

Research Article

Filtering Based Recursive Least Squares Algorithm for Multi-Input Multioutput Hammerstein Models

Ziyun Wang, Yan Wang, and Zhicheng Ji

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

Correspondence should be addressed to Yan Wang; yanwangjn@163.com

Received 28 June 2014; Revised 7 September 2014; Accepted 25 September 2014; Published 16 October 2014

Academic Editor: Haranath Kar

Copyright © 2014 Ziyun Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper considers the parameter estimation problem for Hammerstein multi-input multioutput finite impulse response (FIR-MA) systems. Filtered by the noise transfer function, the FIR-MA model is transformed into a controlled autoregressive model. The key-term variable separation principle is used to derive a data filtering based recursive least squares algorithm. The numerical examples confirm that the proposed algorithm can estimate parameters more accurately and has a higher computational efficiency compared with the recursive least squares algorithm.

1. Introduction

Parameter estimation is an important approach to model dynamical systems and has been widely used in estimating the parameters for nonlinear systems [1–3], deriving system identification methods [4–7], identifying state-space models [8, 9], and developing solutions for matrix equations [10–13]. For example, Dehghan and Hajarian discussed several solution methods for different matrix equations [14–16]. In the area of system control and modeling, Shi and Fang developed a Kalman filter based identification for systems with randomly missing measurements [17], gave output feedback stabilization [18], and presented a robust mixed H_2/H_∞ control of networked control systems [19].

The least squares algorithm is a fundamental method [20–22] and many methods such as the iterative algorithm [23, 24] and the gradient algorithm [25] are widely used in the parameter estimation. In the field of Hammerstein system identification, several methods have been developed [26]. For example, a least squares based iterative algorithm and an auxiliary model based recursive least squares algorithm have been presented, respectively, for Hammerstein nonlinear ARMAX systems and Hammerstein output error systems [27, 28]; a Newton recursive algorithm and a Newton iterative algorithm for Hammerstein controlled autoregressive systems are presented in [29].

Consider a multi-input multioutput (MIMO) Hammerstein finite impulse response (FIR) system depicted by

$$\mathbf{y}(t) = \mathbf{B}(z) \bar{\mathbf{u}}(t) + N(z) \mathbf{v}(t), \quad (1)$$

where $\mathbf{u}(t) := [u_1(t), u_2(t), \dots, u_r(t)]^T \in \mathbb{R}^r$ is the nonlinear system input vector with zero mean and unit variances, $\mathbf{y}(t) \in \mathbb{R}^m$ is the measurement of $\mathbf{x}(t) := \mathbf{B}(z)\bar{\mathbf{u}}(t)$ but is corrupted by $\mathbf{w}(t) := N(z)\mathbf{v}(t)$, $\mathbf{v}(t) \in \mathbb{R}^m$ is the white noise vector with zero mean, and $\mathbf{B}(z)$ and $N(z)$ are polynomials in the unit backward shift operator z^{-1} [$z^{-1}\mathbf{y}(t) = \mathbf{y}(t-1)$]:

$$\mathbf{B}(z) := \mathbf{I}_{m \times r} + \sum_{i=1}^{n_b} \mathbf{B}_i z^{-i}, \quad \mathbf{B}_i \in \mathbb{R}^{m \times r}, \quad (2)$$

$$N(z) := 1 + \sum_{j=1}^{n_d} d_j z^{-j}, \quad d_j \in \mathbb{R}^1.$$

It is obvious that the relation between sizes m and r would influence the model identification of this multi-input multi-output Hammerstein system. For example, the dimension of the output vector is not less than that of the input vector if $m \geq r$; otherwise, when $m < r$, the output size is smaller compared with that of the input vector. In this paper, we discuss the identification problem with $m \geq r$.

The nonlinear block in the Hammerstein model is a linear combination of the known basis $\mathbf{f} := (f_1, f_2, \dots, f_r)$:

$$\begin{aligned} \bar{\mathbf{u}}(t) &= \mathbf{f}(\mathbf{u}(t)) \\ &:= [f_1(u_1(t)), f_2(u_2(t)), \dots, f_r(u_r(t))]^T \in \mathbb{R}^r, \end{aligned} \quad (3)$$

where the superscript T denotes the matrix transpose. The function $f_i(u_i(t))$ in (3) is a nonlinear function of a known basis $(\gamma_1, \gamma_2, \dots, \gamma_{n_c})$:

$$f_i(u_i(t)) := \sum_{j=1}^{n_c} c_j \gamma_j(u_i(t)), \quad (4)$$

where the coefficients $\mathbf{c} := (c_1, c_2, \dots, c_{n_c})$ are unknown. Substituting (4) into (3) yields

$$\bar{\mathbf{u}}(t) = \sum_{i=1}^{n_c} c_i \gamma_i(\mathbf{u}(t)),$$

$$\gamma_i(\mathbf{u}(t)) := [\gamma_i(u_1(t)), \gamma_i(u_2(t)), \dots, \gamma_i(u_r(t))]^T \in \mathbb{R}^r. \quad (5)$$

Assume that $\mathbf{u}(t) = \mathbf{0}$, $\mathbf{y}(t) = \mathbf{0}$, and $\mathbf{v}(t) = \mathbf{0}$ for $t \leq 0$, and the orders n_b , n_c , and n_d are known but can be obtained by trial and error. In general, the orders of the Hammerstein model should be large when the nonlinear system is used in prediction; otherwise, the orders should be small if the system is applied for control.

The objective of this paper is to estimate the unknown parameter matrices: \mathbf{B}_i , c_i , d_i from the available input-output data $\{\mathbf{u}(t), \mathbf{y}(t)\}$ of the multivariable Hammerstein finite impulse response moving average (FIR-MA) models [30].

Recently, the filtering idea has received much attention [31–33]. Xiao and Yue studied input nonlinear dynamical adjustment models and presented a recursive generalized least squares algorithm and a filtering based least squares algorithm by replacing the unknown terms in the information vectors with their estimates [34]. The overparameterization method in [34] leads to a redundant estimated product of the nonlinear systems and requires extra computation. Differing from the work in [30, 34, 35], this paper discusses the estimation problem of the MIMO Hammerstein systems using the data filtering idea and transfers the FIR-MA system to controlled autoregressive model by means of the key-term variable separate principle in [36–38]. The proposed algorithm used in this paper can extend to study parameter estimation problems of dual-rate/multirate sampled systems [39–42] and other linear or nonlinear systems [43–46].

Briefly, the rest of this paper is recognized as follows. Section 2 discusses a recursive least squares algorithm for the Hammerstein systems. Section 3 presents a filtering based recursive least squares algorithm by transferring an FIR-MA model to a controlled autoregressive model. Section 4 provides an illustrative example. Finally, some concluding remarks are offered in Section 5.

2. The MRLS Algorithm

For comparison, the MRLS algorithm is listed in Section 2. Here we introduce some notations. t represents the current time in this paper and “ $A =: X$ ” or “ $X := A$ ” means that “ A is defined as X ”. The symbol $\mathbf{I}_{m \times r}$ represents an identity matrix of size r followed by a null matrix of the last $m - r$ rows when $m \geq r$ and vice versa. The norm of a matrix (or a column vector) \mathbf{X} is defined by $\|\mathbf{X}\|^2 := \text{tr}[\mathbf{X}\mathbf{X}^T]$; \otimes denotes the Kronecker product or direct product: if $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$, $\mathbf{B} = [b_{ij}] \in \mathbb{R}^{p \times q}$, then $\mathbf{A} \otimes \mathbf{B} := [a_{ij}\mathbf{B}] \in \mathbb{R}^{mp \times nq}$; $\text{col}[\mathbf{X}]$ is supposed to be the vector formed by the column of the matrix \mathbf{X} : if $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$, then $\text{col}[\mathbf{X}] := [\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_n^T]^T \in \mathbb{R}^{mn}$.

From (1)–(5), the intermediate variables $\mathbf{x}(t)$ and $\mathbf{w}(t)$ and output of the system $\mathbf{y}(t)$ can be expressed as

$$\mathbf{x}(t) = \left[\mathbf{I}_{m \times r} + \sum_{i=1}^{n_b} \mathbf{B}_i z^{-i} \right] \bar{\mathbf{u}}(t), \quad (6)$$

$$\mathbf{w}(t) = \left[1 + \sum_{i=1}^{n_d} d_i z^{-i} \right] \mathbf{v}(t),$$

$$\begin{aligned} \mathbf{y}(t) &= \sum_{i=1}^{n_c} c_i \begin{bmatrix} \gamma_i(\mathbf{u}(t)) \\ \mathbf{0} \end{bmatrix} + \sum_{i=1}^{n_b} \mathbf{B}_i \bar{\mathbf{u}}(t-i) \\ &\quad + \sum_{i=1}^{n_d} d_i \mathbf{v}(t-i) + \mathbf{v}(t). \end{aligned} \quad (7)$$

Note that the subscripts (Roman) s and n denote the first letters of “system” and “noise” for distinguishing the types of the unknown parameter vectors or matrices, respectively. Define the parameter matrix $\boldsymbol{\theta}_1$, the parameter vectors $\boldsymbol{\theta}_2$, $\boldsymbol{\theta}_n$, and the information vectors $\boldsymbol{\varphi}_1(t)$, $\boldsymbol{\varphi}_2(t)$, and $\boldsymbol{\varphi}_n(t)$ as

$$\boldsymbol{\theta}_1^T := [\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_{n_b}] \in \mathbb{R}^{m \times (n_b r)},$$

$$\boldsymbol{\theta}_2 := [c_1, c_2, \dots, c_{n_c}]^T \in \mathbb{R}^{n_c},$$

$$\boldsymbol{\theta}_n := [d_1, d_2, \dots, d_{n_d}]^T \in \mathbb{R}^{n_d},$$

$$\boldsymbol{\varphi}_1(t) := [\bar{\mathbf{u}}^T(t-1), \bar{\mathbf{u}}^T(t-2), \dots, \bar{\mathbf{u}}^T(t-n_b)]^T \in \mathbb{R}^{n_b r},$$

$$\boldsymbol{\varphi}_2(t) := [\boldsymbol{\eta}_1(t), \boldsymbol{\eta}_2(t), \dots, \boldsymbol{\eta}_{n_c}(t)] \in \mathbb{R}^{n_c n_c},$$

$$\boldsymbol{\eta}_i(t) := \begin{bmatrix} \gamma_i(\mathbf{u}(t)) \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^m,$$

$$\boldsymbol{\varphi}_n(t) := [\mathbf{v}(t-1), \mathbf{v}(t-2), \dots, \mathbf{v}(t-n_d)] \in \mathbb{R}^{n_d n_d}. \quad (8)$$

Then, we have

$$\begin{aligned} \mathbf{x}(t) &= \boldsymbol{\theta}_1^T \boldsymbol{\varphi}_1(t) + \boldsymbol{\varphi}_2(t) \boldsymbol{\theta}_2, \\ \mathbf{w}(t) &= \boldsymbol{\varphi}_n(t) \boldsymbol{\theta}_n + \mathbf{v}(t), \\ \mathbf{y}(t) &= \mathbf{x}(t) + \mathbf{w}(t) \\ &= \boldsymbol{\theta}_1^T \boldsymbol{\varphi}_1(t) + \boldsymbol{\varphi}_2(t) \boldsymbol{\theta}_2 + \boldsymbol{\varphi}_n(t) \boldsymbol{\theta}_n + \mathbf{v}(t). \end{aligned} \quad (9)$$

Distinguished from the hierarchical identification methods, we reparameterize the model in (7) by using the Kronecker product to get a parameter matrix $\boldsymbol{\vartheta}$ and by gathering the input information vectors $\boldsymbol{\varphi}_1(t)$ and $\boldsymbol{\varphi}_2(t)$ and output information matrix $\boldsymbol{\varphi}_n(t)$ into one information matrix $\boldsymbol{\Psi}(t)$ to obtain an information matrix $\boldsymbol{\Psi}(t)$ as follows:

$$\begin{aligned}\boldsymbol{\vartheta} &:= \begin{bmatrix} \boldsymbol{\theta}_s \\ \boldsymbol{\theta}_n \end{bmatrix} \in \mathbb{R}^{n_0}, \quad n_0 := mrn_b + n_c + n_d, \\ \boldsymbol{\theta}_s &:= \begin{bmatrix} \text{col} \left[\boldsymbol{\theta}_1^T \right] \\ \boldsymbol{\theta}_2 \end{bmatrix} \in \mathbb{R}^{mrn_b + n_c}, \\ \boldsymbol{\Psi}(t) &:= [\boldsymbol{\varphi}_s(t), \boldsymbol{\varphi}_n(t)] \in \mathbb{R}^{m \times n_0}, \\ \boldsymbol{\varphi}_s(t) &:= [\boldsymbol{\varphi}_1^T(t) \otimes \mathbf{I}_m, \boldsymbol{\varphi}_2(t)] \in \mathbb{R}^{m \times (mrn_b + n_c)}.\end{aligned}\quad (10)$$

Thus, we obtain

$$\mathbf{y}(t) = \boldsymbol{\Psi}(t) \boldsymbol{\vartheta} + \mathbf{v}(t). \quad (11)$$

Equation (11) is the identification model of the multivariable Hammerstein FIR-MA system. Defining and minimizing the cost function

$$J(\boldsymbol{\vartheta}) := \|\mathbf{y}(t) - \boldsymbol{\Psi}(t) \boldsymbol{\vartheta}\|^2 \quad (12)$$

and using the least squares search principle, we obtain the following recursive least squares algorithm [35] to obtain parameter estimates $\hat{\boldsymbol{\vartheta}}(t)$:

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \mathbf{L}(t) [\mathbf{y}(t) - \boldsymbol{\Psi}(t) \hat{\boldsymbol{\vartheta}}(t-1)], \quad (13)$$

$$\mathbf{L}(t) = \mathbf{P}(t-1) \boldsymbol{\Psi}^T(t) [\mathbf{I}_m + \boldsymbol{\Psi}(t) \mathbf{P}(t-1) \boldsymbol{\Psi}^T(t)]^{-1}, \quad (14)$$

$$\mathbf{P}(t) = [\mathbf{I}_{n_0} - \mathbf{L}(t) \boldsymbol{\Psi}(t)] \mathbf{P}(t-1), \quad \mathbf{P}(0) = p_0 \mathbf{I}_{n_0}. \quad (15)$$

Since the information matrix $\boldsymbol{\Psi}(t)$ in (13) contains the unknown intermediate variables $\bar{\mathbf{u}}(t-i)$ and the unmeasurable terms $\mathbf{v}(t-i)$, the recursive algorithm in (13)–(15) cannot compute the parameter estimate $\hat{\boldsymbol{\vartheta}}(t)$. The solution here is replacing the unknown intermediate variables $\bar{\mathbf{u}}(t-i)$ and the unmeasurable terms $\mathbf{v}(t-i)$ in $\boldsymbol{\Psi}(t)$ with the variable estimates (or the outputs of the auxiliary model) $\hat{\mathbf{u}}(t-i)$ and the estimates $\hat{\mathbf{v}}(t-i)$ based on the auxiliary model identification idea. The replaced information matrices are defined as

$$\begin{aligned}\hat{\boldsymbol{\Psi}}(t) &:= [\hat{\boldsymbol{\varphi}}_s(t), \hat{\boldsymbol{\varphi}}_n(t)] \in \mathbb{R}^{m \times n_0}, \\ \hat{\boldsymbol{\varphi}}_s(t) &:= [\hat{\boldsymbol{\varphi}}_1^T(t) \otimes \mathbf{I}_m, \boldsymbol{\varphi}_2(t)] \in \mathbb{R}^{m \times (mrn_b + n_c)}, \\ \hat{\boldsymbol{\varphi}}_1(t) &:= [\hat{\mathbf{u}}^T(t-1), \hat{\mathbf{u}}^T(t-2), \dots, \hat{\mathbf{u}}^T(t-n_b)]^T \in \mathbb{R}^{rn_b}, \\ \hat{\boldsymbol{\varphi}}_n(t) &:= [\hat{\mathbf{v}}(t-1), \hat{\mathbf{v}}(t-2), \dots, \hat{\mathbf{v}}(t-n_d)] \in \mathbb{R}^{m \times n_d}.\end{aligned}\quad (16)$$

Define the parameter estimation matrices

$$\begin{aligned}\hat{\boldsymbol{\vartheta}}(t) &:= \begin{bmatrix} \hat{\boldsymbol{\theta}}_s(t) \\ \hat{\boldsymbol{\theta}}_n(t) \end{bmatrix} \in \mathbb{R}^{n_0}, \\ \hat{\boldsymbol{\theta}}_s(t) &:= \begin{bmatrix} \text{col} \left[\hat{\boldsymbol{\theta}}_1^T(t) \right] \\ \hat{\boldsymbol{\theta}}_2(t) \end{bmatrix} \in \mathbb{R}^{mrn_b + n_c}, \\ \hat{\boldsymbol{\theta}}_1^T(t) &:= [\hat{\mathbf{B}}_1(t), \hat{\mathbf{B}}_2(t), \dots, \hat{\mathbf{B}}_{n_b}(t)] \in \mathbb{R}^{m \times (rn_b)}, \\ \hat{\boldsymbol{\theta}}_2(t) &:= [\hat{c}_1(t), \hat{c}_2(t), \dots, \hat{c}_{n_c}(t)]^T \in \mathbb{R}^{n_c}, \\ \hat{\boldsymbol{\theta}}_n(t) &:= [\hat{d}_1(t), \hat{d}_2(t), \dots, \hat{d}_{n_d}(t)]^T \in \mathbb{R}^{n_d}.\end{aligned}\quad (17)$$

By replacing the parameters c_i ($i = 1, 2, \dots, n_c$) in (4) with $\hat{c}_i(t)$, the output of the proposed auxiliary model $\hat{\mathbf{u}}(t)$ is given by

$$\begin{aligned}\hat{\mathbf{u}}(t) &= \sum_{i=1}^{n_c} \hat{c}_i(t) \boldsymbol{\gamma}_i(\mathbf{u}(t)) \\ &= \hat{c}_1(t) \boldsymbol{\gamma}_1(\mathbf{u}(t)) + \hat{c}_2(t) \boldsymbol{\gamma}_2(\mathbf{u}(t)) \\ &\quad + \dots + \hat{c}_{n_c}(t) \boldsymbol{\gamma}_{n_c}(\mathbf{u}(t)).\end{aligned}\quad (18)$$

From (11), we obtain $\mathbf{v}(t) = \mathbf{y}(t) - \boldsymbol{\Psi}(t) \boldsymbol{\vartheta}$. Replacing $\boldsymbol{\Psi}(t)$ and $\boldsymbol{\vartheta}$ with $\hat{\boldsymbol{\Psi}}(t)$ and $\hat{\boldsymbol{\vartheta}}(t-1)$, the residual $\hat{\mathbf{v}}(t)$ can be written as $\hat{\mathbf{v}}(t) = \mathbf{y}(t) - \hat{\boldsymbol{\Psi}}(t) \hat{\boldsymbol{\vartheta}}(t-1)$.

To summarize, we conclude the following recursive least squares algorithm for multivariable Hammerstein FIR-MA models (the MRLS algorithm for short):

$$\begin{aligned}\hat{\boldsymbol{\vartheta}}(t) &= \hat{\boldsymbol{\vartheta}}(t-1) + \mathbf{L}(t) [\mathbf{y}(t) - \hat{\boldsymbol{\Psi}}(t) \hat{\boldsymbol{\vartheta}}(t-1)], \\ \mathbf{L}(t) &= \mathbf{P}(t-1) \hat{\boldsymbol{\Psi}}^T(t) [\mathbf{I}_m + \hat{\boldsymbol{\Psi}}(t) \mathbf{P}(t-1) \hat{\boldsymbol{\Psi}}^T(t)]^{-1}, \\ \mathbf{P}(t) &= [\mathbf{I}_{n_0} - \mathbf{L}(t) \hat{\boldsymbol{\Psi}}(t)] \mathbf{P}(t-1), \\ \hat{\mathbf{u}}(t) &= \sum_{i=1}^{n_c} \hat{c}_i(t) \boldsymbol{\gamma}_i(\mathbf{u}(t)), \\ \hat{\mathbf{v}}(t) &= \mathbf{y}(t) - \hat{\boldsymbol{\Psi}}(t) \hat{\boldsymbol{\vartheta}}(t-1), \\ \hat{\boldsymbol{\Psi}}(t) &= [\hat{\boldsymbol{\varphi}}_1^T(t) \otimes \mathbf{I}_m, \boldsymbol{\varphi}_2, \hat{\boldsymbol{\varphi}}_n(t)], \\ \hat{\boldsymbol{\varphi}}_1(t) &= [\hat{\mathbf{u}}^T(t-1), \hat{\mathbf{u}}^T(t-2), \dots, \hat{\mathbf{u}}^T(t-n_b)]^T, \\ \boldsymbol{\varphi}_2(t) &= [\boldsymbol{\eta}_1(t), \boldsymbol{\eta}_2(t), \dots, \boldsymbol{\eta}_{n_c}(t)], \\ \boldsymbol{\eta}_i(t) &= \begin{bmatrix} \boldsymbol{\gamma}_i(\mathbf{u}(t)) \\ \mathbf{0} \end{bmatrix}, \\ \hat{\boldsymbol{\varphi}}_n(t) &= [\hat{\mathbf{v}}(t-1), \hat{\mathbf{v}}(t-2), \dots, \hat{\mathbf{v}}(t-n_d)],\end{aligned}$$

$$\begin{aligned}\widehat{\boldsymbol{\theta}}(t) &= \begin{bmatrix} \text{col} \left[\widehat{\boldsymbol{\theta}}_1^T(t) \right] \\ \widehat{\boldsymbol{\theta}}_2(t) \\ \widehat{\boldsymbol{\theta}}_n(t) \end{bmatrix}, \\ \widehat{\boldsymbol{\theta}}_1^T(t) &= [\widehat{\mathbf{B}}_1(t), \widehat{\mathbf{B}}_2(t), \dots, \widehat{\mathbf{B}}_{n_b}(t)], \\ \widehat{\boldsymbol{\theta}}_2(t) &= [\widehat{c}_1(t), \widehat{c}_2(t), \dots, \widehat{c}_{n_c}(t)]^T, \\ \widehat{\boldsymbol{\theta}}_n(t) &= [\widehat{d}_1(t), \widehat{d}_2(t), \dots, \widehat{d}_{n_d}(t)]^T.\end{aligned}\quad (19)$$

3. The F-MRLS Algorithm

The convergence rate of the MRLS algorithm in Section 2 is slow because the noise information intermediate variables $\mathbf{w}(t)$ contain unmeasurable time-delay noise $\mathbf{v}(t-i)$. The solution here is to present a filtering based recursive least squares algorithm (the F-MRLS algorithm) for the multivariable Hammerstein models by filtering the rational function $N(z)$ and transferring the FIR-MA model in (1) into a controlled autoregressive (CAR) model. Multiplying both sides of (1) by $N^{-1}(z)$ yields

$$N^{-1}(z) \mathbf{y}(t) = \mathbf{B}(z) N^{-1}(z) \bar{\mathbf{u}}(t) + \mathbf{v}(t), \quad (20)$$

or

$$\mathbf{y}_f(t) = \mathbf{B}(z) \bar{\mathbf{u}}_f(t) + \mathbf{v}(t), \quad (21)$$

where

$$\begin{aligned}\bar{\mathbf{u}}_f(t) &:= \frac{1}{N(z)} \bar{\mathbf{u}}(t) \\ &= \frac{1}{N(z)} \sum_{i=1}^{n_c} \widehat{c}_i(t) \boldsymbol{\gamma}_i(\mathbf{u}(t)) \\ &= c_1 \mathbf{U}_1(t) + c_2 \mathbf{U}_2(t) + \dots + c_{n_c} \mathbf{U}_{n_c}(t), \\ \mathbf{y}_f(t) &:= \frac{1}{N(z)} \mathbf{y}(t) \\ &= [1 - N(z)] \mathbf{y}_f(t) + \mathbf{y}(t) \\ &= \mathbf{y}(t) - \sum_{i=1}^{n_d} \widehat{d}_i(t) \mathbf{y}_f(t-i),\end{aligned}\quad (22)$$

$$\mathbf{U}_i(t) := \frac{1}{N(z)} \boldsymbol{\gamma}_i(\mathbf{u}(t)) \in \mathbb{R}^r, \quad i = 1, 2, \dots, n_c,$$

$$\boldsymbol{\zeta}_j(t) := \begin{bmatrix} \mathbf{U}_j(t) \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^m, \quad j = 1, 2, \dots, n_c.$$

Thus, (21) can be rewritten as

$$\begin{aligned}\mathbf{y}_f(t) &= \left[\mathbf{I}_{m \times r} + \sum_{i=1}^{n_b} \mathbf{B}_i z^{-i} \right] \bar{\mathbf{u}}_f(t) + \mathbf{v}(t) \\ &= \begin{bmatrix} \bar{\mathbf{u}}_f(t) \\ \mathbf{0} \end{bmatrix} + \mathbf{B}_1 \bar{\mathbf{u}}_f(t-1) + \mathbf{B}_2 \bar{\mathbf{u}}_f(t-2) \\ &\quad + \dots + \mathbf{B}_{n_b} \bar{\mathbf{u}}_f(t-n_b) + \mathbf{v}(t), \\ &= c_1 \boldsymbol{\zeta}_1(t) + c_2 \boldsymbol{\zeta}_2(t) + \dots + c_{n_c} \boldsymbol{\zeta}_{n_c}(t) + \mathbf{B}_1 \bar{\mathbf{u}}_f(t-1) \\ &\quad + \mathbf{B}_2 \bar{\mathbf{u}}_f(t-2) + \dots + \mathbf{B}_{n_b} \bar{\mathbf{u}}_f(t-n_b) + \mathbf{v}(t).\end{aligned}\quad (23)$$

Define the filtered information matrices:

$$\boldsymbol{\Psi}_f(t) := [\boldsymbol{\varphi}_{f1}^T(t) \otimes \mathbf{I}_m, \boldsymbol{\varphi}_{f2}(t)] \in \mathbb{R}^{m \times (mr n_b + n_c)}, \quad (24)$$

$$\boldsymbol{\varphi}_{f1}(t) := [\bar{\mathbf{u}}_f^T(t-1), \bar{\mathbf{u}}_f^T(t-2), \dots, \bar{\mathbf{u}}_f^T(t-n_b)]^T \in \mathbb{R}^{(r n_b)}, \quad (25)$$

$$\boldsymbol{\varphi}_{f2}(t) := [\boldsymbol{\zeta}_1(t), \boldsymbol{\zeta}_2(t), \dots, \boldsymbol{\zeta}_{n_c}(t)] \in \mathbb{R}^{m \times n_c}. \quad (26)$$

Since the polynomial $N(z)$ is unknown and to be estimated, it is impossible to use $\bar{\mathbf{u}}_f(t)$ to construct $\boldsymbol{\varphi}_{f1}(t)$ in (25). Here, we adopt the principle of the MRLS algorithm in Section 2 and replace the unmeasurable variables and vectors with their estimates to derive the following algorithm.

By using the parameter estimates $\widehat{\boldsymbol{\theta}}_1(t)$ and $\widehat{\boldsymbol{\theta}}_n(t)$, the estimates of polynomials $\mathbf{B}(z)$ and $N(z)$ at time t can be constructed as

$$\widehat{\mathbf{B}}(t, z) := \mathbf{I}_{m \times r} + \sum_{i=1}^{n_b} \widehat{\mathbf{B}}_i(t) z^{-i},$$

$$\widehat{N}(t, z) := 1 + \widehat{d}_1(t) z^{-1} + \widehat{d}_2(t) z^{-2} + \dots + \widehat{d}_{n_d}(t) z^{-n_d}. \quad (27)$$

$\mathbf{w}(t)$ in model (1) can be rewritten as

$$\begin{aligned}\mathbf{w}(t) &= \mathbf{y}(t) - \mathbf{x}(t) \\ &= \mathbf{y}(t) - \left[\mathbf{I}_{m \times r} + \sum_{i=1}^{n_b} \mathbf{B}_i(t) z^{-i} \right] \bar{\mathbf{u}}(t) \\ &= \mathbf{y}(t) - \boldsymbol{\varphi}_s(t) \boldsymbol{\theta}_s.\end{aligned}\quad (28)$$

Let $\widehat{\mathbf{w}}(t)$ be the estimate of $\mathbf{w}(t)$. Replacing $\mathbf{w}(t)$, $\mathbf{y}(t)$, $\boldsymbol{\varphi}_s(t)$, $\boldsymbol{\varphi}_n(t)$, $\boldsymbol{\theta}_n$, and $\boldsymbol{\theta}_s$ with their estimates $\widehat{\mathbf{w}}(t)$, $\widehat{\mathbf{y}}(t)$, $\widehat{\boldsymbol{\varphi}}_s(t)$, $\widehat{\boldsymbol{\varphi}}_n(t)$, $\widehat{\boldsymbol{\theta}}_n(t)$, and $\widehat{\boldsymbol{\theta}}_s(t-1)$ leads to

$$\widehat{\mathbf{w}}(t) = \widehat{\mathbf{y}}(t) - \widehat{\boldsymbol{\varphi}}_s(t) \widehat{\boldsymbol{\theta}}_s(t-1). \quad (29)$$

Defining and minimizing the cost function

$$J(\boldsymbol{\theta}_n) := \sum_{j=1}^t [\mathbf{w}(j) - \boldsymbol{\varphi}_n(j) \boldsymbol{\theta}_n]^2 \quad (30)$$

and using the least squares search principle, we list the recursive least squares algorithm to compute $\widehat{\boldsymbol{\theta}}_n(t)$:

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_n(t) &= \widehat{\boldsymbol{\theta}}_n(t-1) + \mathbf{L}_n(t) [\widehat{\mathbf{w}}(t) - \widehat{\boldsymbol{\varphi}}_n(t) \widehat{\boldsymbol{\theta}}_n(t-1)], \\ \mathbf{L}_n(t) &= \mathbf{P}_n(t-1) \widehat{\boldsymbol{\varphi}}_n^T(t) [\mathbf{I}_m + \widehat{\boldsymbol{\varphi}}_n(t) \mathbf{P}_n(t-1) \widehat{\boldsymbol{\varphi}}_n^T(t)]^{-1}, \\ \mathbf{P}_n(t) &= [\mathbf{I}_{n_d} - \mathbf{L}_n(t) \widehat{\boldsymbol{\varphi}}_n(t)] \mathbf{P}_n(t-1), \quad \mathbf{P}_n(0) = p_0 \mathbf{I}.\end{aligned}\quad (31)$$

Let $\widehat{\mathbf{c}}(t) := [\widehat{c}_1(t), \widehat{c}_2(t), \dots, \widehat{c}_{n_c}(t)]^T \in \mathbb{R}^{n_c}$ be the estimate of \mathbf{c} at time t , filtering $\mathbf{y}(t)$ with $1/\widehat{N}(t, z)$ to obtain the estimate $\widehat{\mathbf{y}}_f(t)$:

$$\begin{aligned}\widehat{\mathbf{y}}_f(t) &= \frac{1}{\widehat{N}(t, z)} \mathbf{y}(t) \\ &= [1 - \widehat{N}(t, z)] \widehat{\mathbf{y}}_f(t) + \mathbf{y}(t) \\ &= - \sum_{i=1}^{n_d} \widehat{d}_i(t) \widehat{\mathbf{y}}_f(t-i) + \mathbf{y}(t).\end{aligned}\quad (32)$$

The estimate of $\mathbf{U}_j(t)$ can be computed by

$$\begin{aligned}\widehat{\mathbf{U}}_j(t) &= \frac{1}{\widehat{N}(t, z)} \boldsymbol{\gamma}_j(\mathbf{u}(t)) \\ &= [1 - \widehat{N}(t, z)] \widehat{\mathbf{U}}_j(t) + \boldsymbol{\gamma}_j(\mathbf{u}(t)) \\ &= - \sum_{i=1}^{n_d} \widehat{d}_i(t) \widehat{\mathbf{U}}_j(t-i) + \boldsymbol{\gamma}_j(\mathbf{u}(t)).\end{aligned}\quad (33)$$

Define the estimate of $\boldsymbol{\zeta}_j(t)$ by

$$\widehat{\boldsymbol{\zeta}}_j(t) := \begin{bmatrix} \widehat{\mathbf{U}}_j(t) \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^m, \quad j = 1, 2, \dots, n_c, \quad (34)$$

and construct the estimate of $\boldsymbol{\Psi}_f(t)$ with $\widehat{\boldsymbol{\varphi}}_{f1}(t)$ and $\widehat{\boldsymbol{\varphi}}_{f2}(t)$ as follows:

$$\begin{aligned}\widehat{\boldsymbol{\Psi}}_f(t) &= [\widehat{\boldsymbol{\varphi}}_{f1}^T(t) \otimes \mathbf{I}_m, \widehat{\boldsymbol{\varphi}}_{f2}^T(t)], \\ \widehat{\boldsymbol{\varphi}}_{f1}(t) &= [\widehat{\mathbf{u}}_f^T(t-1), \widehat{\mathbf{u}}_f^T(t-2), \dots, \widehat{\mathbf{u}}_f^T(t-n_b)]^T, \\ \widehat{\boldsymbol{\varphi}}_{f2}(t) &= [\widehat{\boldsymbol{\zeta}}_1(t), \widehat{\boldsymbol{\zeta}}_2(t), \dots, \widehat{\boldsymbol{\zeta}}_{n_c}(t)].\end{aligned}\quad (35)$$

The filtered model in (21) can be rewritten in a matrix form:

$$\mathbf{y}_f(t) = \boldsymbol{\Psi}_f(t) \boldsymbol{\theta}_s + \mathbf{v}(t), \quad (36)$$

or

$$\mathbf{v}(t) = \mathbf{y}_f(t) - \boldsymbol{\Psi}_f(t) \boldsymbol{\theta}_s. \quad (37)$$

Based on the MRLS search principle, we can obtain the estimate of $\boldsymbol{\theta}_s$ by the following algorithm:

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_s(t) &= \widehat{\boldsymbol{\theta}}_s(t-1) + \mathbf{L}_f(t) [\widehat{\mathbf{y}}_f(t) - \widehat{\boldsymbol{\Psi}}_f(t) \widehat{\boldsymbol{\theta}}_s(t-1)], \\ \mathbf{L}_f(t) &= \mathbf{P}_f(t-1) \widehat{\boldsymbol{\Psi}}_f^T(t) [\mathbf{I}_m + \widehat{\boldsymbol{\Psi}}_f(t) \mathbf{P}_f(t-1) \widehat{\boldsymbol{\Psi}}_f^T(t)]^{-1}, \\ \mathbf{P}_f(t) &= [\mathbf{I}_{r_{mb+n_c}} - \mathbf{L}_f(t) \widehat{\boldsymbol{\Psi}}_f(t)] \mathbf{P}_f(t-1), \\ \mathbf{P}_f(0) &= p_0 \mathbf{I}.\end{aligned}\quad (38)$$

The estimate $\widehat{\mathbf{u}}(t)$ can be computed by

$$\begin{aligned}\widehat{\mathbf{u}}(t) &= \widehat{c}_1(t) \boldsymbol{\gamma}_1(\mathbf{u}(t)) + \widehat{c}_2(t) \boldsymbol{\gamma}_2(\mathbf{u}(t)) \\ &\quad + \dots + \widehat{c}_{n_c}(t) \boldsymbol{\gamma}_{n_c}(\mathbf{u}(t)).\end{aligned}\quad (39)$$

Filter $\widehat{\mathbf{u}}(t)$ by $1/\widehat{N}(t, z)$ to obtain the estimate $\widehat{\mathbf{u}}_f(t)$:

$$\begin{aligned}\widehat{\mathbf{u}}_f(t) &= \frac{1}{\widehat{N}(t, z)} \widehat{\mathbf{u}}(t) \\ &= \frac{1}{\widehat{N}(t, z)} \sum_{i=1}^{n_c} \widehat{c}_i(t) \boldsymbol{\gamma}_i(\mathbf{u}(t)) \\ &= \widehat{c}_1(t) \widehat{\mathbf{U}}_1(t) + \widehat{c}_2(t) \widehat{\mathbf{U}}_2(t) + \dots + \widehat{c}_{n_c}(t) \widehat{\mathbf{U}}_{n_c}(t).\end{aligned}\quad (40)$$

Replacing $\mathbf{v}(t)$, $\mathbf{y}_f(t)$, $\boldsymbol{\Psi}_f(t)$, and $\boldsymbol{\theta}_s$ in (37) with their estimates $\widehat{\mathbf{v}}(t)$, $\widehat{\mathbf{y}}_f(t)$, $\widehat{\boldsymbol{\Psi}}_f(t)$, and $\widehat{\boldsymbol{\theta}}_s(t)$ at time t , the noise vector can be computed by

$$\widehat{\mathbf{v}}(t) = \widehat{\mathbf{y}}_f(t) - \widehat{\boldsymbol{\Psi}}_f(t) \widehat{\boldsymbol{\theta}}_s(t). \quad (41)$$

In conclusion, we can summarize the following filtering based recursive least squares algorithm for multivariable Hammerstein models (the F-MRLS algorithm for short):

$$\widehat{\boldsymbol{\theta}}_s(t) = \widehat{\boldsymbol{\theta}}_s(t-1) + \mathbf{L}_f(t) [\widehat{\mathbf{y}}_f(t) - \widehat{\boldsymbol{\Psi}}_f(t) \widehat{\boldsymbol{\theta}}_s(t-1)], \quad (42)$$

$$\mathbf{L}_f(t) = \frac{\mathbf{P}_f(t-1) \widehat{\boldsymbol{\Psi}}_f^T(t)}{\mathbf{I}_m + \widehat{\boldsymbol{\Psi}}_f(t) \mathbf{P}_f(t-1) \widehat{\boldsymbol{\Psi}}_f^T(t)}, \quad (43)$$

$$\mathbf{P}_f(t) = [\mathbf{I}_{r_{mb+n_c}} - \mathbf{L}_f(t) \widehat{\boldsymbol{\varphi}}_f(t)] \mathbf{P}_f(t-1), \quad (44)$$

$$\widehat{\boldsymbol{\Psi}}_f(t) = [\widehat{\boldsymbol{\varphi}}_{f1}^T(t) \otimes \mathbf{I}_m, \widehat{\boldsymbol{\varphi}}_{f2}^T(t)], \quad (45)$$

$$\widehat{\boldsymbol{\varphi}}_{f1}(t) = [\widehat{\mathbf{u}}_f^T(t-1), \widehat{\mathbf{u}}_f^T(t-2), \dots, \widehat{\mathbf{u}}_f^T(t-n_b)]^T, \quad (46)$$

$$\widehat{\boldsymbol{\varphi}}_{f2}(t) = [\widehat{\boldsymbol{\zeta}}_1(t), \widehat{\boldsymbol{\zeta}}_2(t), \dots, \widehat{\boldsymbol{\zeta}}_{n_c}(t)], \quad (47)$$

$$\widehat{\boldsymbol{\zeta}}_j(t) = \begin{bmatrix} \widehat{\mathbf{U}}_j(t) \\ \mathbf{0} \end{bmatrix}, \quad (48)$$

$$\widehat{\mathbf{U}}_j(t) = -\sum_{i=1}^{n_d} \widehat{d}_i(t) \widehat{\mathbf{U}}_j(t-i) + \boldsymbol{\gamma}_j(\mathbf{u}(t)), \quad (49)$$

$$\widehat{\mathbf{u}}_f(t) = \widehat{c}_1(t) \widehat{\mathbf{U}}_1(t) + \widehat{c}_2(t) \widehat{\mathbf{U}}_2(t) + \cdots + \widehat{c}_{n_c}(t) \widehat{\mathbf{U}}_{n_c}(t), \quad (50)$$

$$\widehat{\mathbf{y}}_f(t) = -\sum_{i=1}^{n_d} \widehat{d}_i(t) \widehat{\mathbf{y}}_f(t-i) + \mathbf{y}(t), \quad (51)$$

$$\widehat{\boldsymbol{\theta}}_n(t) = \widehat{\boldsymbol{\theta}}_n(t-1) + \mathbf{L}_n(t) [\widehat{\mathbf{w}}(t) - \widehat{\boldsymbol{\varphi}}_n(t) \widehat{\boldsymbol{\theta}}_n(t-1)], \quad (52)$$

$$\mathbf{L}_n(t) = \mathbf{P}_n(t-1) \widehat{\boldsymbol{\varphi}}_n^T(t) [\mathbf{I}_m + \widehat{\boldsymbol{\varphi}}_n(t) \mathbf{P}_n(t-1) \widehat{\boldsymbol{\varphi}}_n^T(t)]^{-1}, \quad (53)$$

$$\mathbf{P}_n(t) = [\mathbf{I}_{n_d} - \mathbf{L}_n(t) \widehat{\boldsymbol{\varphi}}_n(t)] \mathbf{P}_n(t-1), \quad (54)$$

$$\widehat{\mathbf{w}}(t) = \mathbf{y}(t) - \widehat{\boldsymbol{\varphi}}_s(t) \widehat{\boldsymbol{\theta}}_s(t-1), \quad (55)$$

$$\widehat{\boldsymbol{\Psi}}(t) = [\widehat{\boldsymbol{\varphi}}_s(t), \widehat{\boldsymbol{\varphi}}_n(t)], \quad (56)$$

$$\widehat{\boldsymbol{\varphi}}_s(t) = [\widehat{\boldsymbol{\varphi}}_1^T(t) \otimes \mathbf{I}_m, \boldsymbol{\varphi}_2(t)], \quad (57)$$

$$\widehat{\boldsymbol{\varphi}}_1(t) = [\widehat{\mathbf{u}}^T(t-1), \widehat{\mathbf{u}}^T(t-2), \dots, \widehat{\mathbf{u}}^T(t-n_b)]^T, \quad (58)$$

$$\boldsymbol{\varphi}_2(t) = [\boldsymbol{\eta}_1(t), \boldsymbol{\eta}_2(t), \dots, \boldsymbol{\eta}_{n_c}(t)], \quad (59)$$

$$\boldsymbol{\eta}_i(t) = \begin{bmatrix} \boldsymbol{\gamma}_i(\mathbf{u}(t)) \\ \mathbf{0} \end{bmatrix}, \quad (60)$$

$$\widehat{\mathbf{v}}(t) = [\widehat{\mathbf{v}}(t-1), \widehat{\mathbf{v}}(t-2), \dots, \widehat{\mathbf{v}}(t-n_d)], \quad (61)$$

$$\widehat{\mathbf{v}}(t) = \widehat{\mathbf{y}}_f(t) - \widehat{\boldsymbol{\Psi}}_f(t) \widehat{\boldsymbol{\theta}}_s(t), \quad (62)$$

$$\widehat{\mathbf{u}}(t) = \sum_{i=1}^{n_c} \widehat{c}_i(t) \boldsymbol{\gamma}_i(\mathbf{u}(t)), \quad (63)$$

$$\widehat{\boldsymbol{\theta}}(t) = \begin{bmatrix} \widehat{\boldsymbol{\theta}}_s(t) \\ \widehat{\boldsymbol{\theta}}_n(t) \end{bmatrix}, \quad (64)$$

$$\widehat{\boldsymbol{\theta}}_s(t) = \begin{bmatrix} \text{col}[\widehat{\boldsymbol{\theta}}_1^T(t)] \\ \widehat{\boldsymbol{\theta}}_2(t) \end{bmatrix}, \quad (65)$$

$$\widehat{\boldsymbol{\theta}}_1^T(t) = [\widehat{\mathbf{B}}_1(t), \widehat{\mathbf{B}}_2(t), \dots, \widehat{\mathbf{B}}_{n_b}(t)], \quad (66)$$

$$\widehat{\boldsymbol{\theta}}_2(t) = [\widehat{c}_1(t), \widehat{c}_2(t), \dots, \widehat{c}_{n_c}(t)]^T, \quad (67)$$

$$\widehat{\boldsymbol{\theta}}_n(t) = [\widehat{d}_1(t), \widehat{d}_2(t), \dots, \widehat{d}_{n_d}(t)]^T. \quad (68)$$

The steps involved in the F-MRLS algorithm for multi-variable Hammerstein systems are listed in the following.

- (1) To initialize, let $t = 1$, set the initial values of the parameter estimation variables and covariance matrices as follows: $\widehat{\boldsymbol{\theta}}_s(i) = \mathbf{1}_{m+n_b+n_c}/p_0$, $\widehat{\boldsymbol{\theta}}_n(i) = \mathbf{1}_{n_d}/p_0$, $\widehat{\mathbf{y}}_f(i) = \mathbf{1}_m/p_0$, $\widehat{\mathbf{u}}_f(i) = \mathbf{1}_r/p_0$, $\widehat{\mathbf{u}}(i) = \mathbf{1}_r/p_0$, $\widehat{\mathbf{v}}(i) = \mathbf{1}_m/p_0$ for $i \leq 0$, $\widehat{\mathbf{U}}_j(i) = \mathbf{1}_r/p_0$ for $i \leq 0$

and $j = 1, 2, \dots, n_c$, $\mathbf{P}_f(0) = p_0 \mathbf{I}_{m+n_b+n_c}$, $\mathbf{P}_n(0) = p_0 \mathbf{I}_{n_d}$, $p_0 = 10^6$, and give the basis functions $\boldsymbol{\gamma}_j(\cdot)$.

- (2) Collect the input-output data $\mathbf{u}(t)$ and $\mathbf{y}(t)$, and construct $\boldsymbol{\eta}_i(t)$ by (60), and $\widehat{\boldsymbol{\varphi}}_1(t)$, $\boldsymbol{\varphi}_2(t)$ by (58), (59). Form the information vectors $\widehat{\boldsymbol{\varphi}}_s(t)$ by (57), $\widehat{\boldsymbol{\varphi}}_n(t)$ by (61), and $\widehat{\boldsymbol{\Psi}}(t)$ by (56), respectively.
- (3) Compute $\widehat{\mathbf{w}}(t)$ by (55), the gain vector $\mathbf{L}_n(t)$ by (53), and the covariance matrix $\mathbf{P}_n(t)$ by (54), respectively. Update the parameter estimate $\widehat{\boldsymbol{\theta}}_n(t)$ by (52).
- (4) Compute $\widehat{\mathbf{y}}_f(t)$ by (51), $\widehat{\mathbf{U}}_j(t)$ by (49). Construct $\widehat{\boldsymbol{\zeta}}_j(t)$, $\widehat{\boldsymbol{\varphi}}_{f2}(t)$, and $\widehat{\boldsymbol{\varphi}}_{f1}(t)$ by (48), (47), and (46), respectively. Compute $\widehat{\boldsymbol{\Psi}}_f(t)$ by (45).
- (5) Compute the gain vector $\mathbf{L}_f(t)$ by (43) and the covariance matrix $\mathbf{P}_f(t)$ by (44). Update the parameter estimate $\widehat{\boldsymbol{\theta}}_s(t)$ by (42).
- (6) Construct the parameter vectors $\widehat{\boldsymbol{\theta}}_1(t)$, $\widehat{\boldsymbol{\theta}}_2(t)$, and $\widehat{\boldsymbol{\theta}}_n(t)$ by (66), (67), and (68). Form $\widehat{\boldsymbol{\theta}}_s(t)$ and $\widehat{\boldsymbol{\theta}}(t)$ by (65) and (64). Compute $\widehat{\mathbf{u}}_f(t)$ by (50), $\widehat{\mathbf{v}}(t)$ by (62), and $\widehat{\mathbf{u}}(t)$ by (63).
- (7) Increase t by 1 and go to step 2.

4. Examples

Example 1. Consider the following 2-input 2-output Hammerstein FIR-MA system:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} \bar{u}_1(t) \\ \bar{u}_2(t) \end{bmatrix} + \mathbf{B} \begin{bmatrix} \bar{u}_1(t-1) \\ \bar{u}_2(t-1) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} + d_1 \begin{bmatrix} v_1(t-1) \\ v_2(t-1) \end{bmatrix}, \quad (69)$$

where

$$\begin{aligned} \mathbf{B}(z) &= \begin{bmatrix} 0.13 & 0.25 \\ -1.21 & 0.13 \end{bmatrix} z^{-1}, \\ D(z) &= 1 + d_1 z^{-1} = 1 + 0.68 z^{-1}, \\ \bar{\mathbf{u}}(t) &= c_1 \mathbf{u}(t) + c_2 \mathbf{u}^2(t), \\ \mathbf{c} &= [c_1, c_2]^T = [1.00, -0.14]^T, \\ \boldsymbol{\vartheta} &= [0.13, -1.21, 0.25, 0.13, -0.14, 0.68]^T. \end{aligned} \quad (70)$$

In simulation, the inputs $\{u_1(t)\}$ and $\{u_2(t)\}$ are taken as two uncorrelated persistent excitation signal sequences with zero mean and unit variances and $\{v_1(t)\}$ and $\{v_2(t)\}$ as two white noise sequences with zero mean and variances $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ($\sigma^2 = 0.50^2, \sigma^2 = 1.00^2$). Applying the MRLS algorithm in (19) and the F-MRLS algorithm in (42)–(68) to estimate the parameters of this multivariable Hammerstein system, the F-MRLS parameter estimates and their estimation errors are shown in Table 1, the comparison between the F-MRLS algorithm and the MRLS algorithm in the estimation error $\delta := \|\widehat{\boldsymbol{\theta}}(t) - \boldsymbol{\vartheta}\|/\|\boldsymbol{\vartheta}\|$ versus t is shown in Table 2, and the

TABLE 1: The F-MRLS estimates and errors in Example 1 ($\sigma^2 = 0.50^2$ and $\sigma^2 = 1.00^2$).

σ^2	t	$B(1, 1)$	$B(2, 1)$	$B(1, 2)$	$B(2, 2)$	c_2	d_1	δ (%)
0.50^2	100	0.10919	-1.60279	0.39771	0.12515	-0.16795	0.46319	33.14292
	200	0.08155	-1.44298	0.31145	0.09590	-0.13875	0.51078	21.01533
	500	0.10144	-1.33934	0.25923	0.08705	-0.15442	0.60335	11.18663
	1000	0.12452	-1.30385	0.26703	0.11599	-0.14796	0.65282	7.04142
	2000	0.13930	-1.27241	0.26711	0.12703	-0.14550	0.69136	4.66378
	3000	0.14008	-1.25781	0.26567	0.11928	-0.14361	0.70371	4.03373
1.00^2	100	0.09530	-1.72159	0.52726	0.13387	-0.15869	0.52551	42.21841
	200	0.03631	-1.52223	0.36553	0.07113	-0.11792	0.58022	25.56872
	500	0.07400	-1.40400	0.25492	0.04715	-0.16037	0.64763	15.50868
	1000	0.12015	-1.36311	0.27803	0.10606	-0.15335	0.68001	11.08080
	2000	0.14849	-1.31457	0.28056	0.12704	-0.15048	0.70624	7.98334
	3000	0.15005	-1.29138	0.27909	0.10998	-0.14680	0.71335	6.79480
True values		0.13000	-1.21000	0.25000	0.13000	-0.14000	0.68000	

 TABLE 2: The MRLS and F-MRLS estimates and errors in Example 1 ($\sigma^2 = 0.50^2$).

Algorithms	t	$B(1, 1)$	$B(2, 1)$	$B(1, 2)$	$B(2, 2)$	c_2	d_1	δ (%)
MRLS	100	-0.01946	-1.84785	0.44644	0.04712	-0.16528	0.39802	52.11950
	200	-0.03340	-1.60645	0.31299	0.05091	-0.14456	0.46017	34.45327
	500	0.03788	-1.39775	0.28600	0.03749	-0.14940	0.56178	18.20046
	1000	0.08074	-1.34458	0.28178	0.07949	-0.14078	0.61722	11.71737
	2000	0.11167	-1.28751	0.27285	0.10106	-0.13785	0.66411	6.24292
	3000	0.12600	-1.27455	0.26866	0.10656	-0.13729	0.68367	4.99772
F-MRLS	100	0.10919	-1.60279	0.39771	0.12515	-0.16795	0.46319	33.14292
	200	0.08155	-1.44298	0.31145	0.09590	-0.13875	0.51078	21.01533
	500	0.10144	-1.33934	0.25923	0.08705	-0.15442	0.60335	11.18663
	1000	0.12452	-1.30385	0.26703	0.11599	-0.14796	0.65282	7.04142
	2000	0.13930	-1.27241	0.26711	0.12703	-0.14550	0.69136	4.66378
	3000	0.14008	-1.25781	0.26567	0.11928	-0.14361	0.70371	4.03373
True values		0.13000	-1.21000	0.25000	0.13000	-0.14000	0.68000	

estimation errors $\delta := \|\hat{\boldsymbol{\theta}}(t) - \boldsymbol{\theta}\|/\|\boldsymbol{\theta}\|$ versus t are shown in Figures 1 and 2.

Example 2. Consider the following 2-input 2-output Hammerstein FIR-MA system:

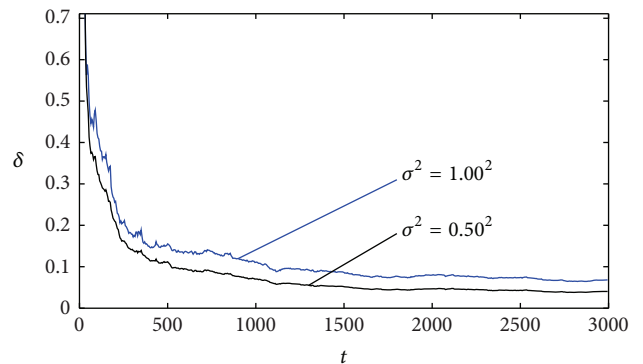
$$\begin{aligned} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} &= \begin{bmatrix} \bar{u}_1(t) \\ \bar{u}_2(t) \end{bmatrix} + \mathbf{B} \begin{bmatrix} \bar{u}_1(t-1) \\ \bar{u}_2(t-1) \end{bmatrix} \\ &+ \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} + d_1 \begin{bmatrix} v_1(t-1) \\ v_2(t-1) \end{bmatrix}, \end{aligned} \quad (71)$$

where

$$\mathbf{B}(z) = \begin{bmatrix} 0.14 & 0.25 \\ -0.15 & 0.125 \end{bmatrix} z^{-1},$$

$$D(z) = 1 + d_1 z^{-1} = 1 + 0.50z^{-1},$$

$$\bar{\mathbf{u}}(t) = c_1 \mathbf{u}(t) + c_2 \mathbf{u}^2(t) + c_3 \mathbf{u}^3(t),$$


 FIGURE 1: The F-RLS estimation errors δ versus t in Example 1 ($\sigma^2 = 0.50^2$ and $\sigma^2 = 1.00^2$).

$$\mathbf{c} = [c_1, c_2, c_3]^T = [1.00, -0.19, 1.19]^T,$$

$$\boldsymbol{\theta} = [0.14, -0.15, 0.25, 0.125, -0.19, 1.19, 0.50]^T. \quad (72)$$

TABLE 3: The F-RLS estimates and errors in Example 2 ($\sigma^2 = 0.50^2$ and $\sigma^2 = 1.00^2$).

σ^2	t	$B(1, 1)$	$B(2, 1)$	$B(1, 2)$	$B(2, 2)$	c_2	c_3	d_1	δ (%)
0.50 ²	100	0.12373	-1.02307	0.27283	0.10358	-0.39300	1.25093	0.32512	16.33820
	200	0.12306	-1.03981	0.24977	0.11748	-0.29306	1.19925	0.35407	10.57344
	500	0.13109	-1.05853	0.24655	0.11563	-0.25013	1.19098	0.40948	6.44687
	1000	0.13728	-1.06400	0.25207	0.12027	-0.22484	1.21159	0.45612	3.63495
	2000	0.14372	-1.06203	0.25315	0.12346	-0.20822	1.20110	0.49970	1.46876
	3000	0.14389	-1.05999	0.25321	0.12156	-0.20281	1.19897	0.51458	1.43078
1.00 ²	100	0.10227	-0.88895	0.29620	0.02428	-0.45977	1.39972	0.43173	23.55652
	200	0.10022	-0.94211	0.26126	0.06755	-0.30135	1.31216	0.41341	13.31668
	500	0.11893	-1.01753	0.24656	0.08815	-0.27093	1.24343	0.43981	7.40054
	1000	0.13195	-1.05047	0.25483	0.10514	-0.24217	1.26386	0.47397	5.67199
	2000	0.14551	-1.05983	0.25634	0.11672	-0.21751	1.22939	0.51038	3.02299
	3000	0.14632	-1.06041	0.25650	0.11459	-0.20918	1.22061	0.52227	2.68962
True values		0.14000	-1.05000	0.25000	0.12500	-0.19000	1.19000	0.50000	

TABLE 4: The comparison of parameter estimates and errors in Example 2 ($\sigma^2 = 0.50^2$).

Algorithms	t	$B(1, 1)$	$B(2, 1)$	$B(1, 2)$	$B(2, 2)$	c_2	c_3	d_1	δ (%)
MRLS	100	0.10379	-1.49022	0.30877	0.11454	-0.37596	0.96745	0.25108	34.46417
	200	0.08551	-1.31358	0.29070	0.10159	-0.30829	1.05130	0.24601	24.37082
	500	0.10950	-1.17837	0.27162	0.09651	-0.24919	1.13023	0.30548	14.80320
	1000	0.12218	-1.13050	0.26705	0.10410	-0.21956	1.17878	0.36159	9.76634
	2000	0.13294	-1.09517	0.26234	0.11047	-0.20567	1.18593	0.42481	5.36992
	3000	0.13715	-1.08439	0.25990	0.11297	-0.20141	1.18517	0.45571	3.49648
F-MRLS	100	0.12373	-1.02307	0.27283	0.10358	-0.39300	1.25093	0.32512	16.33820
	200	0.12306	-1.03981	0.24977	0.11748	-0.29306	1.19925	0.35407	10.57344
	500	0.13109	-1.05853	0.24655	0.11563	-0.25013	1.19098	0.40948	6.44687
	1000	0.13728	-1.06400	0.25207	0.12027	-0.22484	1.21159	0.45612	3.63495
	2000	0.14372	-1.06203	0.25315	0.12346	-0.20822	1.20110	0.49970	1.46876
	3000	0.14389	-1.05999	0.25321	0.12156	-0.20281	1.19897	0.51458	1.43078
True values		0.14000	-1.05000	0.25000	0.12500	-0.19000	1.19000	0.50000	

TABLE 5: Comparison of the computational efficiency of the F-MRLS and MRLS algorithms.

Algorithms	Number of multiplications	Number of additions
MRLS	$2n^2 + 5n$	$2n^2 + 2n$
F-MRLS	$2n_b^2 + 2n_c^2 + 2n_d^2 + 5n$	$2n_b^2 + 2n_c^2 + 2n_d^2 + 2n + 1$

In simulation, the inputs $\{u_1(t)\}$ and $\{u_2(t)\}$ and noise data $\{v_1(t)\}$ and $\{v_2(t)\}$ are settled in the same way as in Example 1, where variances $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ($\sigma^2 = 0.50^2$). Applying the MRLS algorithm and the F-MRLS algorithm in (42)–(68), the F-MRLS parameter estimates and their estimation errors are shown in Table 3, the comparison between the F-MRLS algorithm and the MRLS algorithm in the estimation error $\delta := \|\hat{\Theta}(t) - \Theta\|/\|\Theta\|$ versus t is shown in Table 4, and the estimation errors $\delta := \|\hat{\Theta}(t) - \Theta\|/\|\Theta\|$ versus t are shown in Figures 3 and 4.

To illustrate the advantages of the proposed algorithm, the numbers of multiplications and additions for each step of the F-MRLS algorithm and the MRLS algorithm are listed in Table 5.

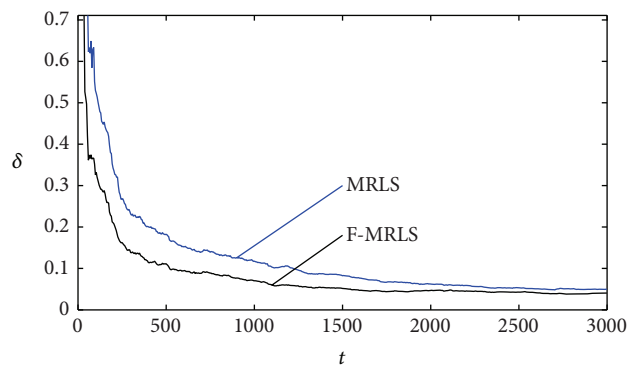


FIGURE 2: The parameter estimation errors δ versus t in Example 1 ($\sigma^2 = 0.50^2$).

From Tables 1–5, Figures 1–4, we can draw the following conclusions.

- (i) The parameter estimation errors are getting smaller with t increasing, which proves that the proposed algorithms are effective.

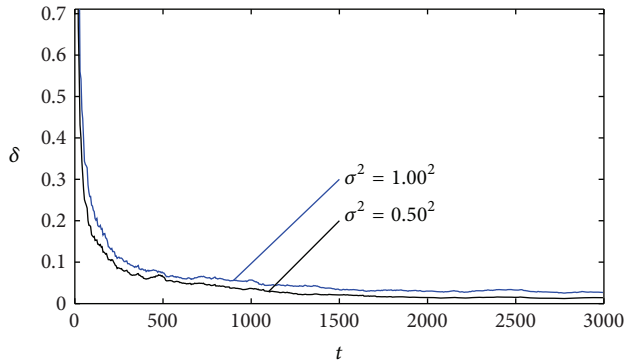


FIGURE 3: The F-RLS estimation errors δ versus t in Example 2 ($\sigma^2 = 0.50^2$ and $\sigma^2 = 1.00^2$).

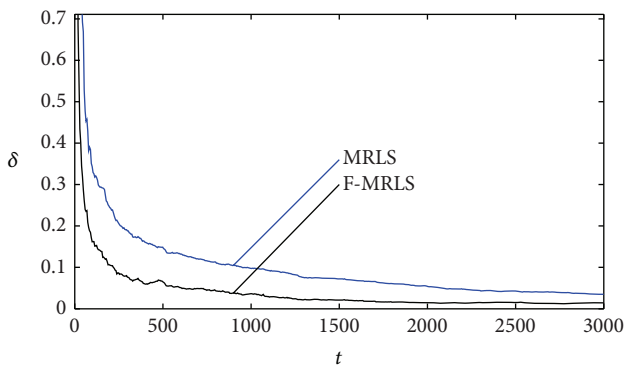


FIGURE 4: The parameter estimation errors δ versus t in Example 2 ($\sigma^2 = 0.50^2$).

- (ii) The F-MRLS algorithm is more accurate than the MRLS algorithm, which means the proposed F-MRLS algorithm has a better performance compared with the MRLS algorithm.
- (iii) The parameter estimates given by the F-MRLS algorithm have faster convergence than those given by the MRLS algorithm.

5. Conclusions

This paper presents a data filtering based recursive least squares algorithm for MIMO nonlinear FIR-MA systems. The simulation results show that the proposed data filtering based recursive least squares algorithm is more accurate and reduces computational burden compared with the recursive least squares algorithm.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (nos. 61174032 and 61202473),

the Doctoral Foundation of Higher Education Priority Fields of Scientific Research (no. 20110093130001), and the Scientific Research Foundation of Jiangnan University (no. 1252050205135110) and by the PAPD of Jiangsu Higher Education Institutions.

References

- [1] W.-H. Ho, J.-H. Chou, and C.-Y. Guo, "Parameter identification of chaotic systems using improved differential evolution algorithm," *Nonlinear Dynamics*, vol. 61, no. 1-2, pp. 29-41, 2010.
- [2] M. D. Narayanan, S. Narayanan, and C. Padmanabhan, "Parametric identification of nonlinear systems using multiple trials," *Nonlinear Dynamics*, vol. 48, no. 4, pp. 341-360, 2007.
- [3] W. Silva, "Identification of nonlinear aeroelastic systems based on the volterra theory: progress and opportunities," *Nonlinear Dynamics*, vol. 39, no. 1-2, pp. 25-62, 2005.
- [4] Y. Zhang and G. Cui, "Bias compensation methods for stochastic systems with colored noise," *Applied Mathematical Modelling: Simulation and Computation for Engineering and Environmental Systems*, vol. 35, no. 4, pp. 1709-1716, 2011.
- [5] Y. Zhang, "Unbiased identification of a class of multi-input single-output systems with correlated disturbances using bias compensation methods," *Mathematical and Computer Modelling*, vol. 53, no. 9-10, pp. 1810-1819, 2011.
- [6] Y. S. Xiao, Y. Zhang, J. Ding, and J. Y. Dai, "The residual based interactive least squares algorithms and simulation studies," *Computers & Mathematics with Applications*, vol. 58, no. 6, pp. 1190-1197, 2009.
- [7] Q.-G. Wang, X. Guo, and Y. Zhang, "Direct identification of continuous time delay systems from step responses," *Journal of Process Control*, vol. 11, no. 5, pp. 531-542, 2001.
- [8] F. Ding, "State filtering and parameter estimation for state space systems with scarce measurements," *Signal Processing*, vol. 104, pp. 369-380, 2014.
- [9] Y. Gu, F. Ding, and J. H. Li, "State filtering and parameter estimation for linear systems with d-step state-delay," *IET Signal Processing*, vol. 8, no. 6, pp. 639-646, 2014.
- [10] M. Dehghan and M. Hajarian, "Two algorithms for finding the HERmitian reflexive and skew-HERmitian solutions of Sylvester matrix equations," *Applied Mathematics Letters*, vol. 24, no. 4, pp. 444-449, 2011.
- [11] M. Dehghan and M. Hajarian, "A lower bound for the product of eigenvalues of solutions to matrix equations," *Applied Mathematics Letters*, vol. 22, no. 12, pp. 1786-1788, 2009.
- [12] M. Dehghan and M. Hajarian, "SSHI methods for solving general linear matrix equations," *Engineering Computations (Swansea, Wales)*, vol. 28, no. 8, pp. 1028-1043, 2011.
- [13] M. Dehghan and M. Hajarian, "An iterative algorithm for the reflexive solutions of the generalized coupled Sylvester matrix equations and its optimal approximation," *Applied Mathematics and Computation*, vol. 202, no. 2, pp. 571-588, 2008.
- [14] M. Dehghan and M. Hajarian, "Finite iterative algorithms for the reflexive and anti-reflexive solutions of the matrix equation $A_1 X_1 B_1 + A_2 X_2 B_2 = C$," *Mathematical and Computer Modelling*, vol. 49, no. 9-10, pp. 1937-1959, 2009.
- [15] M. Dehghan and M. Hajarian, "Fourth-order variants of Newton's method without second derivatives for solving non-linear equations," *Engineering Computations (Swansea, Wales)*, vol. 29, no. 4, pp. 356-365, 2012.

- [16] M. Dehghan and M. Hajarian, "Iterative algorithms for the generalized centro-symmetric and central anti-symmetric solutions of general coupled matrix equations," *Engineering Computations*, vol. 29, no. 5, pp. 528–560, 2012.
- [17] Y. Shi and H. Fang, "Kalman filter-based identification for systems with randomly missing measurements in a network environment," *International Journal of Control*, vol. 83, no. 3, pp. 538–551, 2010.
- [18] Y. Shi and B. Yu, "Output feedback stabilization of networked control systems with random delays modeled by Markov chains," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1668–1674, 2009.
- [19] Y. Shi and B. Yu, "Robust mixed H_2/H_∞ control of networked control systems with random time delays in both forward and backward communication links," *Automatica*, vol. 47, no. 4, pp. 754–760, 2011.
- [20] Y. Hu, "Iterative and recursive least squares estimation algorithms for moving average systems," *Simulation Modelling Practice and Theory*, vol. 34, pp. 12–19, 2013.
- [21] Y. B. Hu, B. L. Liu, Q. Zhou, and C. Yang, "Recursive extended least squares parameter estimation for Wiener nonlinear systems with moving average noises," *Circuits, Systems, and Signal Processing*, vol. 33, no. 2, pp. 655–664, 2014.
- [22] C. Wang and T. Tang, "Recursive least squares estimation algorithm applied to a class of linear-in-parameters output error moving average systems," *Applied Mathematics Letters*, vol. 29, pp. 36–41, 2014.
- [23] F. Ding, X. Liu, and J. Chu, "Gradient-based and least-squares-based iterative algorithms for Hammerstein systems using the hierarchical identification principle," *IET Control Theory & Applications*, vol. 7, no. 2, pp. 176–184, 2013.
- [24] C. Wang and T. Tang, "Several gradient-based iterative estimation algorithms for a class of nonlinear systems using the filtering technique," *Nonlinear Dynamics*, vol. 77, no. 3, pp. 769–780, 2014.
- [25] Y. B. Hu, B. L. Liu, and Q. Zhou, "A multi-innovation generalized extended stochastic gradient algorithm for output nonlinear autoregressive moving average systems," *Applied Mathematics and Computation*, vol. 247, pp. 218–224, 2014.
- [26] J. H. Li, "Parameter estimation for Hammerstein CARARMA systems based on the Newton iteration," *Applied Mathematics Letters*, vol. 26, no. 1, pp. 91–96, 2013.
- [27] F. Ding and T. Chen, "Identification of Hammerstein nonlinear ARMAX systems," *Automatica*, vol. 41, no. 9, pp. 1479–1489, 2005.
- [28] F. Ding, Y. Shi, and T. Chen, "Auxiliary model-based least-squares identification methods for Hammerstein output-error systems," *Systems and Control Letters*, vol. 56, no. 5, pp. 373–380, 2007.
- [29] F. Ding, X. P. Liu, and G. Liu, "Identification methods for Hammerstein nonlinear systems," *Digital Signal Processing: A Review Journal*, vol. 21, no. 2, pp. 215–238, 2011.
- [30] Z. Y. Wang, Y. X. Shen, Z. Ji, and R. Ding, "Filtering based recursive least squares algorithm for Hammerstein FIR-MA systems," *Nonlinear Dynamics*, vol. 73, no. 1-2, pp. 1045–1054, 2013.
- [31] M. S. Ahmad, O. Kukrer, and A. Hocanin, "Recursive inverse adaptive filtering algorithm," *Digital Signal Processing: A Review Journal*, vol. 21, no. 4, pp. 491–496, 2011.
- [32] D. Q. Wang, "Least squares-based recursive and iterative estimation for output error moving average systems using data filtering," *IET Control Theory & Applications*, vol. 5, no. 14, pp. 1648–1657, 2011.
- [33] B. Yu, Y. Shi, and H. Huang, " l_2 - l_∞ filtering for multirate systems based on lifted models," *Circuits, Systems, and Signal Processing*, vol. 27, no. 5, pp. 699–711, 2008.
- [34] Y. Xiao and N. Yue, "Parameter estimation for nonlinear dynamical adjustment models," *Mathematical and Computer Modelling*, vol. 54, no. 5-6, pp. 1561–1568, 2011.
- [35] F. Ding and H. Duan, "Two-stage parameter estimation algorithms for Box-Jenkins systems," *IET Signal Processing*, vol. 7, no. 8, pp. 646–654, 2013.
- [36] J. Vörös, "Iterative algorithm for parameter identification of Hammerstein systems with two-segment nonlinearities," *IEEE Transactions on Automatic Control*, vol. 44, no. 11, pp. 2145–2149, 1999.
- [37] J. Vörös, "Identification of Hammerstein systems with time-varying piecewise-linear characteristics," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 52, no. 12, pp. 865–869, 2005.
- [38] J. Vörös, "Recursive identification of Hammerstein systems with discontinuous nonlinearities containing dead-zones," *IEEE Transactions on Automatic Control*, vol. 48, no. 12, pp. 2203–2206, 2003.
- [39] Y. Liu, F. Ding, and Y. Shi, "An efficient hierarchical identification method for general dual-rate sampled-data systems," *Automatica*, vol. 50, no. 3, pp. 962–970, 2014.
- [40] J. Ding, C. X. Fan, and J. X. Lin, "Auxiliary model based parameter estimation for dual-rate output error systems with colored noise," *Applied Mathematical Modelling*, vol. 37, no. 6, pp. 4051–4058, 2013.
- [41] L. Xie, H. Yang, and B. Huang, "FIR model identification of multirate processes with random delays using EM algorithm," *AIChE Journal*, vol. 59, no. 11, pp. 4124–4132, 2013.
- [42] J. Ding and J. Lin, "Modified subspace identification for periodically non-uniformly sampled systems by using the lifting technique," *Circuits, Systems, and Signal Processing*, vol. 33, no. 5, pp. 1439–1449, 2014.
- [43] F. Ding, "Hierarchical estimation algorithms for multivariable systems using measurement information," *Information Sciences*, vol. 277, pp. 396–405, 2014.
- [44] F. Ding, "Combined state and least squares parameter estimation algorithms for dynamic systems," *Applied Mathematical Modelling*, vol. 38, no. 1, pp. 403–412, 2014.
- [45] X. Luan, P. Shi, and F. Liu, "Stabilization of networked control systems with random delays," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 9, pp. 4323–4330, 2011.
- [46] X. Luan, S. Zhao, and F. Liu, " H_∞ control for discrete-time Markov jump systems with uncertain transition probabilities," *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, vol. 58, no. 6, pp. 1566–1572, 2013.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

