IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 16, 2019, accepted November 2, 2019, date of publication November 12, 2019,
date of current version December 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2953126

Recursive Parsimonious Subspace Identification for Closed-Loop Hammerstein Nonlinear Systems

JIE HOU[1], FENGWEI CHEN[2], PENGHUA LI[1], LIJIE SUN[3], AND FEN ZHAO[4]

[1]College of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[2]Department of Automation, Wuhan University, Wuhan 430072, China
[3]College of Electronic and Information Engineering, Taizhou University, Taizhou 317000, China
[4]College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Corresponding authors: Fengwei Chen (fengwei.chen@outlook.com) and Jie Hou (jiehou.phd@hotmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61903057, Grant 61703311, and Grant 61703066, in part by the Chongqing Natural Science Foundation under Grant cstc2019jcyj-msxmX0129, Grant cstc2018jcyjAX0167 and Grant cstc2018jcyjAX0536, in part by the Artificial Intelligence Technology Innovation Significant Theme Special Project of Chongqing Science and Technology Commission under Grant cstc2019jscx-zdztzxX0027, and in part by the Ministry of Education China Mobile Research Fund under Grant MCM20180404.

ABSTRACT In this paper, a recursive closed-loop subspace identification method for Hammerstein nonlinear systems is proposed. To reduce the number of unknown parameters to be identified, the original hybrid system is decomposed as two parsimonious subsystems, with each subsystem being related directly to either the linear dynamics or the static nonlinearity. To avoid redundant computations, a recursive least-squares (RLS) algorithm is established for identifying the common terms in the two parsimonious subsystems, while another two RLS algorithms are established to estimate the coefficients of the nonlinear subsystem and the predictor Markov parameters of the linear subsystem, respectively. Subsequently, the system matrices of the linear subsystem are retrieved from the identified predictor Markov parameters. The convergence of the proposed method is analyzed. Two illustrative examples are shown to demonstrate the effectiveness and merit of the proposed method.

INDEX TERMS Hammerstein-type nonlinear system, subspace identification, closed-loop identification, recursive identification, hierarchical identification.

I. INTRODUCTION

Hammerstein systems can effectively represent and capture the nonlinearity and linear dynamics of many real-world nonlinear processes, e.g., turntable servo systems [1], voltage management systems [2], and blast furnace ironmaking systems [3]. A number of identification methods have been well recognised for identifying Hammerstein systems, including the stochastic approximation methods [4], [5], the blind methods [6], [7], the frequency domain methods [8], [9], and the subspace methods [10], [11], and the least-squares based methods [12]–[14].

Presently, the focus of this paper is on the subspace identification methods (SIMs) for Hammerstein state space systems. Various SIMs have been proposed for Hammerstein state space systems, which have captured a particular attention in the control community. Existing Hammerstein SIMs in the literature can be roughly divided into two categories: the over-parametrization like SIMs (OPM-like SIMs) and the parsimonious SIMs.

In the OPM-like SIMs a linear-time-invariant, over-parametrized hybrid model is first estimated by using a conventional SIM. Then, the underlying system matrices of the nonlinear and dynamic linear elements are extracted from the estimated hybrid model. A number of OPM-like SIMs have been well recognised for identifying Hammerstein state space systems, including the MOESP based Hammerstein SIMs [15], the N4SID based Hammerstein SIMs [16], and the parity-space based Hammerstein SIMs [10], [11]. A persistent excitation input design procedure for OPM-like SIMs for MISO Hammerstein state space systems was developed in [17]. Moreover, for Hammerstein state space models whose

The associate editor coordinating the review of this manuscript and approving it for publication was Lei Jiao.

parameters are slowly time-varying, recursive OPM-like SIMs were proposed [18]–[20]. However, the OPM-like SIMs assume an over-parametrized hybrid model containing extra parameters, which may cause the inflated variance in the estimates and the loss of model fidelity. To overcome this problem, the parsimonious SIMs based on a truncated SVD decomposition have been proposed to reduce the number of unknown parameters to be estimated. These methods are developed by using different regression iterative or recursive algorithms to retrieve the underlying system matrices of the nonlinear subsystem and dynamic linear subsystem, separately. For example, the Gauss-Newton iteration based SIM [21], the two-least-square based SIM [22], and the fixed point iteration based SIM [23].

Note that most of the aforementioned subspace identification methods for Hammerstein state space models are only developed for systems operating in open-loop. For systems operating in closed loop, due to the feedback controller, both the future input and output are correlated with measurement noise, and this gives rise to biased estimates. There are no general methods for recursive subspace identification of Hammerstein state space models in closed loop. A few efforts have been done for iterative subspace identification, for example, reference [21] developed a Gauss-Newton iteration based SIM for closed-loop Hammerstein state space models in innovation form. However, direct extension of the method in [21] to the recursive context seems very complicated.

In this paper, a recursive parsimonious SIM is first proposed for closed-loop Hammerstein state space models in innovation form. The main idea is to model the system by two sub-models rather than a single, overparameterized hybrid model. The main contributions of this paper can be summurized as following.

Firstly, the over-parameterized model to characterize the linear and nonlinear elements of the Hammerstein state space system is decomposed into two sub-models with fewer parameters. Secondly, a recursive least-squares (RLS) algorithm is established for identifying the common terms in the two parsimonious sub-models to avoid redundant estimates, while another two RLS algorithms are established to estimate the coefficients of the nonlinear sub-model and the predictor Markov parameters of the linear subsystem, respectively. Thirdly, convergence properties of the proposed method are presented by using the existing analysis method based on subtly reconstructed output.

The paper is organized as follows. In Section 2 the identification problem is presented. In Section 3 the proposed method is presented, followed by the convergence analysis in Section 4. Section 5 demonstrates the performance of the proposed method. Finally, some conclusions are drawn in Section 6.

II. PROBLEM STATEMENT

Consider a general class of closed-loop systems shown in Fig.1. The plant P is a Hammerstein state space system in

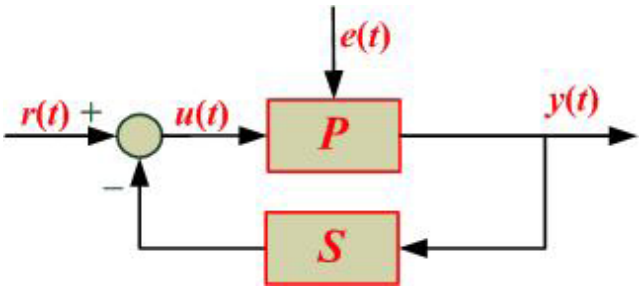


FIGURE 1. The closed-loop Hammerstein nonlinear model structure.

the innovation form described as follows

$$\begin{cases} x(t+1) = Ax(t) + Bf(t) + Ke(t) \\ y(t) = Cx(t) + e(t) \\ f(t) = w_1 f_1(u(t)) + w_2 f_2(u(t)) + \cdots + w_r f_r(u(t)) \end{cases} \tag{1}$$

where $x(t) \in \Re^{n_x}$, $u(t) \in \Re$, $y(t) \in \Re$, $e(t) \in \Re$ are the state vector, the input signal, the output signal, and the innovation, respectively, and (A, B, C, K) are system matrices defined with compatible dimensions. $f(t) \in \Re$ is the static nonlinearity function, which is considered to be a linear combination of the known basis function $f_i(u(t))$ with an unknown coefficient w_i. The feedback controller S is known. $r(t)$ is the external excitation signal specified by the user and is uncorrelated to the measurement noise. We assume that both $u(k)$ and $y(k)$ are rich enough to identify the plant.

Define the following vectors

$$w = [w_1, \cdots, w_r] \tag{2}$$

$$z(t) = [f_1(u(t)), f_2(u(t)), \cdots, f_r(u(t))]^\top \tag{3}$$

Substituting (2) and (3) into (1), we have

$$\begin{cases} x(t+1) = Ax(t) + Bwz(t) + Ke(t) \\ y(t) = Cx(t) + e(t) \end{cases} \tag{4}$$

Further, substituting $e(t) = y(t) - Cx(t)$ into $x(t)$ yields

$$\begin{cases} x(t+1) = \bar{A}x(t) + Bwz(t) + Ky(t) \\ y(t) = Cx(t) + e(t) \end{cases} \tag{5}$$

where $\bar{A} = A - KC$.

To address the concerned identification problem, the following assumptions are made in the paper.

- A1 All the eigenvalues of $\bar{A}$ are within the unit circle.
- A2 (A, C) is observable and (A, B) is controllable.
- A3 The first-nonzero element of w is positive and $\|w\|_2 = 1$.
- A4 The system is minimal and the system order n_x is known.
- A5 The innovation sequence $\{e(t)\}$ is a zero-mean white noise sequence and uncorrelated with the input sequence $\{u(i)\}$ for $t \geq i$.

By iterating $x(t)$ in (5), we have

$$x(t) = \bar{A}^p x(t-p) + L_1 z_p(t) + L_2 y_p(t) \tag{6}$$

with

$$L_1 = [\bar{A}^{p-1}Bw, \bar{A}^{p-2}Bw, \cdots, Bw] \quad (7)$$
$$L_2 = [\bar{A}^{p-1}K, \bar{A}^{p-2}K, \cdots, K] \quad (8)$$
$$z_p(t) = [z^\top(t-p), z^\top(t-p+1), \cdots, z^\top(t-1)]^\top \quad (9)$$
$$y_p(t) = [y^\top(t-p), y^\top(t-p+1), \cdots, y^\top(t-1)]^\top \quad (10)$$

where p is the past window size, which is greater than the system order n_x in general.

Under the assumption A1, the term $\bar{A}^{p-1}$ can be made arbitrarily small by choosing p sufficiently large. Then the first term on the right side of (6) can be neglected. Thus, by substituting (6) into $y(t)$, the following vector autoregressive with exogenous inputs (VARX) model can be obtained

$$\begin{aligned} y(t) &= CL_1 z_p(t) + CL_2 y_p(t) + e(t) \\ &= \sum_{i=1}^{p}((z^\top(t-i)) \otimes 1)\mathrm{vec}(C\bar{A}^{i-1}Bw) \\ &\quad + \sum_{i=1}^{p}((y_p^\top(t) \otimes 1)\mathrm{vec}(C\bar{A}^{i-1}K) + e(t) \\ &= \chi_0(t)\theta_0 + \chi_1(t)\theta_1 + e(t) \end{aligned} \quad (11)$$

with

$$\chi_0(t) = [z^\top(t-1), z^\top(t-2), \cdots, z^\top(t-p)] \quad (12)$$
$$\chi_1(t) = [y_p^\top(t-1), y_p^\top(t-2), \cdots, y_p^\top(t-p)] \quad (13)$$
$$\theta_0 = [(\mathrm{vec}(C\bar{A}^{p-1}Bw))^\top, \cdots, (\mathrm{vec}(CBw))^\top]^\top \quad (14)$$
$$\theta_1 = [(\mathrm{vec}(C\bar{A}^{p-1}K))^\top, \cdots, (\mathrm{vec}(CK))^\top]^\top \quad (15)$$

where $\otimes$ is the Kronecker product operator.

Note that the unknown parameters CL_1 and CL_2 in the overparameterized model (11) can be directly estimated by solving an LS problem. Furthermore, CL_1 is the product of predictor Markov parameter vector $[C\bar{A}^{p-1}B, C\bar{A}^{p-2}B, \cdots, CB]$ and w. In view of this, we have several difficulties to estimate the system matrices A, B, C, w. Firstly, A, B, C, w cannot be directly retrieved from an estimate of CL_1. Secondly, the number of unknown parameters in $[C\bar{A}^{p-1}B, C\bar{A}^{p-2}B, \cdots, CB]$ and w is $p + r$, while the number of unknown parameters in CL_1 is pr. This implies a redundancy in the number of identified parameters, which may give rise to unnecessarily high variance, and also causes a loss of model fidelity. Lastly, an SVD decomposition is needed to retrieve the system parameters B and w from an estimate of Bw. Under model mismatch and using noisy and finite data, a truncated SVD decomposition will result in a loss of model fidelity.

It can be seen that, to get rid of the above mentioned problems, we need to avoid the estimation of CL_1.With this in mind, we are now in the position to state the objective of this paper: develop a recursive method to estimate directly the unknown linear subsystem parameters (A, B, C) and the nonlinear subsystem parameter vector w from the observed data $\{u(t), y(t)\}_{t=1}^{N}$.

III. THE PROPOSED METHOD

In this section, a recursive computationally efficient SIM is developed to identify closed-loop Hammerstein systems.

A. RECURSIVE ESTIMATION OF THE PREDICTOR MARKOV PARAMETERS AND W

To circumvent the problems stemming from estimating CL_1, we suggest to split it into a linear part and a nonlinear part. Then, (11) can be decoupled into two parsimonious submodels as following

$$\begin{aligned} y(t) &= \sum_{i=1}^{p}((wz(t-i))^\top \otimes 1)\mathrm{vec}(C\bar{A}^{i-1}B) + \chi_1(t)\theta_1 + e(t) \\ &= \chi_2(t)\theta_2 + \chi_1(t)\theta_1 + e(t) \end{aligned} \quad (16)$$

$$\begin{aligned} y(t) &= \sum_{i=1}^{p}(z^\top(t-i) \otimes C\bar{A}^{i-1}B)\mathrm{vec}(w) + \chi_1(t)\theta_1 + e(t) \\ &= \chi_3(t)w^\top + \chi_1(t)\theta_1 + e(t) \end{aligned} \quad (17)$$

with

$$\chi_2(t) = [wz(t-1), wz(t-2), \cdots, wz(t-p)] \quad (18)$$
$$\chi_3(t) = \sum_{i=1}^{p}(z^\top(t-i) \otimes C\bar{A}^{i-1}B) \quad (19)$$
$$\theta_2 = [(\mathrm{vec}(C\bar{A}^{p-1}B))^\top, \cdots, (\mathrm{vec}(CB))^\top]^\top \quad (20)$$

Note that θ_1 and θ_2 include all of the predictor Markov parameters for $i = 1, \ldots, p$ as following

$$\bar{h}_i = C\bar{A}^{i-1}K \quad (21)$$
$$\bar{g}_i = C\bar{A}^{i-1}B \quad (22)$$

To avoid redundant estimates of common parameter vector θ_1 in parsimonious models, three objective functions of prediction error with forgetting factors are defined for minimisation:

$$J_1(\hat{\theta}_1(t)) = \sum_{i=1}^{t}\lambda^{t-i}[y(i) - \chi_0(i)\theta_0(i) - \chi_1(i)\hat{\theta}_1(i)]^2 \quad (23)$$
$$J_2(\hat{\theta}_2(t)) = \sum_{i=1}^{t}\lambda^{t-i}[y(i) - \chi_1(i)\theta_1(i) - \chi_2(i)\hat{\theta}_2(i)]^2 \quad (24)$$
$$J_3(\hat{w}^\top(t)) = \sum_{i=1}^{t}\lambda^{t-i}[y(i) - \chi_1(i)\theta_1(i) - \chi_3(i)\hat{w}^\top(i)]^2 \quad (25)$$

where $J_1(\hat{\theta}_1(t))$ is used to estimate the common parameter vector θ_1, $J_2(\hat{\theta}_2(t))$ and $J_3(\hat{w}^\top(t))$ are for estimating the predictor Markov parameters of the linear subsystem, and the coefficients of the nonlinearity function, respectively. $\lambda \in (0, 1]$ is the forgetting factor to leave out the earlier observation data.

Taking the first derivative of $J_1(\hat{\theta}_1(t))$, $J_2(\hat{\theta}_2(t))$ and $J_3(\hat{w}^\top(t))$ with respect to $\hat{\theta}_1(t), \hat{\theta}_2(t)$ and $\hat{w}^\top(t)$, and

then equating them to zero, we obtain

$$\hat{\theta}_1(t) = \left[\sum_{i=1}^{t} \lambda^{t-i}\chi_1^\top(i)\chi_1(i)\right]^{-1} \times \left[\sum_{i=1}^{t} \lambda^{t-i}\chi_1^\top(i)[y(i)-\chi_0(i)\theta_0(i)]\right] \tag{26}$$

$$\hat{\theta}_2(t) = \left[\sum_{i=1}^{t} \lambda^{t-i}\chi_2^\top(i)\chi_2(i)\right]^{-1} \times \left[\sum_{i=1}^{t} \lambda^{t-i}\chi_2^\top(i)[y(i)-\chi_1(i)\theta_1(i)]\right] \tag{27}$$

$$\hat{w}^\top(t) = \left[\sum_{i=1}^{t} \lambda^{t-i}\chi_3^\top(i)\chi_3(i)\right]^{-1} \times \left[\sum_{i=1}^{t} \lambda^{t-i}\chi_3^\top(i)[y(i)-\chi_1(i)\theta_1(i)]\right] \tag{28}$$

Note that Equations (26), (27) and (28) cannot be computed for the time being since $w(t)$ and $\bar{g}_i(t)$ for $i = 1, \ldots, p$ are unknown. Inspired by the hierarchical identification methods [24]–[27], it is proposed to replace $w(t)$ and $\bar{g}_i(t)$ in $\theta_0(t)$ by $\hat{w}(t-1)$ and $\hat{\bar{g}}_i(t-1)$ for computing (26), replace $\theta_1(t)$ and $w(t)$ in $\chi_2(t)$ by $\hat{\theta}_1(t)$ and $\hat{w}(t-1)$ for computing (27), and replace $\theta_1(t)$ and $\bar{g}_i(t))$ in $\chi_3(t)$ by $\hat{\theta}_1(t)$ and $\hat{\bar{g}}_i(t)$ for computing (28). Then, (26), (27) and (28) admit the recursive solutions as follows

$$\hat{\theta}_0(t) = [(\mathrm{vec}(\hat{\bar{g}}_p(t-1)\hat{w}(t-1)))^\top, \cdots, (\hat{\bar{g}}_1(t-1)\hat{w}(t-1)))^\top]^\top \tag{29}$$

$$\hat{\theta}_1(t) = \hat{\theta}_1(t-1) + K_1(t)[y(t) - \chi_0(t)\hat{\theta}_0(t) - \chi_1(t)\hat{\theta}_1(t-1)] \tag{30}$$

$$K_1(t) = P_1(t-1)\chi_1^\top(t)[\lambda + \chi_1(t)P_1(t-1)\chi_1^\top(t)]^{-1} \tag{31}$$

$$P_1(t) = \frac{1}{\lambda}[P_1(t-1) - K_1(t)\chi_1(t)P_1(t-1)] \tag{32}$$

$$\hat{\chi}_2(t) = [\hat{w}(t-1)z(t-1), \cdots, \hat{w}(t-1)z(t-p)] \tag{33}$$

$$\hat{\theta}_2(t) = \hat{\theta}_2(t-1) + K_2(t)[y(t) - \chi_1(t)\hat{\theta}_1(t) - \hat{\chi}_2(t)\hat{\theta}_2(t-1)] \tag{34}$$

$$K_2(t) = P_2(t-1)\chi_2^\top(t)[\lambda + \hat{\chi}_2(t)P_2(t-1)\hat{\chi}_2^\top(t)]^{-1} \tag{35}$$

$$P_2(t) = \frac{1}{\lambda}[P_2(t-1) - K_2(t)\hat{\chi}_2(t)P_2(t-1)] \tag{36}$$

$$\hat{\chi}_3(t) = \sum_{i=1}^{p}(z^\top(t-i)\hat{\bar{g}}_i(t)) \tag{37}$$

$$\hat{w}^\top(t) = \hat{w}^\top(t-1) + K_3(t)[y(t) - \chi_1(t)\hat{\theta}_1(t) - \hat{\chi}_3(t)\hat{w}^\top(t-1)] \tag{38}$$

$$K_3(t) = P_3(t-1)\chi_3^\top(t)[\lambda + \chi_3(t)P_3(t-1)\chi_3^\top(t)]^{-1} \tag{39}$$

$$P_3(t) = \frac{1}{\lambda}[P_3(t-1) - K_3(t)\chi_3(t)P_3(t-1)] \tag{40}$$

To provide a unique solution based on assumption A3, let s be the sign of first nonzero element of $\hat{w}(t)$, the estimate $\hat{w}(t)$ can be obtained after normalization

$$\hat{w}^\top(t) = \hat{w}^\top(t)s/\|\hat{w}^\top(t)\|^2 \tag{41}$$

and then the estimate $\hat{\chi}_2(t)$ can be obtained as follows

$$\hat{\chi}_2(t) = \hat{\chi}_2(t)s \tag{42}$$

To initialize the algorithm, we take $P_1(0) = \eta I, P_1(0) = \eta I$ and $P_3(0) = \eta I$ with η normally a large positive number (e.g., $\eta = 10^6$) and $\hat{\theta}_1(0) = 10^{-6}$, $\hat{\theta}_2(0) = 10^{-6}$, and $\hat{w}(0) = 10^{-6}$.

B. RETRIEVING THE LINEAR SUBSYSTEM PARAMETERS

To obtain the system matrices, we need to recover the system Markov parameters $g_i = CA^{i-1}B$ for $i = 1, \ldots, p$ from predictor Markov parameters $\hat{\bar{g}}_i$ and $\hat{\bar{h}}_i$ in $\hat{\theta}_1(t)$ and $\hat{\theta}_2(t)$, which can be calculated recursively as $\hat{g}_1 = \hat{\bar{g}}_1$ and $\hat{g}_i = \hat{\bar{g}}_i + \sum_{m=1}^{i-1}\hat{\bar{h}}_m\hat{\bar{g}}_{i-m}$ for $2 \le i \le p$. Then construct following Hankel matrix and performing an SVD, we have

$$\hat{T} = \begin{bmatrix} \hat{g}_1 & \hat{g}_2 & \cdots & \hat{g}_{p/2} \\ \hat{g}_2 & \hat{g}_3 & \cdots & \hat{g}_{p/2+1} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{g}_{p/2} & \hat{g}_{p/2+1} & \cdots & \hat{g}_p \end{bmatrix} = USV^\top = \begin{bmatrix} U_{n_x} & U_s \end{bmatrix} \begin{bmatrix} R_{n_x} & 0 \\ 0 & R_s \end{bmatrix} \begin{bmatrix} V_{n_x}^\top \\ V_s^\top \end{bmatrix} \tag{43}$$

where R_{n_x} is the largest n_x eigenvectors of $\hat{T}$. U_{n_x} and $V_{n_x}^\top$ contain the corresponding left and right eigenvectors of U and $V^\top$, respectively. Since the observability matrix and the controllability matrix are [28], [29]

$$\Gamma = [C^\top, (CA)^\top, \cdots, (CA^{p/2})^\top]^\top \tag{44}$$

$$\Delta = [B, AB, \cdots A^{p/2}B]^\top \tag{45}$$

based on the assumption A2, we obtain

$$\hat{\Gamma} = U_{n_x}R_{n_x}^{1/2} \tag{46}$$

$$\hat{\Delta} = R_{n_x}^{1/2}V_{n_x}^\top \tag{47}$$

Then, it is straightforward to obtain

$$\hat{B} = \hat{\Delta}(:, 1:n_x) \tag{48}$$

$$\hat{C} = \hat{\Gamma}(1, :) \tag{49}$$

$$\hat{A} = \hat{\Gamma}^\dagger(1:p/2-1, :)\hat{\Gamma}(2:p/2, :) \tag{50}$$

where Matlab notation : is used for partitioning matrices.

The proposed Recursive Parsimonious Subspace Identification Method, named as RPSIM, is summarized below.

1. Specify the initial values: p, λ, $P_1(0)$, $P_1(0)$, $P_3(0)$, $\hat{\theta}_1(0)$, $\hat{\theta}_2(0)$, and $\hat{w}(0)$.
2. Estimate the predictor Markov parameters and $w(t)$
   - 2.1 Construct $\chi_0(t)$ and $\chi_1(t)$ as (12) and (13).
   - 2.2 Compute $\hat{\theta}_1(t)$, $\hat{\theta}_2(t)$, and $\hat{w}(t)$ as (29)-(40).
   - 2.3 Perform a normalization operation to $\hat{w}$ as (41).
3. Estimate the linear subsystem parameters
   - 3.1 Extract $\hat{\bar{g}}_i$ and $\hat{\bar{h}}_i$ from $\hat{\theta}_1(t)$ and $\hat{\theta}_2(t)$.
   - 3.1 Compute system Markov parameters $\hat{g}_i$.
   - 3.3 Construct the Hankel matrix $\hat{T}$ as (43).
   - 3.4 Perform SVD for $\hat{T}$.
   - 3.5 Compute $\hat{\Gamma}$ and $\hat{\Delta}$ as (46) and (47).
   - 3.5 Compute the linear subsystem matrices from (48), (49), and (50).

IV. CONVERGENCE ANALYSIS

Note that the proposed method can be extended to identify time-varying Hammerstein state space systems without changing any of the arguments or insights of the paper. We will then establish the convergence properties of the proposed method for time-varying Hammerstein state space systems. For this purpose, the following corollary is first given.

Corollary 1 [30]: For the system in (2) and the proposed algorithm in (29), (33) and (37), if there exist α_j and β_j for $j = 1, 2,$ and 3 satisfying $0 < \alpha_i \leq \beta_i < \infty$ and an integer $N \geq \max(p, r)$ such that the following strong persistent excitation (SPE) condition holds

$$\alpha_1 I \leq \frac{1}{N}\sum_{i=0}^{N} \chi_1^\top(t+i)\chi_1(t+i) \leq \beta_1 I \tag{51}$$

$$\alpha_o I \leq \frac{1}{N}\sum_{i=0}^{N} \hat{\chi}_o^\top(t+i)\hat{\chi}_o(t+i) \leq \beta_o I, \quad o = 2, 3 \tag{52}$$

Then $P_j(t)$ in (29), (33) and (37) satisfies [30]

$$\frac{1-\lambda}{\beta_j N} I \leq \lim_{N\to\infty} P_j(t) \leq \frac{1-\lambda}{\alpha_j \lambda^{N-1}} I \tag{53}$$

□

Define the parameter changing rate vector of $\theta_1(t)$

$$\tilde{\theta}_1(t) = \hat{\theta}_1(t) - \theta_1(t) \tag{54}$$

and the parameter changing rate vector of $\theta_1(t)$

$$\bar{\theta}_1(t) = \theta_1(t) - \theta_1(t-1) \tag{55}$$

and all other terms $\tilde{(.)}$ and $\bar{(.)}$, such as, $\tilde{\theta}_2(t)$, $\tilde{w}^\top(t)$, $\bar{\theta}_2(t)$, and $\bar{w}^\top(t)$ are defined in a similar way.

Then the convergence properties of the proposed method can be directly established in the following theorem.

Theorem 1: For the system described by Equation (2), with the SPE condition in Equations (51) and (52) and the assumptions of $E[||\tilde{\theta}_1(0)||^2] = \xi_1 < \infty$, $E[||\tilde{\theta}_2(0)||^2] = \xi_2 < \infty$, and $E[||\tilde{w}(0)||^2] = \xi_3 < \infty$, the estimation errors of $\tilde{\theta}_1(t)$, $\tilde{\theta}_2(t)$, and $\tilde{w}^\top(t)$ given in Equations (29), (33) and (37) of the proposed algorithm satisfy for $t \geq N$

$$E[||\tilde{\theta}_1(t)||^2] \leq \frac{3\lambda^{2(t-N+1)}(1-\lambda)^2\xi_1}{\alpha_1^2\eta^2} + \frac{3p(1-\lambda)\sigma_1^2}{\alpha_1\lambda^{N-1}} + \frac{3(N\beta)^2\sigma_{\theta_1}^2}{\alpha_1^2\lambda^{2(N-1)}(1-\lambda)^2} \tag{56}$$

$$E[||\tilde{\theta}_2(t)||^2] \leq \frac{3\lambda^{2(t-N+1)}(1-\lambda)^2\xi_2}{\alpha_2^2\eta^2} + \frac{3p(1-\lambda)\sigma_2^2}{\alpha_2\lambda^{N-1}} + \frac{3(N\beta)^2\sigma_{\theta_2}^2}{\alpha_2^2\lambda^{2(N-1)}(1-\lambda)^2} \tag{57}$$

$$E[||\tilde{w}(t)||^2] \leq \frac{3\lambda^{2(t-N+1)}(1-\lambda)^2\xi_3}{\alpha_3^2\eta^2} + \frac{3r(1-\lambda)\sigma_3^2}{\alpha_3\lambda^{N-1}} + \frac{3(N\beta)^2\sigma_w^2}{\alpha_3^2\lambda^{2(N-1)}(1-\lambda)^2} \tag{58}$$

where $\sigma_j^2 = \frac{1}{t}\sum_{i=1}^{t}\varepsilon_j^\top(i)\varepsilon_j(i)$ for $j = 1, 2$ and 3, $\sigma_{\theta_1}^2 = \frac{1}{t}\sum_{i=1}^{t}\bar{\theta}_1^\top(i)\bar{\theta}_1(i)$, $\sigma_{\theta_2}^2 = \frac{1}{t}\sum_{i=1}^{t}\bar{\theta}_2^\top(i)\bar{\theta}_2(i)$, $\sigma_w^2 = \frac{1}{t}\sum_{i=1}^{t}\bar{w}^\top(i)\bar{w}(i)$, $\varepsilon_1(t) = -\chi_0(t)\tilde{\theta}_0(t) + e(t)$, $\varepsilon_2(t) = e(t) - \tilde{\chi}_2(t)\theta_2(t) - \chi_1(t)\tilde{\theta}_1(t)$ and $\varepsilon_3(t) = e(t) - \tilde{\chi}_3(t)\theta_3(t) - \chi_1(t)\tilde{\theta}_1(t)$. And $\tilde{\theta}_0(t)$, $\tilde{\chi}_2(t)$, $\tilde{\chi}_3(t)$ are defined similar to $\tilde{\theta}_1(t)$.

Proof: In order to compute the difference between $\hat{\theta}_1(t)$ and the true system parameters $\theta_1(t)$, the output in (11) can readily be elaborated to

$$\begin{aligned} y(t) &= \chi_0(t)\theta_0(t) + \chi_1(t)\theta_1(t) + e(t) \\ &= \chi_0(t)\theta_0(t) + \chi_0(t)\hat{\theta}_0(t) - \chi_0(t)\hat{\theta}_0(t) + \chi_1(t)\theta_1(t) + e(t) \\ &= \chi_0(t)\hat{\theta}_0(t) + \chi_1(t)\theta_1(t) + \varepsilon_1(t) \end{aligned} \tag{59}$$

where

$$\varepsilon_1(t) = -\chi_0(t)\tilde{\theta}_0(t) + e(t) \tag{60}$$

Note that the output is reconstructed as a function about $\hat{\theta}_0(t)$ with estimate error term $\varepsilon_1(t)$ caused by hierarchical computational procedure and noise terms. In this way, the parameters estimated error can be easily derived based on the standard convergence results of the forgetting factor recursive least squares methods in [30], we only gives a brief convergence proof process.

By substituting (59) into (29), we get

$$\begin{aligned} \tilde{\theta}_1(t) &= (I - P_1(t)\chi_1^\top(t)\chi_1(t))\tilde{\theta}_1(t-1) + P_1(t)\chi_1^\top(t) \\ &\quad + (P_1(t)\chi_1^\top(t)\chi_1(t) - I)\bar{\theta}_1(t)\varepsilon_1(t) \\ &= \lambda^t P_1(t)P_1^{-1}(t)\tilde{\theta}_1(0) - P_1(t)\sum_{i=1}^{t}\lambda^{t-i}P_1^{-1}(i)\bar{\theta}_1(i) \\ &\quad + P_1(t)H_1(t)W_1(t) \end{aligned} \tag{61}$$

where

$$\Phi(t) = I - P_1(t)\chi_1^\top(t)\chi_1(t) = \lambda P_1(t)P_1^{-1}(t-1) \tag{62}$$

$$H_1(t) = [\sqrt{\lambda}^{t-1}\chi_1(1), \sqrt{\lambda}^{t-2}\chi_1(2), \cdots, \chi_1(t)] \tag{63}$$

$$W_1(t) = [[\sqrt{\lambda}^{t-1}\varepsilon_1(1)]^\top, [\sqrt{\lambda}^{t-2}\varepsilon_1(2)]^\top, \cdots, [\varepsilon_1(t)]^\top]^\top \tag{64}$$

Referring to the methods in [30], and taking the expectation on the norm of each term at the right-side of (61), respectively, we have

$$E[||\lambda^t P_1(t)P_1^{-1}(t)\tilde{\theta}_1(0)||^2] \leq \frac{\lambda^{2(t-N+1)}(1-\lambda)^2\xi_2}{\alpha_2^2\eta^2} \tag{65}$$

$$E[||P_1(t)H_1(t)W_1(t)||^2] \leq \frac{p(1-\lambda)\sigma_1^2}{\alpha_1\lambda^{N-1}} \tag{66}$$

$$E[||P_1(t)\sum_{i=1}^{t}\lambda^{t-i}P_1^{-1}(i)\bar{\theta}_1(i)||^2] \leq \frac{(N\beta)^2\sigma_{\theta_2}^2}{\alpha_1^2\lambda^{2(N-1)}(1-\lambda)^2} \tag{67}$$

where the maximal eigenvalue of $P_1(t)$ is $\lambda_{max}[P_1(t)]$, $\sigma_1^2 = \frac{1}{t}\sum_{i=1}^{t}\varepsilon_1^\top(i)\varepsilon_1(i)$, $\sigma_{\theta_1}^2 = \frac{1}{t}\sum_{i=1}^{t}\bar{\theta}_1^\top(i)\bar{\theta}_1(i)$.

Taking the expectation on the norm of (61), we have

$$E[||\tilde{\theta}_1(t)||^2] \leq 3E[||\lambda^t P_1(t)P_1^{-1}(t)\tilde{\theta}_1(0)||^2] + 3E[||P_1(t)H_1(t)W_1(t)||^2] + 3E[||P_1(t)\sum_{i=1}^{t}\lambda^{t-i}P_1^{-1}(i)\bar{\theta}_1(i)||^2] \quad (68)$$

Substituting (65), (66), and (67) into (68), gives rise to (56) shown in the theorem.

To obtain the estimation errors of $\tilde{\theta}_2(t)$ and $\tilde{w}^\top(t)$, we rewrite the outputs in (16) and (17) as following,

$$y(t) = \hat{\chi}_2(t)\theta_2(t) + \chi_1(t)\hat{\theta}_1(t) + \varepsilon_2(t) \quad (69)$$
$$y(t) = \hat{\chi}_3(t)w^\top(t) + \chi_1(t)\hat{\theta}_1(t) + \varepsilon_3(t) \quad (70)$$

with

$$\varepsilon_2(t) = e(t) - \tilde{\chi}_2(t)\theta_2(t) - \chi_1(t)\tilde{\theta}_1(t) \quad (71)$$
$$\varepsilon_3(t) = e(t) - \tilde{\chi}_3(t)\theta_3(t) - \chi_1(t)\tilde{\theta}_1(t) \quad (72)$$

The results (57) and (58) shown in the theorem can be easily obtained based on a similar proof procedure for $\tilde{\theta}_1(t)$. □

From the above theorem, we can draw the following conclusions.

- For time-invariant systems, the first terms are equal to zero when $t \to \infty$, the second terms are equal to zero when $\lambda \to 1$, and the third terms are equal to zero in (56), (57) and (58). Namely, the proposed algorithms in (29), (33) and (37) can give a consistent parameter estimation for $\lambda \to 1$ and $t \to \infty$.
- For time-varying systems, due to $\lambda^t \to 0$ for $t \to \infty$, the error upper bounds of the parameter estimates approach to the second and third terms in (56), (57) and (58).

The estimation errors of $\tilde{A}(t), \tilde{B}(t)$ and $\tilde{C}(t)$ are linear functions of $\tilde{\theta}_1(t)$ and $\tilde{\theta}_2(t)$ based on the first-order sensitivity analysis as clarified in the references [31]–[35], which have similar convergence rates with $\tilde{\theta}_1(t)$ and $\tilde{\theta}_2(t)$.

V. ILLUSTRATIVE EXAMPLES

In this section, the performance of the proposed method is illustrated by two examples. One is a benchmark example in [36] and the other is an industrial injection molding precess studied in [23], [37].

Example 1: Consider the following time-varying system with abrupt change and different parameter changing rates extended from [36]

$$\begin{cases} x(t+1) = A(t)x(t) + [\,1 \quad 0\,]^\top f(t) + v_1(t) \\ y(t) = [0.6804 \quad 0.6303]x(t) + v(t) \\ f(t) = w_1(t)u(t) + w_2(t)u^2(k) \end{cases}$$

where the noises $v(t)$ and $v_1(t)$ are assumed to be Gaussian white with zero-mean and variance of 0.1. The time-varying system matrices are given as following

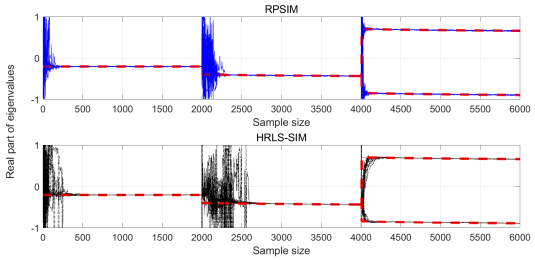


FIGURE 2. Comparison of estimated real part of eigenvalues for Example 1 using RPSIM and HRLS-SIM. The true values are indicated with the red lines.

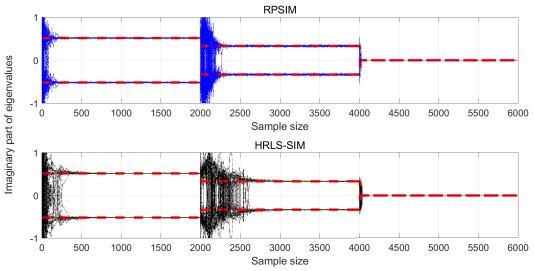


FIGURE 3. Comparison of estimated imaginary part of eigenvalues for Example 1 using RPSIM and HRLS-SIM. The true values are indicated with the red lines.

The input is given by

$$u(t) = r(t) - 0.1y(t)$$

where the external input $r(t)$ is taken as Gaussian white noise with zero-mean and variance of 7. The signal-to-noise ratio (SNR) of the process is 24 dB.

To empirically study the statistical properties of RPSIM, a Monte-Carlo (MC) test with 20 runs is carried out. The model parameters (A, B, C, w) are estimated from an input/output data set of length $N = 6000$ by using RPSIM with $p = 10$ and $\lambda = 0.99$. The proposed method is compared with a SIM by estimating $\hat{w}^\top$, $\hat{\theta}_1$ and $\hat{\theta}_2$ from (16) and (17) using HRLS [38], named as HRLS-SIM.

Figs. 2 and 3 shows the estimation results of the eigenvalues of $A(t)$. Fig.4 shows the estimation results of w of the nonlinear subsystem. Fig.5 shows the identified model transfer function matrix (TFM), i.e.

$$\hat{G}(t) = ||\hat{C}(t)(zI - \hat{A}(t))^{-1}\hat{B}(t)||_2$$

From Figs. 2-5, we can draw the following conclusions.

- All methods can track the time-varying system parameters with high accuracy. While the proposed algorithm shows better performances in terms of initial convergence speed and computational efficiency compared to HRLS-SIM.
- All methods are able to detect and track the system parameter jumps. The RPSIM show better performances in terms of initial convergence speed compared to HRLS-SIM.

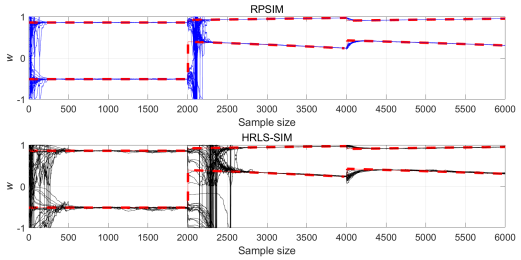


FIGURE 4. Comparison of estimated w for Example 1 using RPSIM and HRLS-SIM. The true poles are indicated with the red lines.

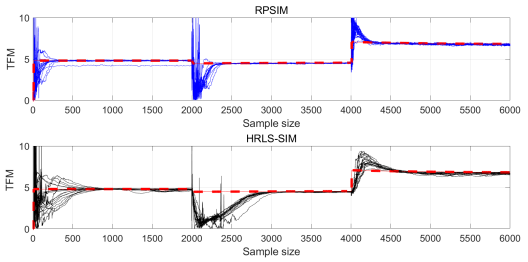


FIGURE 5. Comparison of estimated TFM for Example 1 using RPSIM and HRLS-SIM. The true poles are indicated with the red lines.

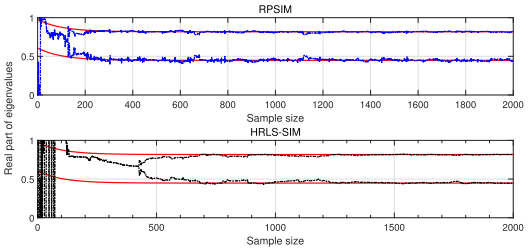


FIGURE 6. Comparison of estimated eigenvalues for Example 2 using RPSIM and HRLS-SIM. The true poles are indicated with the red lines.

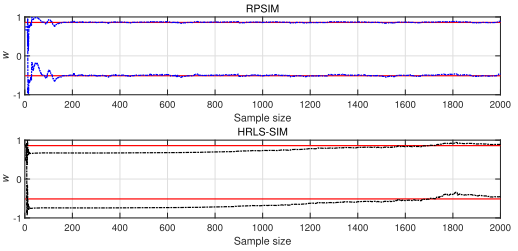


FIGURE 7. Comparison of estimated w for Example 2 using RPSIM and HRLS-SIM. The true poles are indicated with the red lines.

Example 2: Consider an injection moulding process studied in the reference [23], [37]:

$$\begin{cases} x(t+1) = A(t)x(t) + [\,1 \quad 0\,]^\top f(t) + v_2(t) \\ y(t) = \begin{bmatrix} 1.69 & 1.419 \end{bmatrix} x(t) + v_3(t) \\ f(t) = 0.8593u(k) - 0.5115u^2(k) \end{cases}$$

where

$$A(t) = \begin{bmatrix} 1.582 - \dfrac{0.1e^{-(t-1)/2000} - 1}{e^{-1} - 1} & -0.5916 \\ 1 & -\dfrac{0.1e^{-(t-100)/2000} - 1}{e^{-1} - 1} \end{bmatrix}$$

and the input is given by

$$u(t) = r(t) - 0.01y(t)$$

where the external input $r(t)$ is designed as a Pseudo-Random-Binary-Sequence (PRBS) to ensure proper excitation, and the noise is taken as white noise with variance of 1%. The SNR of the process is 20 dB.

The initial conditions are chosen as follows: $p = 10$, $\lambda = 0.99$ and $N = 2000$ for the RPSIM and HRLS-SIM algorithms. Figs. 6, 7, and 8 show the estimated eigenvalues, w, and $G(t)$, respectively. As can be seen from Fig. 7, both the considered algorithms yield consistent estimates: the estimated w converge to the true ones. The results

$$A(t) = \begin{cases} \begin{bmatrix} -0.4120 & -0.3090 \\ 1 & 0 \end{bmatrix} & t \in [1, 2000] \\ \begin{bmatrix} -0.6 - \dfrac{0.1e^{-(t-1)/2000} - 1}{e^{-1} - 1} & -0.2 \\ 1 & -\dfrac{0.1e^{-(t-100)/2000} - 1}{e^{-1} - 1} \end{bmatrix} & t \in [2001, 4000] \\ \begin{bmatrix} 0.5 - \dfrac{0.2e^{-(t-600)/2000} - 1}{e^{-1} - 1} & 0.5 \\ 1 & -\dfrac{0.3e^{-(t-600)/2000} - 1}{e^{-1} - 1} \end{bmatrix} & t \in [4001, 6000] \end{cases}$$

$$w_x(t) = \begin{cases} [0.8593, -0.5115] & t \in [1, 2000] \\ [0.866\frac{0.5e^{-(t-1)/6000}-1}{e^{-1}-1}, 0.5 - 0.5\frac{0.1e^{-(t-1)/6000}-1}{e^{-1}-1}] & t \in [2001, 4000] \\ [0.866\frac{0.5e^{-(t-6000)/4000}-1}{e^{-1}-1}, 0.5 - 0.5\frac{0.1e^{-(t-4000)/6000}-1}{e^{-1}-1}] & t \in [4001, 6000] \end{cases} \quad w(t) = w_x(t)/\|w_x(t)\|_2$$

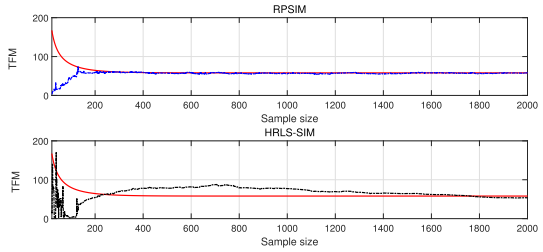


FIGURE 8. Comparison of estimated TFM for Example 2 using RPSIM and HRLS-SIM. The true poles are indicated with the red lines.

(see Figs. 6 and 8) show that both techniques present interesting abilities to track variations in the system parameters. On the other hand, in the considered example, the proposed algorithm shows a better performance in terms of transient behaviour with respect to HRLS-SIM. It is again seen that the proposed method can track the time-variant system parameters with high accuracy.

VI. CONCLUSION

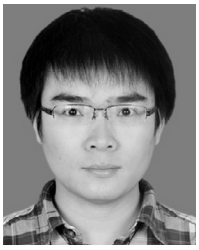
In this paper, a recursive computationally efficient SIM for nonlinear Hammerstein state space model identification has been presented. The first step of the proposed identification method is to estimate the coefficient vector of the nonlinear subsystem and the predictor Markov parameters of the linear subsystem separately by using a computationally efficient three-stage RLS based method. Then the system matrices are extracted from the identified predictor Markov parameters using a SVD based method. Convergence of the proposed method has been analyzed by using existing analysis method based on subtly reconstructed output. The applications to two time-varying illustrative examples have well demonstrated the effectiveness and merit of the proposed method.

REFERENCES

[1] X. Gao, X. Ren, C. Zhu, and C. Zhang, "Identification and control for Hammerstein systems with hysteresis non-linearity," *IET Control Theory Appl.*, vol. 9, no. 13, pp. 1935–1947, Aug. 2015.
[2] Y. Bao, L. Y. Wang, C. Wang, and Y. Wang, "Hammerstein models and real-time system identification of load dynamics for voltage management," *IEEE Access*, vol. 6, pp. 34598–34607, 2018.
[3] P. Zhou, H. Song, H. Wang, and T. Chai, "Data-driven nonlinear subspace modeling for prediction and control of molten iron quality indices in blast furnace ironmaking," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 5, pp. 1761–1774, Sep. 2017.
[4] B. Mu, H.-F. Chen, L. Y. Wang, G. Yin, and W. X. Zheng, "Recursive identification of Hammerstein systems: Convergence rate and asymptotic normality," *IEEE Trans. Autom. Control*, vol. 62, no. 7, pp. 3277–3292, Jul. 2017.
[5] B.-Q. Mu and H.-F. Chen, "Recursive identification of multi-input multi-output errors-in-variables Hammerstein systems," *IEEE Trans. Autom. Control*, vol. 60, no. 3, pp. 843–849, Mar. 2015.
[6] E. W. Bai and F. Y. Fu, "Recursive identification of multi-input multi-output errors-in-variables Hammerstein systems," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1610–1619, Jul. 2002.
[7] F. R. Ávila, L. T. Duarte, and L. W. P. Biscainho,"On the sparsity-based identification and compensation of Hammerstein systems," *IEEE Signal Process. Lett.*, vol. 24, no. 9, pp. 1363–1367, Sep. 2017.
[8] E.-W. Bai, "Frequency domain identification of Hammerstein models," *IEEE Trans. Autom. Control*, vol. 48, no. 4, pp. 530–542, Apr. 2003.
[9] J. P. Noël and G. Kerschen, "Frequency-domain subspace identification for nonlinear mechanical systems," *Mech. Syst. Signal Process.*, vol. 40, no. 2, pp. 701–717, Nov. 2013.
[10] J. Hou, T. Liu, B. Wahlberg, and M. Jansson, "Subspace Hammerstein model identification under periodic disturbance," *IFAC-PapersOnLine*, vol. 51, no. 15, pp. 335–340, Jul. 2018.
[11] J. Hou, T. Liu, and Q.-G. Wang, "Subspace identification of Hammerstein-type nonlinear systems subject to unknown periodic disturbance," *Int. J. Control*, to be published, doi: 10.1080/00207179.2019.1621385.
[12] F. Ding, H. Chen, L. Xu, J. Dai, Q. Li, and T. Hayat, "A hierarchical least squares identification algorithm for Hammerstein nonlinear systems using the key term separation," *J. Franklin Inst.*, vol. 355, no. 8, pp. 3737–3752, May 2018.
[13] Y. Mao, F. Ding, L. Xu, and T. Hayat, "Highly efficient parameter estimation algorithms for Hammerstein non-linear systems," *IET Control Theory Appl.*, vol. 13, no. 4, pp. 477–485, Mar. 2019.
[14] F. Ding, Y. Shi, and T. Chen, "Auxiliary model-based least-squares identification methods for Hammerstein output-error systems," *Syst. Control Lett.*, vol. 56, no. 5, pp. 373–380, Mar. 2007.
[15] M. Verhaegen and D. Westwick, "Identifying MIMO Hammerstein systems in the context of subspace model identification methods," *IFAC Proc. Volumes*, vol. 29, no. 1, pp. 4092–4097, Jun./Jul. 1996.
[16] I. Goethals, K. Pelckmans, J. A. K. Suykens, and B. De Moor, "Subspace identification of Hammerstein systems using least squares support vector machines," *IEEE Trans. Autom. Control*, vol. 50, no. 10, pp. 1509–1519, Oct. 2005.
[17] A. Naitali and F. Giri, "Persistent excitation by deterministic signals for subspace parametric identification of MISO Hammerstein systems," *IEEE Trans. Autom. Control*, vol. 61, no. 1, pp. 258–263, Jan. 2016.
[18] L. Bako, G. Mercere, S. Lecoeuche, and M. Lovera, "Recursive subspace identification of Hammerstein models based on least squares support vector machines," *IET Control Theory Appl.*, vol. 3, no. 9, pp. 1209–1216, Sep. 2009.
[19] X. Chen and H.-T. Fang, "Recursive identification for Hammerstein systems with state-space model," *Acta Automatica Sinica.*, vol. 36, no. 10, pp. 1460–1467, 2010.
[20] S. Zhang, J. Hou, J. Du, and T. Liu, "Recursive subspace identification of Hammerstein-type nonlinear systems under slow time-varying load disturbance," in *Proc. Chin. Automat. Congr. (CAC)*, Nov./Dec. 2018, pp. 870–874.
[21] G. J. van der Veen, J.-W. van Wingerden, and M. Verhaegen, "Global identification of wind turbines using a Hammerstein identification method," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 4, pp. 1471–1478, Jul. 2013.
[22] K. Jalaleddini and R. E. Kearney, "Subspace identification of SISO Hammerstein systems: Application to stretch reflex identification," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2725–2734, Oct. 2013.
[23] J. Hou, F. Chen, P. Li, and Z. Zhu, "Fixed point iteration-based subspace identification of Hammerstein state-space models," *IET Control Theory Appl.*, vol. 13, no. 8, pp. 1173–1181, May 2019.
[24] F. Ding and T. Chen, "Hierarchical gradient-based identification of multivariable discrete-time systems," *Automatica*, vol. 41, no. 2, pp. 315–325, Feb. 2005.
[25] L. Xu, W. Xiong, A. Alsaedi, and T. Hayat, "Hierarchical parameter estimation for the frequency response based on the dynamical window data," *Int. J. Control Autom. Syst.*, vol. 16, no. 4, pp. 1756–1764, Aug. 2018.
[26] D. Wang, X. Zhang, and B. Xue, "Decoupled parameter estimation methods for Hammerstein systems by using filtering technique," *IEEE Access*, vol. 6, pp. 66612–66620, 2018.
[27] J. Hou, T. Liu, and Q.-G. Wang, "Recursive subspace identification subject to relatively slow time-varying load disturbance," *Int. J. Control*, vol. 91, no. 3, pp. 622–638, Mar. 2018.
[28] C. P. Yu, L. Ljung, and M. Verhaegen, "Identification of structured state-space models," *Automatica*, vol. 90, pp. 54–61, Apr. 2018.
[29] C. Yu, J. Chen, and M. Verhaegen, "Subspace identification of individual systems in a large-scale heterogeneous network," *Automatica*, vol. 109, Nov. 2019, Art. no. 108517.
[30] F. Ding and T. Chen, "Performance bounds of forgetting factor least-squares algorithms for time-varying systems with finite measurement data," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 3, pp. 555–566, Mar. 2005.
[31] J. Hou, F. Chen, P. Li, Z. Zhu, and F. Liu, "An improved consistent subspace identification method using parity space for state-space models," *Int. J. Control, Autom. Syst.*, vol. 17, no. 5, pp. 1167–1176, May 2019.
[32] J. Hou, T. Liu, and F. Chen, "Orthogonal projection based subspace identification against colored noise," *Control Theory Technol.*, vol. 15, no. 1, pp. 69–77, Feb. 2017.

[33] J. Hou, F. Chen, P. Li, and Z. Zhu, "Prior-knowledge-based subspace identification for batch processes," *J. Process Control*, vol. 82, pp. 22–30, Oct. 2019.
[34] J. Hou and T. Liu, "Closed-loop subspace model identification using innovation estimation and orthogonal projection," *Acta Automatica Sinica*, vol. 42, no. 11, pp. 1657–1663, Nov. 2016.
[35] M. Jansson, "Asymptotic variance analysis of subspace identification methods," *IFAC Proc. Volumes*, vol. 33, no. 15, pp. 91–96, Jun. 2000.
[36] C. Yan, J. Wang, and Q. Zhang, "Subspace identification methods for Hammerstein systems: Rank constraint and dimension problem," *Int. J. Control*, vol. 83, pp. 2204–2216, Nev. 2010.
[37] T. Liu, B. Huang, and S. J. Qin, "Bias-eliminated subspace model identification under time-varying deterministic type load disturbance," *J. Process Control*, vol. 25, pp. 41–49, Jan. 2015.
[38] F. Ding, G. Liu, and X. P. Liu, "Partially coupled stochastic gradient identification methods for non-uniformly sampled systems," *IEEE Trans. Autom. Control*, vol. 55, no. 8, pp. 1976–1981, Aug. 2010.

JIE HOU was born in Yunnan, China. He received the B.Eng. degree from North Minzu University, China, in 2010, the M.Eng. degree from Chongqing University, China, in 2013, and the Ph.D. degree in control theory and control engineering from the Dalian University of Technology, China, in 2018. Since 2018, he has been with the Chongqing University of Posts and Telecommunications, where he is currently a Lecturer. His current research interests include system identification and modeling.



FENGWEI CHEN was born in Chongqing, China. He received the B.Eng. and M.Eng. degrees from Wuhan University, Wuhan, China, in 2009 and 2011, respectively, and the Ph.D. degree from the University de Lorraine, Nancy, France, in 2014. He was a Lecturer with the Dalian University of Technology, Dalian, China, till 2016. He is currently a Postdoctoral Researcher with Wuhan University. His research interests include system identification, parameter estimation, and wireless power transfer.

PENGHUA LI was born in 1984. He received the M.Sc. degree in electronic engineering and the Ph.D. degree in control theory and control engineering from the University of Chongqing, Chongqing, China, in 2008, and 2012, respectively. He is currently an Associate Professor with the Chongqing University of Posts and Telecommunications. His research fields focus on principle and application of neural computation. His main research interests include estimation and prediction of time series, recognition of image and speech signals, and fault diagnosis of energy storage systems.

LIJIE SUN was born in Liaoning, China. She received the B.Eng. degree from BOHAI University, China, in 2010, the M.Eng. degree from Liaoning University, China, in 2013, and the Ph.D. degree in control theory and control engineering from the Dalian University of Technology, China, in 2018. Since 2018, she has been with Taizhou University, where she is currently a Lecturer. Her current research interests include fault diagnosis and prediction control.

FEN ZHAO was born in Shandong, China. She is currently pursuing the Ph.D. degree with the Chongqing University of Posts and Telecommunications, China. Her research interests include NLP, intelligent control theory, pattern recognition, and fault diagnosis.