbol{c} \in \mathbb{R}^{l\times n}$, and $\boldsymbol{d} \in \mathbb{R}^{l\times m}$ are the parameter matrices of the system.

Assume that the state estimation algorithm adopts as follows:

$$\hat{\boldsymbol{x}}_{k+1} = \boldsymbol{A}\hat{\boldsymbol{x}}_k + \sum_{i=1}^{m} \boldsymbol{B}_i \hat{\boldsymbol{x}}_k u_{i,k} + \boldsymbol{f}\boldsymbol{u}_k + \boldsymbol{L}_{1,k}(\boldsymbol{y}_k - \boldsymbol{c}\hat{\boldsymbol{x}}_k - \boldsymbol{d}\boldsymbol{u}_k), \tag{56}$$

where $\boldsymbol{L}_{1,k} \in \mathbb{R}^{n\times l}$ is the gain matrix. Then, the state estimation error vector is formed as

$$\tilde{\boldsymbol{x}}_{k+1} = (\boldsymbol{A} - \boldsymbol{L}_{1,k}\boldsymbol{c})\tilde{\boldsymbol{x}}_k + \boldsymbol{w}_k - \boldsymbol{L}_{1,k}\boldsymbol{v}_k. \tag{57}$$

Define the state estimation error covariance matrix

$$\boldsymbol{P}_{1,k+1} := \mathrm{E}\left[\tilde{\boldsymbol{x}}_{k+1}\tilde{\boldsymbol{x}}_{k+1}^{\mathrm{T}}\right]. \tag{58}$$

Hence, we have

$$\begin{aligned}\boldsymbol{P}_{1,k+1} = {} & (\boldsymbol{A} - \boldsymbol{L}_{1,k})\boldsymbol{P}_{1,k}\left[(\boldsymbol{A} - \boldsymbol{L}_{1,k}\boldsymbol{c})^{\mathrm{T}} + \sum_{i=1}^{m} \boldsymbol{B}_i^{\mathrm{T}} u_{i,k}\right] \\ & + \sum_{i=1}^{m} \boldsymbol{B}_i \boldsymbol{P}_{1,k}\left[(\boldsymbol{A} - \boldsymbol{L}_{1,k}\boldsymbol{c})^{\mathrm{T}} u_{i,k} + u_{i,k}\sum_{j=1}^{m} \boldsymbol{B}_j^{\mathrm{T}} u_{j,k}\right] + \boldsymbol{R}_w + \boldsymbol{L}_{1,k}\boldsymbol{R}_v\boldsymbol{L}_{1,k}^{\mathrm{T}}.\end{aligned} \tag{59}$$

Similarly, introducing the departure $\delta\boldsymbol{L}_{1,k}$ from the filtering gain vector to the optimal gain vector makes the estimation error covariance matrix obtained from (59) reach $\boldsymbol{P}_{1,k+1} + \delta\boldsymbol{P}_{1,k+1}$. $\delta\boldsymbol{P}_{1,k+1}$ is the nonnegative definite matrix. From (59), we find that $\boldsymbol{L}_{1,k} + \delta\boldsymbol{L}_{1,k}$ and $\boldsymbol{P}_{1,k+1} + \delta\boldsymbol{P}_{1,k+1}$ satisfy

$$\begin{aligned}\boldsymbol{P}_{1,k+1} + \delta\boldsymbol{P}_{1,k+1} = {} & [\boldsymbol{A} - (\boldsymbol{L}_{1,k} + \delta\boldsymbol{L}_{1,k})\boldsymbol{c}]\boldsymbol{P}_{1,k}\left[\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{c}^{\mathrm{T}}(\boldsymbol{L}_{1,k} + \delta\boldsymbol{L}_{1,k})^{\mathrm{T}} + \sum_{i=1}^{m} \boldsymbol{B}_i^{\mathrm{T}} u_{i,k}\right] \\ & + \sum_{i=1}^{m} \boldsymbol{B}_i \boldsymbol{P}_{1,k}\left\{[\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{c}^{\mathrm{T}}(\boldsymbol{L}_{1,k} + \delta\boldsymbol{L}_{1,k})^{\mathrm{T}}]u_{i,k} + u_{i,k}\sum_{j=1}^{m} \boldsymbol{B}_j^{\mathrm{T}} u_{j,k}\right\} \\ & + \boldsymbol{R}_w + (\boldsymbol{L}_{1,k} + \delta\boldsymbol{L}_{1,k})\boldsymbol{R}_v(\boldsymbol{L}_{1,k} + \delta\boldsymbol{L}_{1,k})^{\mathrm{T}},\end{aligned} \tag{60}$$

where $\boldsymbol{P}_{1,k+1}$ and $\boldsymbol{L}_{1,k}$ satisfy (59). Substituting (59) into (60) obtains

$$\begin{aligned}
\delta\boldsymbol{P}_{1,k+1} &= (\boldsymbol{A}-\boldsymbol{L}_{1,k}\boldsymbol{c}-\delta\boldsymbol{L}_{1,k}\boldsymbol{c})\boldsymbol{P}_{1,k}\left(\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_{1,k}^{\mathrm{T}}-\boldsymbol{c}^{\mathrm{T}}\delta\boldsymbol{L}_{1,k}^{\mathrm{T}}+\sum_{i=1}^{m}\boldsymbol{B}_i^{\mathrm{T}}u_{i,k}\right)\\
&\quad+\sum_{i=1}^{m}\boldsymbol{B}_i\boldsymbol{P}_{1,k}\left(\boldsymbol{A}^{\mathrm{T}}u_{i,k}-\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_{1,k}^{\mathrm{T}}+u_{i,k}\sum_{j=1}^{m}\boldsymbol{B}_j^{\mathrm{T}}u_{j,k}-\boldsymbol{c}^{\mathrm{T}}\delta\boldsymbol{L}_{1,k}^{\mathrm{T}}u_{i,k}\right)+\boldsymbol{R}_w+\boldsymbol{L}_{1,k}\boldsymbol{R}_v\boldsymbol{L}_{1,k}^{\mathrm{T}}\\
&\quad+\boldsymbol{L}_{1,k}\boldsymbol{R}_v\delta\boldsymbol{L}_{1,k}^{\mathrm{T}}+\delta\boldsymbol{L}_{1,k}\boldsymbol{R}_v\boldsymbol{L}_{1,k}^{\mathrm{T}}+\delta\boldsymbol{L}_{1,k}\boldsymbol{R}_v\delta\boldsymbol{L}_{1,k}^{\mathrm{T}}-\boldsymbol{P}_{1,k+1}\\
&= -\delta\boldsymbol{L}_{1,k}\left(\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_{1,k}^{\mathrm{T}}+\boldsymbol{c}\boldsymbol{P}_{1,k}\sum_{i=1}^{m}\boldsymbol{B}_i^{\mathrm{T}}u_{i,k}-\boldsymbol{R}_v\boldsymbol{L}_{1,k}^{\mathrm{T}}\right)\\
&\quad-\left(\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_{1,k}^{\mathrm{T}}+\boldsymbol{c}\boldsymbol{P}_{1,k}\sum_{i=1}^{m}\boldsymbol{B}_i^{\mathrm{T}}u_{i,k}-\boldsymbol{R}_v\boldsymbol{L}_{1,k}^{\mathrm{T}}\right)^{\mathrm{T}}\delta\boldsymbol{L}_{1,k}^{\mathrm{T}}+\delta\boldsymbol{L}_{1,k}(\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{L}_v)\delta\boldsymbol{L}_{1,k}^{\mathrm{T}}\\
&= \boldsymbol{W}_k+\boldsymbol{W}_k^{\mathrm{T}}+\delta\boldsymbol{L}_{1,k}(\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{R}_v)\delta\boldsymbol{L}_{1,k}^{\mathrm{T}},
\end{aligned} \tag{61}$$

where

$$\boldsymbol{W}_k := -\delta\boldsymbol{L}_{1,k}\left(\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_{1,k}^{\mathrm{T}}+\boldsymbol{c}\boldsymbol{P}_{1,k}\sum_{i=1}^{m}\boldsymbol{B}_i^{\mathrm{T}}u_{i,k}-\boldsymbol{R}_v\boldsymbol{L}_{1,k}^{\mathrm{T}}\right). \tag{62}$$

Taking

$$\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{L}_{1,k}^{\mathrm{T}}+\boldsymbol{c}\boldsymbol{P}_{1,k}\sum_{i=1}^{m}\boldsymbol{B}_i^{\mathrm{T}}u_{i,k}-\boldsymbol{R}_v\boldsymbol{L}_{1,k}^{\mathrm{T}}=\boldsymbol{0}$$

obtains the optimal gain matrix

$$\boldsymbol{L}_{1,k}=\left(\boldsymbol{P}\boldsymbol{A}_{1,k}\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}\sum_{i=1}^{m}\boldsymbol{B}_iu_{i,k}\right)(\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{R}_v)^{-1}. \tag{63}$$

Thus, we have

$$\delta\boldsymbol{P}_{1,k+1}=\delta\boldsymbol{L}_{1,k}(\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{R}_v)\delta\boldsymbol{L}_{1,k}^{\mathrm{T}}. \tag{64}$$

According to the explanation in Remark 4, we can conclude that $\boldsymbol{L}_{1,k}$ in (63) is the optimal gain vector that makes the state estimation error covariance matrix minimal. Then, the bilinear state estimator for the multiple-input–multiple-output bilinear system is as follows:

$$\hat{\boldsymbol{x}}_{k+1}=\boldsymbol{A}\hat{\boldsymbol{x}}_k+\sum_{i=1}^{m}\boldsymbol{B}_i\hat{\boldsymbol{x}}_ku_{i,k}+\boldsymbol{f}\boldsymbol{u}_k+\boldsymbol{L}_{1,k}(\boldsymbol{y}_k-\boldsymbol{c}\hat{\boldsymbol{x}}_k-\boldsymbol{d}\boldsymbol{u}_k), \tag{65}$$

$$\boldsymbol{L}_{1,k}=\left(\boldsymbol{A}\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}\sum_{i=1}^{m}\boldsymbol{B}_iu_{i,k}\right)(\boldsymbol{c}\boldsymbol{P}_{1,k}\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{R}_v)^{-1}, \tag{66}$$

$$\begin{aligned}
\boldsymbol{P}_{1,k+1} &= (\boldsymbol{A}-\boldsymbol{L}_{1,k})\boldsymbol{P}_{1,k}\left[(\boldsymbol{A}-\boldsymbol{L}_{1,k}\boldsymbol{c})^{\mathrm{T}}+\sum_{i=1}^{m}\boldsymbol{B}_i^{\mathrm{T}}u_{i,k}\right]\\
&\quad+\sum_{i=1}^{m}\boldsymbol{B}_i\boldsymbol{P}_{1,k}\left[(\boldsymbol{A}-\boldsymbol{L}_{1,k}\boldsymbol{c})^{\mathrm{T}}u_{i,k}+u_{i,k}\sum_{j=1}^{m}\boldsymbol{B}_j^{\mathrm{T}}u_{j,k}\right]+\boldsymbol{R}_w+\boldsymbol{L}_{1,k}\boldsymbol{R}_v\boldsymbol{L}_{1,k}^{\mathrm{T}}.
\end{aligned} \tag{67}$$

Replacing the unknown $\boldsymbol{R}_w$ and $\boldsymbol{R}_v$ in (66)-(67) with their estimates $\hat{\boldsymbol{R}}_{w,k}$ and $\hat{\boldsymbol{R}}_{v,k}$ obtains the following:

$$\hat{\boldsymbol{x}}_{k+1}=\boldsymbol{A}\hat{\boldsymbol{x}}_k+\sum_{i=1}^{m}\boldsymbol{B}_i\hat{\boldsymbol{x}}_ku_{i,k}+\boldsymbol{f}\boldsymbol{u}_k+\boldsymbol{L}_{1,k}(\boldsymbol{y}_k-\boldsymbol{c}\hat{\boldsymbol{x}}_k-\boldsymbol{d}\boldsymbol{u}_k), \tag{68}$$

$$L_{1,k} = \left(AP_{1,k}c^{\mathrm{T}} + P_{1,k}c^{\mathrm{T}}\sum_{i=1}^{m}B_i u_{i,k}\right)(cP_{1,k}c^{\mathrm{T}} + \hat{R}_{v,k})^{-1}, \tag{69}$$

$$P_{1,k+1} = (A - L_{1,k})P_{1,k}\left[(A - L_{1,k}c)^{\mathrm{T}} + \sum_{i=1}^{m}B_i^{\mathrm{T}}u_{i,k}\right] + \sum_{i=1}^{m}B_iP_{1,k}\left[(A - L_{1,k}c)^{\mathrm{T}}u_{i,k} + u_{i,k}\sum_{j=1}^{m}B_j^{\mathrm{T}}u_{j,k}\right] + \hat{R}_{w,k} + L_{1,k}\hat{R}_{v,k}L_{1,k}^{\mathrm{T}}, \tag{70}$$

$$\hat{R}_{w,k} = \frac{1}{k}\sum_{j=1}^{k}\left[\hat{x}_{j+1} - A\hat{x}_j - \sum_{i=1}^{m}B_ix_ju_{i,j} - fu_j\right]\left[\hat{x}_{j+1} - A\hat{x}_j - \sum_{i=1}^{m}B_ix_ju_{i,j} - fu_j\right]^{\mathrm{T}} \in \mathbb{R}^{n\times n}, \tag{71}$$

$$\hat{R}_{v,k} = \frac{1}{k}\sum_{j=1}^{k}[y_j - c\hat{x}_j - du_j][y_j - c\hat{x}_j - du_j]^{\mathrm{T}} \in \mathbb{R}^{l\times l}. \tag{72}$$

Equations (68)-(72) form the state estimation algorithm for the multiple-input–multiple-output bilinear system in (54)-(55). The proposed methods can be used to study parameter estimation and state filtering of linear systems[36-40] and nonlinear systems with colored noise.[41-43]

5 | EXAMPLE

The numerical example is selected to test the effectiveness of the proposed bilinear state estimator. Consider the following bilinear system:

$$x_{k+1} = Ax_k + Bx_ku_k + fu_k + w_k,$$
$$y_k = cx_k + du_k + v_k.$$

In simulation, input $\{u_k\}$ is taken as a persistently excited signal, and $\{v_k\}$ and $\{w_k\}$ as uncorrelated Gaussian white noise sequence with zero mean and variance R_v and R_w, respectively. The initial values include $\hat{x}_1 = \mathbf{1}_n, P_1 = I_n$.

Case I: Consider a second-order bilinear state-space system, whose parameters are

$$A = \begin{bmatrix} 0.20 & 0.25 \\ 0.25 & -0.35 \end{bmatrix}, \quad B = \begin{bmatrix} 0.20 & -0.15 \\ 0.10 & -0.17 \end{bmatrix},$$
$$f = \begin{bmatrix} -0.45 \\ -0.25 \end{bmatrix}, \quad c = [0.30, \ 0.25], \quad d = 0.70.$$

In simulation, input $\{u_k\}$ is taken as a persistent excitation signal sequence with zero mean and unit variance. The covariance matrix of the process noise w_k is R_w, and the variance of the measurement noise v_k is R_v. Figure 1 shows the simulated

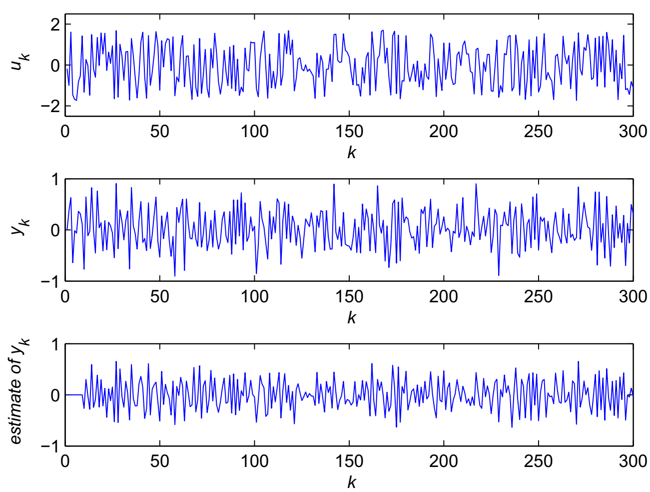

FIGURE 1 System input u_k, output y_k, and its estimate $\hat{y}_k$ [Colour figure can be viewed at wileyonlinelibrary.com]

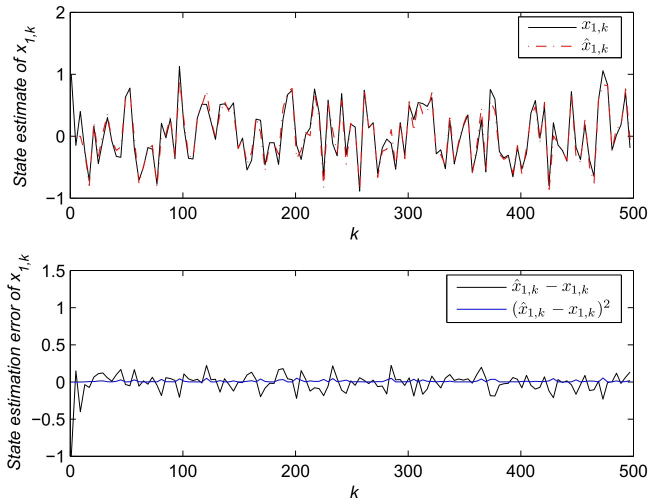

FIGURE 2 The state estimate $\hat{x}_{1,k}$ and the estimation errors versus k ($R_v = 0.20^2$, $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_2$) [Colour figure can be viewed at wileyonlinelibrary.com]

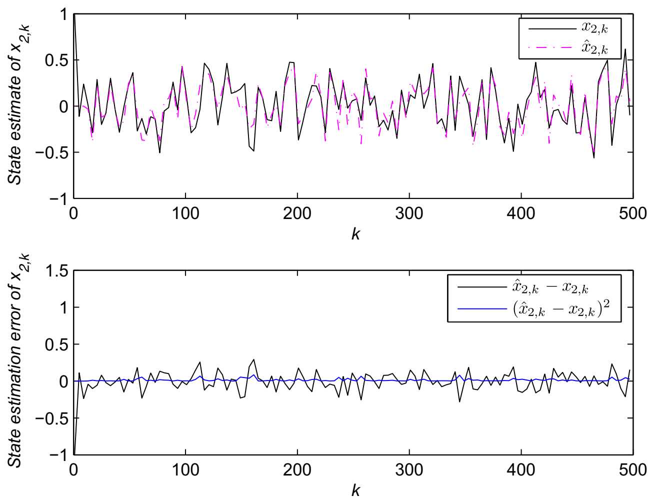

FIGURE 3 The state estimate $\hat{x}_{2,k}$ and the estimation errors versus k ($R_v = 0.20^2$, $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_2$) [Colour figure can be viewed at wileyonlinelibrary.com]

input-output data and the predicted output data. Figure 2 shows the true states $x_{1,k}$ and its estimated value $\hat{x}_{1,k}$ and their estimation errors computed by $\hat{x}_{1,k} - x_{1,k}$ and $(\hat{x}_{1,k} - x_{1,k})^2$. Figure 3 shows the true state $\hat{x}_{2,k}$ and its estimate $\hat{x}_{2,k}$ and their estimation errors $\hat{x}_{2,k} - x_{2,k}$ and $(\hat{x}_{2,k} - x_{2,k})^2$. Then, the root mean square error (RMSE) is used to describe the error between the true state $x_{i,k}$ and its estimated value $\hat{x}_{i,k}$, and the error between the true output y_k and its predicted output $\hat{y}_k$, which are defined as

$$Error_x = \left\{ \frac{1}{L} \sum_{k=1}^{L} [\hat{x}_{i,k} - x_{i,k}]^2 \right\}^{1/2},$$

$$Error_y = \left\{ \frac{1}{L} \sum_{k=1}^{L} [\hat{y}_k - y_k]^2 \right\}^{1/2}.$$

To show the influence of the noise level on the proposed algorithm, we study the bilinear state estimator in (20)-(22) with the noise variances $R_v = 0.10^2$ and $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_2$, $R_v = 0.15^2$ and $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_2$, $R_v = 0.20^2$ and $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_2$, $R_v = 0.25^2$ and $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_2$, $R_v = 0.30^2$ and $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_2$. The RMSE results are shown in Table 1.

TABLE 1 The root mean square errors (RMSEs) under different noise levels

Noise Level		RMSE		
R_v	$\boldsymbol{R}_w$	$x_{1,k}$	$x_{2,k}$	y_k
0.10^2	$0.10^2\boldsymbol{I}_2$	0.08545	0.07354	0.14067
0.15^2	$0.10^2\boldsymbol{I}_2$	0.08549	0.07346	0.18116
0.20^2	$0.10^2\boldsymbol{I}_2$	0.08548	0.07341	0.22609
0.25^2	$0.10^2\boldsymbol{I}_2$	0.08546	0.07338	0.27327
0.30^2	$0.10^2\boldsymbol{I}_2$	0.08544	0.07336	0.32173

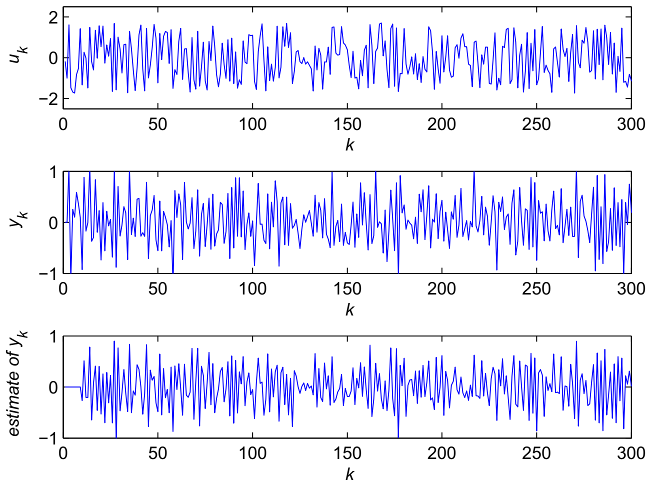

FIGURE 4 System input u_k, output y_k, and its estimate $\hat{y}_k$ [Colour figure can be viewed at wileyonlinelibrary.com]

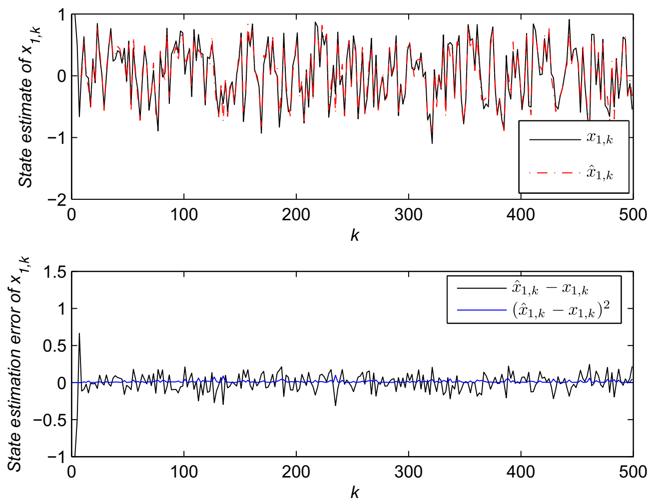

FIGURE 5 The state estimate $\hat{x}_{1,k}$ and the estimation errors versus k ($R_v = 0.20^2$, $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_3$) [Colour figure can be viewed at wileyonlinelibrary.com]

Case II: Consider a third-order bilinear state-space system, whose parameters are

$$\boldsymbol{A} = \begin{bmatrix} 0.38 & -0.15 & -0.21 \\ 0.20 & -0.35 & -0.15 \\ 0.32 & -0.25 & -0.20 \end{bmatrix}, \quad \boldsymbol{f} = \begin{bmatrix} -0.45 \\ -0.65 \\ -0.35 \end{bmatrix},$$

$$\boldsymbol{B} = \begin{bmatrix} 0.20 & -0.15 & -0.05 \\ 0.19 & -0.15 & -0.10 \\ 0.15 & -0.15 & -0.10 \end{bmatrix}, \quad d = 0.20, \quad \boldsymbol{c} = [0.30, \ 0.25, \ 0.15]. \tag{73}$$

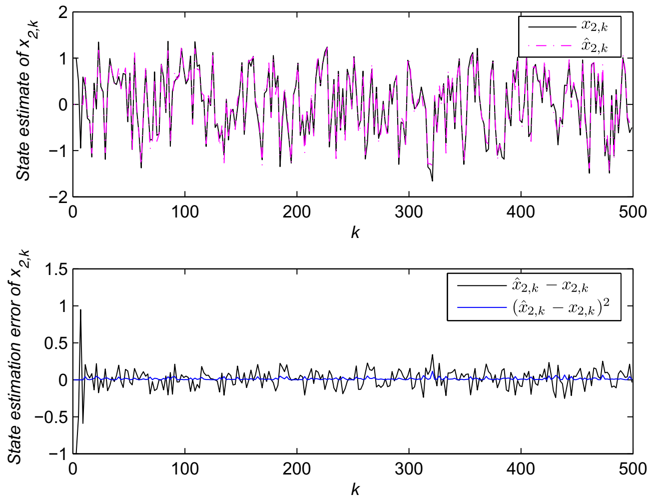

FIGURE 6 The state estimate $\hat{x}_{2,k}$ and the estimation errors versus k ($R_v = 0.20^2$, $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_3$) [Colour figure can be viewed at wileyonlinelibrary.com]

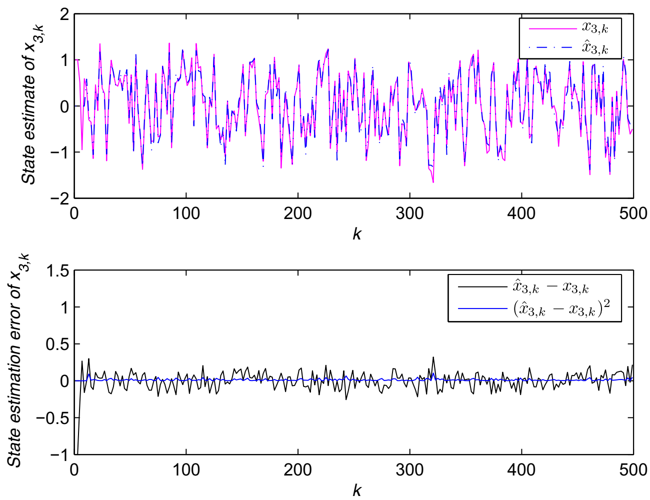

FIGURE 7 The state estimate $\hat{x}_{3,k}$ and the estimation errors versus k ($R_v = 0.20^2$, $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_3$) [Colour figure can be viewed at wileyonlinelibrary.com]

Assume that the covariance matrix of the process noise $\boldsymbol{w}_k$ is $\boldsymbol{R}_w = 0.10\boldsymbol{I}_3$, and the variance of the measurement noise v_k is $R_v = 0.20$. The simulation conditions are same as Case I. The proposed bilinear state estimator is used to obtain the state estimates. Figure 4 shows the system input-output data and the estimated output. Figures 5 to 7 show the true system states $x_{1,k}$, $x_{2,k}$, and $x_{3,k}$, the estimated states $\hat{x}_{1,k}$, $\hat{x}_{2,k}$, and $\hat{x}_{3,k}$, and their estimation errors $\hat{x}_{1,k} - x_{1,k}$ and $(\hat{x}_{1,k} - x_{1,k})^2$, $\hat{x}_{2,k} - x_{2,k}$ and $(\hat{x}_{2,k} - x_{2,k})^2$, and $\hat{x}_{3,k} - x_{3,k}$ and $(\hat{x}_{3,k} - x_{3,k})^2$. Under different noise levels $R_v = 0.10^2$ and $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_3$, $R_v = 0.15^2$ and $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_3$, $R_v = 0.20^2$ and $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_3$, $R_v = 0.25^2$ and $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_3$, $R_v = 0.30^2$ and $\boldsymbol{R}_w = 0.10^2\boldsymbol{I}_3$, the RMSEs between the true state $x_{i,k}$ and its estimated value $\hat{x}_{i,k}$, and the error between the true output y_k and its predicted output $\hat{y}_k$ are shown in Table 2.

From Figures 1 to 7 and Tables 1 to 2, we can draw the following conclusions.

- The bilinear state estimator has good performance because the estimated states are close to their true values with k increasing and the estimation errors between the true states and the estimated states are quite small (see Figures 2 and 3 and Figures 5 to 7).
- The bilinear state estimator can generate good estimates because the estimated output is close to the true output and the RMSEs under different noise levels are close to the noise standard deviation (see Tables 1 and 2, and Figures 1 and 4).

TABLE 2 The root mean square errors (RMSEs) under different noise levels

Noise Level		RMSE			
R_v	R_w	$x_{1,k}$	$x_{2,k}$	$x_{3,k}$	y_k
0.10^2	$0.10^2 I_3$	0.10490	0.11541	0.09902	0.12590
0.15^2	$0.10^2 I_3$	0.10497	0.11550	0.09908	0.16902
0.20^2	$0.10^2 I_3$	0.10502	0.11555	0.09911	0.21551
0.25^2	$0.10^2 I_3$	0.10506	0.11559	0.09913	0.26359
0.30^2	$0.10^2 I_3$	0.10509	0.11563	0.09914	0.31253

6 | CONCLUSIONS

This paper proposes a bilinear state estimator for the single-input–single-output bilinear state-space system based on the delta operator. Different from the previous linearization method like Taylor expansion, we take use of the special structure of the bilinear system and propose the state filtering algorithm to obtain the unknown states by minimizing the covariance matrix of the state estimation errors based on the extremum principle. Moreover, the bilinear state estimator is extended for the multiple-input–multiple-output bilinear system. Finally, the convergence analysis and the simulation results show that the proposed state estimator has good performance in the state estimation of bilinear systems. The methods proposed in this paper can combine some statistical optimal strategies[44-47] to study the parameter estimation algorithms of linear and nonlinear systems[48-52] and can be applied to other fields,[53-59] such as fault detection, image processing, and sliding mode control.[60-63]

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China under Grant 61873111, the 111 Project (B12018), and the Postgraduate Research and Practice Innovation Program of Jiangsu Province under Grant KYCX18_1854.

ORCID

Xiao Zhang https://orcid.org/0000-0002-6413-6148
Feng Ding https://orcid.org/0000-0002-2721-2025

REFERENCES

1. Xu L. A proportional differential control method for a time-delay system using the Taylor expansion approximation. *Appl Math Comput*. 2014;236:391-399.
2. Xu L. Application of the Newton iteration algorithm to the parameter estimation for dynamical systems. *J Comput Appl Math*. 2015;288:33-43.
3. Xu L, Chen L, Xiong W. Parameter estimation and controller design for dynamic systems from the step responses based on the Newton iteration. *Nonlinear Dynamics*. 2015;79(3):2155-2163.
4. Chen G-Y, Gan M, Chen CP, Li H-X. A regularized variable projection algorithm for separable nonlinear least-squares problems. *IEEE Trans Autom Control*. 2019;64(2):526-537.
5. Chen G-Y, Gan M, Ding F, Chen CP. Modified Gram-Schmidt method-based variable projection algorithm for separable nonlinear models. *IEEE Trans Neural Netw Learn Syst*. 2019. https://doi.org/10.1109/TNNLS.2018.2884909
6. dos Santos PL, Ramos JA, de Carvalho JLM. Identification of bilinear systems with white noise inputs: an iterative deterministic-stochastic subspace approach. *IEEE Trans Control Syst Technol*. 2009;17(5):1145-1153.
7. Vicario F, Phan MQ, Betti R, Longman RW. Linear state representations for identification of bilinear discrete-time models by interaction matrices. *Nonlinear Dynamics*. 2014;77(4):1561-1576.
8. Dai H, Sinha NK. Robust identification of systems using block-pulse functions. *IEE Proc D Control Theory Appl*. 1992;139(3):308-316.
9. Hizir NB, Phan MQ, Betti R, Longman RW. Identification of discrete-time bilinear systems through equivalent linear models. *Nonlinear Dynamics*. 2012;69(4):2065-2078.
10. Singh TSD, Chatterjee A. A comparative study of adaptation algorithms for nonlinear system identification based on second order Volterra and bilinear polynomial filters. *Measurement*. 2011;44(10):1915-1923.
11. Cao Y, Li P, Zhang Y. Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing. *Futur Gener Comput Syst*. 2018;88:279-283.

12. Cao Y, Ma L, Xiao S, Zhang X, Xu W. Standard analysis for transfer delay in CTCS-3. *Chin J Electron*. 2017;26(5):1057-1063.
13. Cao Y, Lu H, Wen T. A safety computer system based on multi-sensor data processing. *Sensors*. 2019;19(4). https://doi.org/10.3390/s19040818
14. Cao Y, Zhang Y, Wen T, Li P. Research on dynamic nonlinear input prediction of fault diagnosis based on fractional differential operator equation in high-speed train control system. *Chaos*. 2019;29(1). Article ID 013130. https://aip.scitation.org/doi/10.1063/1.5085397
15. Ding F, Wang X, Mao L, Xu L. Joint state and multi-innovation parameter estimation for time-delay linear systems and its convergence based on the Kalman filtering. *Digit Signal Process*. 2017;62:211-223.
16. Xu L, Ding F. Recursive least squares and multi-innovation stochastic gradient parameter estimation methods for signal modeling. *Circuits Syst Signal Process*. 2017;36(4):1735-1753.
17. Xu L, Ding F. Iterative parameter estimation for signal models based on measured data. *Circuits Syst Signal Process*. 2018;37(7):3046-3069.
18. Xu L, Xiong W, Alsaedi A, Hayat T. Hierarchical parameter estimation for the frequency response based on the dynamical window data. *Int J Control Autom Syst*. 2018;16(4):1756-1764.
19. Xu L. The parameter estimation algorithms based on the dynamical response measurement data. *Adv Mech Eng*. 2017;9(11):1-12. https://doi.org/10.1177/1687814017730003
20. Ding F. State filtering and parameter estimation for state space systems with scarce measurements. *Signal Processing*. 2014;104:369-380.
21. Zhao S, Shmaliy YS, Liu F. Fast Kalman-like optimal unbiased FIR filtering with applications. *IEEE Trans Signal Process*. 2016;64(9):2284-2297.
22. Zhao S, Shmaliy YS, Liu F, Khan SH. Unbiased, optimal, and in-betweens: the trade-off in discrete finite impulse response filtering. *IET Signal Process*. 2016;10(4):325-334.
23. Johnston LA, Krishnamurthy V. Finite dimensional smoothers for MAP state estimation of bilinear systems. *IEEE Trans Signal Process*. 1999;47(9):2444-2459.
24. Germani A, Manes C, Palumbo P. Linear filtering for bilinear stochastic differential systems with unknown inputs. *IEEE Trans Autom Control*. 2002;47(10):1726-1730.
25. Kulikov GY, Kulikova MV. Accurate continuous-discrete unscented Kalman filtering for estimation of nonlinear continuous-time stochastic models in radar tracking. *Signal Processing*. 2017;139:25-35.
26. Hara S, Furuta K. Minimal order state observers for bilinear systems. *Int J Control*. 1976;24(5):705-718.
27. Tsai SH. A global exponential fuzzy observer design for time-delay Takagi-Sugeno uncertain discrete fuzzy bilinear systems with disturbance. *IEEE Trans Fuzzy Syst*. 2012;20(6):1063-1075.
28. Gomez-Exposito A, Gomez-Quiles C, de la Villa Jaen A. Bilinear power system state estimation. *IEEE Trans Power Syst*. 2012;27(1):493-501.
29. Phan MQ, Vicario F, Longman RW, Betti R. Optimal bilinear observers for bilinear state-space models by interaction matrices. *Int J Control*. 2015;88(8):1504-1522.
30. Zhang X, Xu L, Ding F, Hayat T. Combined state and parameter estimation for a bilinear state space system with moving average noise. *J Franklin Inst*. 2018;355(6):3079-3103.
31. Zhang X, Ding F, Alsaadi FE, Hayat T. Recursive parameter identification of the dynamical models for bilinear state space systems. *Nonlinear Dynamics*. 2017;89(4):2415-2429.
32. Zhang X, Ding F, Xu L, Yang E. State filtering-based least squares parameter estimation for bilinear systems using the hierarchical identification principle. *IET Control Theory Appl*. 2018;12(12):1704-1713.
33. Li W, Todorov E. Iterative linearization methods for approximately optimal control and estimation of non-linear stochastic system. *Int J Control*. 2017;80(9):1439-1453.
34. Agniel RG, Jury EI. Almost sure boundedness of randomly sampled systems. *SIAM J Control*. 1971;9(3):372-384.
35. Tarn T-J, Rasis Y. Observers for nonlinear stochastic systems. *IEEE Trans Autom Control*. 1976;21(4):441-448.
36. Xu L. The damping iterative parameter identification method for dynamical systems based on the sine signal measurement. *Signal Processing*. 2016;120:660-667.
37. Xu L, Ding F. Parameter estimation algorithms for dynamical response signals based on the multi-innovation theory and the hierarchical principle. *IET Signal Process*. 2017;11(2):228-237.
38. 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 Processing*. 2017;140:97-103.
39. Xu L, Ding F. Parameter estimation for control systems based on impulse responses. *Int J Control Autom Syst*. 2017;15(6):2471-2479.
40. Ding F, Liu PX, Liu G. Gradient based and least-squares based iterative identification methods for OE and OEMA systems. *Digit Signal Process*. 2010;20(3):664-677.
41. Ding F, Liu X, Chu J. Gradient-based and least-squares-based iterative algorithms for Hammerstein systems using the hierarchical identification principle. *IET Control Theory Appl*. 2013;7(2):176-184.
42. Xu H, Ding F, Yang E. Modeling a nonlinear process using the exponential autoregressive time series model. *Nonlinear Dynamics*. 2019;95(3):2079-2092.
43. Wang Y, Ding F. A filtering based multi-innovation gradient estimation algorithm and performance analysis for nonlinear dynamical systems. *IMA J Appl Math*. 2017;82(6):1171-1191.
44. Wen Y, Yin C. Solution of Hamilton-Jacobi-Bellman equation in optimal reinsurance strategy under dynamic VaR constraint. *J Funct Spaces*. 2019. Article ID 6750892. https://doi.org/10.1155/2019/6750892

45. Tian X, Niu H. A bi-objective model with sequential search algorithm for optimizing network-wide train timetables. *Comput Ind Eng*. 2019;127:1259-1272.
46. Yang F, Zhang P, Li X-X. The truncation method for the Cauchy problem of the inhomogeneous Helmholtz equation. *Applicable Analysis*. 2019;98(5):991-1004.
47. Wu T, Shi X, Liao L, Zhou C, Zhou H, Su Y. A capacity configuration control strategy to alleviate power fluctuation of hybrid energy storage system based on improved particle swarm optimization. *Energies*. 2019;12(4):1-11. Article ID 642. https://doi.org/10.3390/en12040642
48. Ding F. Decomposition based fast least squares algorithm for output error systems. *Signal Processing*. 2013;93(5):1235-1242.
49. Ding F. Two-stage least squares based iterative estimation algorithm for CARARMA system modeling. *Appl Math Model*. 2013;37(7):4798-4808.
50. Ding J, Chen J, Lin J, Wan L. Particle filtering based parameter estimation for systems with output-error type model structures. *J Franklin Inst*. 2019:356. https://doi.org/10.1016/j.jfranklin.2019.04.027
51. Pan J, Ma H, Jiang X, Ding W, Ding F. Adaptive gradient-based iterative algorithm for multivariate controlled autoregressive moving average systems using the data filtering technique. *Complexity*. 2018;2018. Article ID 9598307. https://doi.org/10.1155/2018/9598307
52. Liu Q, Ding F. Auxiliary model-based recursive generalized least squares algorithm for multivariate output-error autoregressive systems using the data filtering. *Circuits Syst Signal Process*. 2019;38(2):590-610.
53. Wu M, Li X, Liu C, et al. Robust global motion estimation for video security based on improved k-means clustering. *J Ambient Intell Humaniz Comput*. 2019;10(2):439-448.
54. Wan X-K, Wu H, Qiao F, et al. Electrocardiogram baseline wander suppression based on the combination of morphological and wavelet transformation based filtering. *Comput Math Methods Med*. 2019;2019. Article ID 7196156. https://doi.org/10.1155/2019/7196156
55. Ma F, Yin Y, Li M. Start-up process modelling of sediment microbial fuel cells based on data driven. *Math Probl Eng*. 2019;2019. Article ID 7403732. https://doi.org/10.1155/2019/7403732
56. Sun Z-Y, Zhang D, Meng Q, Chen C-C. Feedback stabilisation of time-delay nonlinear systems with continuous time-varying output function. *Int J Syst Sci*. 2019;50(2):244-255.
57. Sha X, Xu Z, Yin C. Elliptical distribution-based weight-determining method for ordered weighted averaging operators. *Int J Intell Syst*. 2019;34(5):858-877.
58. Fu B, Ouyang C, Li C, Wang J, Gul E. An improved mixed integer linear programming approach based on symmetry diminishing for unit commitment of hybrid power system. *Energies*. 2019;12(5):1-14. Article ID 833. https://doi.org/10.3390/en12050833
59. Shi WX, Liu N, Zhou YM, Cao XA. Effects of postannealing on the characteristics and reliability of polyfluorene organic light-emitting diodes. *IEEE Trans Electron Devices*. 2019;66(2):1057-1062.
60. Wang T, Liu L, Zhang J, Schaeffer E, Wang Y. A M-EKF fault detection strategy of insulation system for marine current turbine. *Mech Syst Signal Process*. 2019;115:269-280.
61. Feng L, Li Q, Li Y. Imaging with 3-D aperture synthesis radiometers. *IEEE Trans Geosci Remote Sens*. 2019;57(4):2395-2406.
62. Pan J, Li W, Zhang H. Control algorithms of magnetic suspension systems based on the improved double exponential reaching law of sliding mode control. *Int J Control Autom Syst*. 2018;16(6):2878-2887.
63. Liu N, Mei S, Sun D, et al. Effects of charge transport materials on blue fluorescent organic light-emitting diodes with a host-dopant system. *Micromachines*. 2019;10(5):1-10. Article Number 344. https://doi.org/10.3390/mi10050344

How to cite this article: Zhang X, Ding F, Yang E. State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors. *Int J Adapt Control Signal Process*. 2019;1–17. https://doi.org/10.1002/acs.3027