WILEY | Hindawi

Research Article

Instrumental Variable-Based OMP Identification Algorithm for Hammerstein Systems

Shuo Zhang [ID],[1] Dongqing Wang [ID],[1,2] and Yaru Yan [ID][1]

[1]*College of Automation Engineering, Qingdao University, Qingdao 266071, China*
[2]*Collaborative Innovation Center for Eco-Textiles of Shandong Province, Qingdao 266071, China*

Correspondence should be addressed to Dongqing Wang; dqwang64@163.com

Hammerstein systems are formed by a static nonlinear block followed by a dynamic linear block. To solve the parameterizing difficulty caused by parameter coupling between the nonlinear part and the linear part in a Hammerstein system, an instrumental variable method is studied to parameterize the Hammerstein system. To achieve in simultaneously identifying parameters and orders of the Hammerstein system and to promote the computational efficiency of the identification algorithm, a sparsity-seeking orthogonal matching pursuit (OMP) optimization method of compressive sensing is extended to identify parameters and orders of the Hammerstein system. The idea is, by the filtering technique and the instrumental variable method, to transform the Hammerstein system into a simple form with a separated nonlinear expression and to parameterize the system into an autoregressive model, then to perform an instrumental variable-based orthogonal matching pursuit (IV-OMP) identification method for the Hammerstein system. Simulation results illustrate that the investigated method is effective and has advantages of simplicity and efficiency.

1. Introduction

Nonlinear system modeling and identification are very important in theory and application [1–6]; block-oriented nonlinear systems, which combine nonlinear and linear blocks in various styles, are the typical representation of nonlinear systems. Hammerstein systems, a static nonlinear block plus a dynamic linear block, are a type of commonly used block-oriented nonlinear systems, and a lot of work has contributed in identification methods for Hammerstein systems. The traditional identification methods for Hammerstein systems mainly include the overparameterization model-based methods [7–10] and the recursive/iterative identification methods [11–14]. The highly efficient identification methods include the key-term separation principle-based identification methods [15–17], the hierarchical identification methods [18, 19], the filtering technique based-identification methods [20–22], the maximum likelihood estimation methods [23–25], and the evolution optimization methods [26–28]. Recently, Chen et al. investigated a

particle swarm optimization method [29], Krishnanathan et al. discussed a continuous-time nonlinear systems using approximate Bayesian computation [30], and Wang et al. studied a model recovery for Hammerstein systems using the auxiliary model-based orthogonal matching pursuit method [31] and so on.

It is difficult to parameterize Hammerstein systems due to existing parameter coupling between the nonlinear part and the linear part of Hammerstein systems. In order to solve this problem, an instrumental variable-based method is studied to parameterize the Hammerstein systems. Further, to achieve in simultaneously identifying parameters and orders and to promote the computational efficiency of the estimated method, a sparsity-seeking orthogonal matching pursuit optimization method of compressive sensing is extended to identify parameters and orders of the Hammerstein systems. The idea is, by the filtering technique and the instrumental variable method, to transform the Hammerstein system into a simple form with a separated nonlinear expression, so as to easily parameterize the system into

autoregressive form, and to perform the instrumental variable-based orthogonal matching pursuit (IV-OMP) identification method for the Hammerstein systems.

In the compressive sensing theory [32–34], there mainly exist two sparse signal recovery methods: the orthogonal matching pursuit algorithms and the basis pursuit algorithms. The orthogonal matching pursuit (OMP) algorithm is a kind of greedy parameter recovering method [35–37], it selects the best fitting column of the measurement matrix and the corresponding sparse signal in each selected step. Due to the selection being orthogonal, the OMP algorithm has a lower computational complexity compared with the basis pursuit algorithms [38–40].

Recently, Mao et al. investigated parameter estimation algorithms for Hammerstein time-delay systems based on the OMP scheme to estimate the system parameters and the time delay [41] by means of the compressed sensing recovery theory and the auxiliary model identification idea. In this paper, an instrumental variable-based orthogonal matching pursuit (IV-OMP) algorithm is investigated to simultaneously estimate the orders and parameters of a Hammerstein system. The contributions lie in four aspects:

(i) To solve the parameterizing difficulty caused by the coupling between the nonlinear part and the linear part in Hammerstein systems, a filtering technique-based instrumental variable method is studied to parameterize the Hammerstein systems.

(ii) To achieve in simultaneously identifying parameters and orders and to promote the computational efficiency of the system, an instrumental variable-based orthogonal matching pursuit (IV-OMP) optimization method of compressive sensing is extended to identify parameters and orders of the systems.

(iii) The advantage of the proposed IV-OMP method over the traditional methods is that it is not necessary to collect a lot of data and invest a lot of power into the parameter identification based on the sparse principle.

(iv) Simulation results illustrate that the investigated method is effective and has advantages of simplicity and efficiency.

The rest of the paper is organized as follows. Section 2 demonstrates the identification problem of a Hammerstein system. Section 3 studies the parameterizing method of the Hammerstein system. Section 4 presents the IV-OMP identification algorithm by using the instrumental variables. Section 5 provides a numerical example for the proposed method. Finally, the concluding remarks are involved in Section 6.

2. The Problem Formulation

For the narrative convenience, we define some notation. "$\mathbf{M} =: \mathbf{N}$" represents "$\mathbf{M}$ is defined as $\mathbf{N}$"; z represents a unit forward shift operator: $zy(t) = \mathbf{y}(t+1)$ and $z^{-1}y(t) = y(t-1)$;
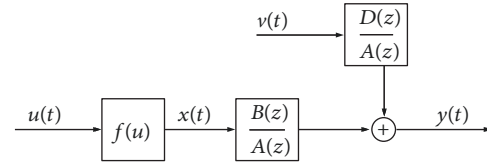

Figure 1: A system described by the Hammerstein CARMA model.

$\widehat{x}(t)$ stand for the estimate of x at time t. The input nonlinear and output linear functions of a Hammerstein CARMA system in Figure 1 are expressed as

$$x(t) = f[u(t)] = \sum_{k=1}^{n_c} c_k f_k[u(t)], \tag{1}$$

$$A(z)y(t) = B(z)x(t) + D(z)v(t), \tag{2}$$

where $u(t)$ and $y(t)$ are the system input and output, $x(t)$ is an internal variable, and $v(t)$ is stochastic white noise with zero mean; the input nonlinearity f is modeled as a linear combination of basis functions f_k; n_c is the number of the basis functions; the linear block is a CARMA model; $A(z)$, $B(z)$, and $D(z)$ are polynomials in the unit backward shift operator z^{-1}($z^{-1}y(t) = y(t-1)$) and defined by

$$A(z) := 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{n_a} z^{-n_a},$$
$$B(z) := 1 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_b} z^{-n_b}, \tag{3}$$
$$D(z) := 1 + d_1 z^{-1} + d_2 z^{-2} + \cdots + d_{n_d} z^{-n_d}.$$

If the order n_a is known and the orders n_b, n_c, and n_d are unknown, then we face two parameterizing problems for the Hammerstein system:

(i) The coupling between the nonlinear part and the linear part in a Hammerstein system makes the parametrization of the Hammerstein system more difficult.

(ii) Simultaneous and efficient identification of system parameters and orders is difficult to be achieved, due to existing unknown orders.

3. The Parametrization of The Hammerstein CARMA System

Divide both sides of (2) by the filtering function $B(z)$, we get

$$\frac{A(z)}{B(z)}y(t) = x(t) + \frac{D(z)}{B(z)}v(t). \tag{4}$$

Define the instrumental variables $z(t)$ and $w(t)$ as

$$z(t) := \frac{A(z)}{B(z)}y(t), \tag{5}$$

$$w(t) := \frac{D(z)}{B(z)}v(t). \tag{6}$$

Then, (4) can be written as

$$z(t) = x(t) + w(t). \tag{7}$$

Replacing the expressions of $A(z)$ and $B(z)$ and $B(z)$ and $D(z)$ into (5) and (6), respectively, we can get

$$z(t) = -\sum_{i=1}^{n_b} b_i z(t-i) + y(t) + \sum_{i=1}^{n_a} a_i y(t-i), \tag{8}$$

$$w(t) = -\sum_{i=1}^{n_b} b_i w(t-i) + v(t) + \sum_{i=1}^{n_d} d_i v(t-i). \tag{9}$$

Substituting $z(t)$ and $w(t)$ in (8) and (9) and $x(t)$ in (1) into (7) gives

$$y(t) = -\sum_{i=1}^{n_a} a_i y(t-i) + \sum_{i=1}^{n_c} c_i f_i[u(t)]$$
$$+ \sum_{i=1}^{n_b} b_i (z(t-i) - w(t-i)) \tag{10}$$
$$+ \sum_{i=1}^{n_d} d_i v(t-i) + v(t).$$

Since the orders n_b, n_c, and n_d are unknown, we assume a sufficient length of order L for n_b, n_c, and n_d ($L > n_b, L > n_c, L > n_d$). Then, (10) can be rewritten as

$$y(t) = -\sum_{i=1}^{n_a} a_i y(t-i) + \sum_{i=1}^{L} c_i f_i[u(t)]$$
$$+ \sum_{i=1}^{L} b_i (z(t-i) - w(t-i)) \tag{11}$$
$$+ \sum_{i=1}^{L} d_i v(t-i) + v(t).$$

Define the information vectors $\boldsymbol{\varphi}(t)$, $\boldsymbol{\varphi}_u(t)$, and $\boldsymbol{\varphi}_v(t)$ as

$$\boldsymbol{\varphi}(t) := \begin{bmatrix} \varphi_u(t) \\ \varphi_v(t) \end{bmatrix} \in \mathbb{R}^N, \quad N := 3L + n_a,$$

$$\boldsymbol{\varphi}_u(t) := [-y(t-1), -y(t-2), \dots, -y(t-n_a),$$
$$\cdot f_1[u(t)], f_2[u(t)], \dots, f_L[u(t)]]^T \in \mathbb{R}^{L+n_a},$$

$$\boldsymbol{\varphi}_v(t) := [z(t-1) - w(t-1), z(t-2) - w(t-2), \dots,$$
$$\cdot z(t-L) - w(t-L), v(t-1), v(t-2), \dots,$$
$$\cdot v(t-L)]^T \in \mathbb{R}^{2L}$$

$$\tag{12}$$

and the parameter vectors $\boldsymbol{\Theta}$, $\boldsymbol{\theta}_u$ and θ_v, as

$$\boldsymbol{\Theta} := \begin{bmatrix} \boldsymbol{\theta}_u \\ \boldsymbol{\theta}_v \end{bmatrix} \in \mathbb{R}^N,$$

$$\boldsymbol{\theta}_u := \left[a_1, a_2, \dots, a_{n_a}, c_1, c_2, \dots, c_{n_c}, \underbrace{0, 0, \dots, 0}_{L-n_c}\right]^T \in \mathbb{R}^{L+n_a},$$

$$\boldsymbol{\theta}_v := \left[b_1, b_2, \dots, b_{n_b}, \underbrace{0, 0, \dots, 0}_{L-n_b}, d_1, d_2, \dots, d_{n_d}, \underbrace{0, 0, \dots, 0}_{L-n_d}\right]^T \in \mathbb{R}^{2L}.$$

$$\tag{13}$$

Then, (11) can be written as

$$y(t) = \varphi_u^T(t)\boldsymbol{\theta}_u + \varphi_v^T(t)\boldsymbol{\theta}_v + v(t) = \varphi^T(t)\boldsymbol{\Theta} + v(t). \tag{14}$$

Sampling m sets of data ($t = 1, 2, \dots, m$) and substituting them into (14) get

$$y(1) = \varphi^T(1)\boldsymbol{\Theta} + v(1),$$
$$y(2) = \varphi^T(2)\boldsymbol{\Theta} + v(2),$$
$$\dots, \tag{15}$$
$$y(m) = \varphi^T(m)\boldsymbol{\Theta} + v(m).$$

Define the accumulated information vectors and the matrix,

$$\mathbf{Y} := \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(m) \end{bmatrix} \in \mathbb{R}^m,$$

$$\mathbf{V} := \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(m) \end{bmatrix} \in \mathbb{R}^m, \tag{16}$$

$$\boldsymbol{\Phi} := \begin{bmatrix} \varphi^T(1) \\ \varphi^T(2) \\ \vdots \\ \varphi^T(m) \end{bmatrix} \in \mathbb{R}^{m \times N},$$

then (15) can be described as

$$\mathbf{Y} = \boldsymbol{\Phi}\boldsymbol{\Theta} + \mathbf{V}. \tag{17}$$

If there are enough measurements, that is, m reaches several thousands, we can get the least squares estimate of $\boldsymbol{\Theta}$, as follows:

$$\widehat{\boldsymbol{\Theta}}_{LS} = \left(\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{-1}\boldsymbol{\Phi}^T\mathbf{Y}. \tag{18}$$

But with an assumed big order length L (≥ 5) and N, the above least squares algorithm leads a big computational burden and is not suitable for solving system parameters and orders on line. To avoid this problem and get an efficient

identification algorithm, we choose a more advantageous method over the traditional least squares method; the aim is to identify the parameter vector Θ with less observations $(K < m < n)$ by using the OMP theory based on the mentioned instrumental variable method.

According to the CS theory, we can regard the parameter vector Θ as a sparse signal. Let $\|\Theta\|_0 = K$ be the number of the nonzero entries in Θ, that is, the sparsity level of Θ, then the identification problem can be described as an orthogonal matching pursuit (OMP) method:

$$\widehat{\Theta} = \arg \min \|\Theta\|_0,$$
$$\text{s.t.} \quad \left\| y - \varphi^{\mathrm{T}}(t)\Theta \right\|_2 < \varepsilon, \tag{19}$$

where $\widehat{\Theta}$ is the estimate of Θ and $\varepsilon(\varepsilon > 0)$ is the error tolerance.

4. The IV-OMP Identification Algorithm

For the sparse parameter vector Θ, if $K = \|\Theta\|_0 (K < N)$, there will be K nonzero scalar components in Θ. What we should do in this identification method is to pick out and recover valid data in Θ from Φ by using the OMP theory.

Define ϕ_i as the ith column vector of Φ (a ϕ_i is also called an atom), and θ_i as the ith element of $\Theta(i = 1, 2, \ldots, N)$, that is,

$$\Phi = [\phi_1, \phi_2, \ldots, \phi_N],$$
$$\Theta = [\theta_1, \theta_2, \ldots, \theta_N]^T. \tag{20}$$

Then, the vector $\mathbf{Y}$ in (17) can be written as

$$\mathbf{Y} = \sum_{i=1}^{N} \phi_i \theta_i + V. \tag{21}$$

Obviously, the output vector $\mathbf{Y}$ contains a linear combination of all atoms $\phi_i (i = 1, 2, \ldots, N)$. The main idea is to find the K nonzero items and corresponding valid atoms at the right-hand side of (21).

To describe the IV-OMP recovery algorithm, we need to give some notations. Let $k = 1, 2, \ldots$ be the iterative number and λ_k be the index of the solution support ϕ_{λ_k} in Φ at the kth iteration; Λ_k is a set composed of λ_i, $i = 1, 2, \ldots, k$; $\Phi_{\Lambda k}$ is the submeasurement matrix composed by the k columns of Φ indexed by Λ_k; r_k denotes the residual at the kth iteration; and $\widehat{\Theta}_k$ is the parameter estimation at the kth iteration. The algorithm is initialized as $r_0 = \mathbf{Y}$, $\Lambda_0 = \varnothing$ and $\widehat{\Theta}_{\Lambda_0} = 0$.

Define a cost function at the kth iteration

$$J(\theta_i) = \|r_{k-1} - \phi_i \theta_i\|^2, \quad i = 1, 2, \ldots, N. \tag{22}$$

Remark 1. Minimizing $J(\theta_i)$ with respect to θ_i means that the derivation of $J(\theta_i)$ with respect to θ_i is equal to 0, that is,

$$\frac{\partial J(\theta_i)}{\partial \theta_i} = -2\phi_i^{\mathrm{T}}[r_{k-1} - \phi_i \theta_i] = 0, \tag{23}$$

and we get

$$\theta_i = \frac{\phi_i^{\mathrm{T}} r_{k-1}}{\|\phi_i\|^2}. \tag{24}$$

Substitute the above θ_i into (22) to get a minimized $J(\theta_i)$ as

$$
\begin{aligned}
J(\theta_i)_{\min} &= \left\| \phi_i \frac{\phi_i^T r_{k-1}}{\|\phi_i\|^2} - r_{k-1} \right\|^2 \\
&= \left(\frac{\phi_i \phi_i^T r_{k-1}}{\|\phi_i\|^2} - r_{k-1} \right)^T \left(\frac{\phi_i \phi_i^T r_{k-1}}{\|\phi_i\|^2} - r_{k-1} \right) \\
&= \left(\frac{r_{k-1}^T \phi_i \phi_i^T}{\|\phi_i\|^2} - r_{k-1}^T \right) \left(\frac{\phi_i \phi_i^T r_{k-1}}{\|\phi_i\|^2} - r_{k-1} \right) \\
&= \frac{r_{k-1}^T \phi_i \phi_i^T \phi_i \phi_i^T r_{k-1}}{\|\phi_i\|^2 \|\phi_i\|^2} - \frac{r_{k-1}^T \phi_i \phi_i^T r_{k-1}}{\|\phi_i\|^2} \\
&\quad - \frac{r_{k-1}^T \phi_i \phi_i^T r_{k-1}}{\|\phi_i\|^2} + r_{k-1}^T r_{k-1} \\
&= \|r_{k-1}\|^2 - \frac{\left(\phi_i^T r_{k-1}\right)^2}{\|\phi_i\|^2}.
\end{aligned}
\tag{25}
$$

Remark 2. The result $[J(\theta_i)_{\min} = \|r_{k-1}\|^2 - ((\phi_i^{\mathrm{T}} r_{k-1})^2 / \|\phi_i\|^2)]$ says that minimizing $J(\theta_i)$ is equivalent to the quest for the largest inner product between the residual r_{k-1} and the normalized column vector ϕ_i of Φ.

The next step is to find the index i and the column ϕ_i corresponding to the largest inner product between the residual r_{k-1} and the normalized column vectors ϕ_i and assign them into λ_k and $\phi_{\lambda k}$ at the kth step:

$$\lambda_k = \arg \max_{i=1,2,\ldots,N} \left| \left\langle r_{k-1}, \frac{\phi_i}{\|\phi_i\|} \right\rangle \right|. \tag{26}$$

Remark 3. Minimizing the least squares cost function $J(\Theta_{\Lambda_k}) = \|\mathbf{Y} - \Phi_{\Lambda_k} \Theta_{\Lambda_k}\|^2$ is equal to taking the derivation of $J(\Theta_{\Lambda_k})$ with respect to Θ_{Λ_k},

$$\frac{\partial J(\Theta_{\Lambda_k})}{\partial \Theta_{\Lambda_k}} = \Phi_{\Lambda_k}^{\mathrm{T}} \left(\mathbf{Y} - \Phi_{\Lambda_k} \Theta_{\Lambda_k} \right) = \Phi_{\Lambda_k}^{\mathrm{T}} r_k = 0. \tag{27}$$

It means that the residual r_k is orthogonal to the subinformation matrix Φ_{Λ_k}. Therefore, the index corresponding to the largest inner product between the residual r_{k-1} and the normalized column vector ϕ_i is computed by

$$\lambda_k = \arg \max_{i \in S \backslash \Lambda_{k-1}} \left| \left\langle r_{k-1}, \frac{\phi_i}{\|\phi_i\|} \right\rangle \right|, \quad S = \{1, 2, \ldots, N\}. \tag{28}$$

Update the support set Λ_k and the subinformation matrix Φ_{Λ_k} by

$$\Lambda_k = [\Lambda_{k-1}, \lambda_k],$$
$$\Phi_{\Lambda_k} = \left[\Phi_{\Lambda_{k-1}}, \phi_{\lambda_k} \right]. \tag{29}$$

With the obtained index set $\mathbf{\Lambda}_k$ and the corresponding subinformation matrix $\mathbf{\Phi}_{\Lambda_k}$, the least squares estimation for the nonzero parameters at the kth iteration can be achieved. Define a cost function:

$$J\left(\mathbf{\Theta}_{\Lambda_k}\right) = \left\|\mathbf{Y} - \mathbf{\Phi}_{\Lambda_k}\mathbf{\Theta}_{\Lambda_k}\right\|^2. \tag{30}$$

Minimizing $J(\mathbf{\Theta}_{\Lambda_k})$ should get the least squares estimates of the nonzero parameters. But there exist the unknown instrumental variables $z(t-i)$ and $w(t-i)$ and the unmeasurable noise term $v(t-i)$ in the information vector $\mathbf{\varphi}(t)$ of $\mathbf{\Phi}$; the parameter vector $\mathbf{\Theta}_{\Lambda_k}$ cannot be estimated by using the standard OMP methods in the CS theory. The method is to use their estimates $\widehat{z}(t-i)$, $\widehat{w}(t-i)$, and $\widehat{v}(t-i)$ to replace them; the computation of the estimates $\widehat{z}(t-i)$, $\widehat{w}(t-i)$, and $\widehat{v}(t-i)$ is as follows:

(i) Expand the orders n_b, n_c, and n_d in (1), (2), (8), and (9).

(ii) $\widehat{z}(t-i)$ is computed by replacing a_i and b_i with their estimates $\widehat{a}_{i,k}$ and $\widehat{b}_{i,k}$ in (8).

(iii) $\widehat{v}(t-i)$ is computed by replacing a_i, b_i, c_i, and d_i with their estimates $\widehat{a}_{i,k}$, $\widehat{b}_{i,k}$, $\widehat{c}_{i,k}$, and $\widehat{d}_{i,k}$ in (2) with (1) being inserted.

(iv) $\widehat{w}(t-i)$ is computed by replacing bi and di with their estimates $\widehat{b}_{i,k}$ and $\widehat{d}_{i,k}$ in (9). The results are

$$\widehat{z}_k(t) = -\sum_{i=1}^{L}\widehat{b}_{i,k}\widehat{z}_k(t-i) + \sum_{i=1}^{n_a}\widehat{a}_{i,k}y(t-1) + y(t),$$

$$\widehat{v}_k(t) = \sum_{i=o}^{n_a}\widehat{a}_{i,k}y(t-i) - \sum_{i=0}^{L}\widehat{b}_{i,k}\left[\sum_{i=1}^{L}\widehat{c}_{i,k}f_k[u(t-i)]\right]$$
$$- \sum_{i=1}^{L}\widehat{d}_{i,k}\widehat{v}(t-i), \quad \widehat{b}_{i,0} = 1$$

$$\widehat{w}_k(t) = -\sum_{i=1}^{L}\widehat{b}_{i,k}\widehat{w}_k(t-i) + \sum_{i=1}^{L}\widehat{d}_{i,k}\widehat{v}(t-i) + v(t),$$
$$\tag{31}$$

and the estimates of $\mathbf{\varphi}_v(t)$, $\mathbf{\varphi}(t)$, and $\mathbf{\Phi}$ at iteration k can be written as

$$\widehat{\varphi}_{v,k}(t) := \left[\widehat{z}(t-1) - \widehat{w}(t-1), \widehat{z}(t-2) - \widehat{w}(t-2), \ldots,\right.$$
$$\cdot \widehat{z}(t-L) - \widehat{w}(t-L), \widehat{v}(t-1), \widehat{v}(t-2), \ldots,$$
$$\left.\cdot \widehat{v}(t-L)\right]^T \in \mathbb{R}^{2L},$$

$$\widehat{\varphi}_k(t) := \begin{bmatrix} \varphi_u(t) \\ \widehat{\varphi}_{v,k}(t) \end{bmatrix} =: \left[-y(t-1), -y(t-2), \ldots,\right.$$
$$- y(t-n_a), f_1[u(t)], f_2[u(t)], \ldots,$$
$$\cdot f_L[u(t)], \widehat{z}_k(t-1) - \widehat{w}_k(t-1), \widehat{z}_k(t-2)$$
$$- \widehat{w}_k(t-2), \ldots, \widehat{z}_k(t-L) - \widehat{w}_k(t-L),$$
$$\left.\cdot v(t-1), v(t-2), \ldots, v(t-L)\right]^T,$$

$$\widehat{\mathbf{\Phi}}_k := \begin{bmatrix} \widehat{\varphi}_k^T(1) \\ \widehat{\varphi}_k^T(2) \\ \vdots \\ \widehat{\varphi}_k^T(m) \end{bmatrix} =: \left[\widehat{\phi}_{1,k}, \widehat{\phi}_{2,k}, \ldots, \widehat{\phi}_{N,k}\right].$$
$$\tag{32}$$

With replacing the above unknown variables with their estimates, the least squares estimate $\widehat{\mathbf{\Theta}}_{\Lambda_k}$ for the parameter vector $\mathbf{\Theta}_{\Lambda_k}$ at the k step is given:

$$\widehat{\mathbf{\Theta}}_{\Lambda_k} = \left(\widehat{\mathbf{\Phi}}_{\Lambda_k}^T \widehat{\mathbf{\Phi}}_{\Lambda_k}\right)^{-1}\widehat{\mathbf{\Phi}}_{\Lambda_k}^T \mathbf{Y}. \tag{33}$$

Remark 4. At the beginning, the IV-OMP algorithm with the estimates $\widehat{z}(t-1)$, $\widehat{w}(t-i)$, and $\widehat{v}(t-i)$ in the subinformation matrix $\widehat{\mathbf{\Phi}}_{\Lambda_k}$ causes an inaccurate support atom selecting. With the iteration k increasing, these estimated variables become more accurate, the misselected support atoms certainly unmeet the threshold requirement, and the corresponding element in the estimated parameter support set will be a small nonzero value. Thus, we set an appropriate small threshold ε to filter the parameter estimate $\widehat{\mathbf{\Theta}}_{\Lambda_k}$. If $|\widehat{\theta}_{h,k}| < \varepsilon$ (where $\widehat{\theta}_{h,k}$ is the hth element of $\widehat{\mathbf{\Theta}}_{\Lambda_K}$), eliminate $\widehat{\theta}_{h,k}$ from $\widehat{\mathbf{\Theta}}_{\Lambda_K}$ and the corresponding $\widehat{\phi}_{h,k}$ from $\widehat{\mathbf{\Phi}}_{\Lambda_k}$,

$$\widehat{\mathbf{\Theta}}_{\Lambda_{k\varepsilon}} = \left[\frac{\widehat{\mathbf{\Theta}}_{\Lambda_{k\varepsilon}}}{\widehat{\theta}_{h,k}} \notin \widehat{\mathbf{\Theta}}_{\Lambda_{k\varepsilon}}\right],$$
$$\widehat{\mathbf{\Phi}}_{\Lambda_{k\varepsilon}} = \left[\frac{\widehat{\mathbf{\Phi}}_{\Lambda_k}}{\widehat{\phi}_{h,k}} \notin \widehat{\mathbf{\Phi}}_{\Lambda_k}\right].$$
$$\tag{34}$$

Because $\widehat{\theta}_{h,k}$ is the hth element in $\widehat{\mathbf{\Theta}}_{\Lambda_k}$ and $\widehat{\phi}_{h,k}$ is the hth column in $\widehat{\mathbf{\Phi}}_{\Lambda_k}$, the expressions of eliminating $\widehat{\theta}_{h,k}$ from $\widehat{\mathbf{\Theta}}_{\Lambda_k}$ and $\widehat{\phi}_{h,k}$ from $\widehat{\mathbf{\Phi}}_{\Lambda_k}$ in MATLAB are

$$\widehat{\mathbf{\Theta}}_{\Lambda_{k\varepsilon}} : \widehat{\mathbf{\Theta}}_{\Lambda_k}(h) = [\,],$$
$$\widehat{\mathbf{\Phi}}_{\Lambda_{k\varepsilon}} : \widehat{\mathbf{\Phi}}_{\Lambda_k}(h) = [\,].$$
$$\tag{35}$$

Then we have

$$r_k = \mathbf{Y} - \widehat{\mathbf{\Phi}}_{\Lambda_{k\varepsilon}}\widehat{\mathbf{\Theta}}_{\Lambda_{k\varepsilon}}. \tag{36}$$

Briefly, the proposed IV-OMP algorithm can be implemented as in Algorithm 1.

IV-OMP for CS Recovery

Measurement matrix: $\mathbf{\Phi} = [\phi_1, \phi_2, \ldots, \phi_N]$,.

output vector: $\mathbf{Y}$, sparsity level: K,

parameter vector $\mathbf{\Theta} = [\theta_1, \theta_2, \ldots, \theta_N]^{\mathrm{T}}$.

Initialization: $k = 0, r_0 = \mathbf{Y}, \Lambda_0 = \varnothing, \widehat{\mathbf{\Theta}}_{\Lambda_0} = 0, \widehat{\mathbf{\Theta}}_k = 0$. **repeat**

   $k = k + 1$;

   Form $\widehat{\varphi}_k(t)$ by Equation (32);

$$\widehat{\mathbf{\Phi}}_k = \begin{bmatrix} \widehat{\varphi}_k^{\mathrm{T}}(1) \\ \widehat{\varphi}_k^{\mathrm{T}}(2) \\ \vdots \\ \widehat{\varphi}_k^{\mathrm{T}}(m) \end{bmatrix} = [\widehat{\phi}_{1,k}, \widehat{\phi}_{2,k}, \ldots, \widehat{\phi}_{N,k}];$$

   $\lambda_k = $ index of the highest amplitude component of $\widehat{\phi}_{i,k}^{\mathrm{T}} r_{k-1}$;

   $\Lambda_k = [\Lambda_{k-1}, \lambda_k]$;

   $\widehat{\mathbf{\Theta}}_{\Lambda_k} = [\widehat{\mathbf{\Phi}}_{\Lambda_k}^{\mathrm{T}} \widehat{\mathbf{\Phi}}_{\Lambda_k}]^{-1} \widehat{\mathbf{\Phi}}_{\Lambda_k}^{\mathrm{T}} \mathbf{Y}$;

   If $|\widehat{\theta}_{h,k}| < \varepsilon$, let $\widehat{\mathbf{\Theta}}_{\Lambda_k}(h) = [\,]$ to get $\widehat{\mathbf{\Theta}}_{\Lambda_{k\varepsilon}}$.

     and let $\widehat{\mathbf{\Phi}}_{\Lambda_k}(h) = [\,]$ to get $\widehat{\mathbf{\Phi}}_{\Lambda_{k\varepsilon}}$;

   $r_k = \mathbf{Y} - \widehat{\mathbf{\Phi}}_{\Lambda_{k\varepsilon}} \widehat{\mathbf{\Theta}}_{\Lambda_{k\varepsilon}}$;

   let $\widehat{\mathbf{\Theta}}_k(\Lambda_{k\varepsilon}) = \widehat{\mathbf{\Theta}}_{\Lambda_{k\varepsilon}}$ to get $\widehat{\mathbf{\Theta}}_k$;

   $\widehat{a}_{:,k} = \widehat{\mathbf{\Theta}}_k(1 : n_a), \widehat{c}_{:,k} = \widehat{\mathbf{\Theta}}_k(n_a + 1 : n_a + L)$,

   $\widehat{b}_{:,k} = \widehat{\mathbf{\Theta}}_k(n_a + L + 1 : n_a + 2L)$,

   $\widehat{d}_{:,k} = \widehat{\mathbf{\Theta}}_k(n_a + 2L + 1 : n_a + 3L)$;

$$\widehat{z}_k(t) = -\sum_{i=1}^{L} \widehat{b}_{i,k} \widehat{z}_k(t-i) + \sum_{i=1}^{n_a} \widehat{a}_{i,k} y(t-i) + y(t),$$

$$\widehat{v}_k(t) = \sum_{i=0}^{n_a} \widehat{a}_{i,k} y(t-i) - \sum_{i=0}^{L} \widehat{b}_{i,k} \left[\sum_{i=1}^{L} \widehat{c}_{i,k} f_k[u(t-i)] \right]$$

$$- \sum_{i=1}^{L} \widehat{d}_{i,k} \widehat{v}(t-i),$$

$$\widehat{w}_k(t) = -\sum_{i=1}^{L} \widehat{b}_{i,k} \widehat{w}_k(t-i) + \sum_{i=1}^{L} \widehat{d}_{i,k} \widehat{v}(t-i) + v(t);$$

until $k = $ a fixed positive integer $(k > K)$, s,.t. $\|r_k\| < \varepsilon$;.

Estimated parameter vector: $\widehat{\mathbf{\Theta}}_{\Lambda_{k\varepsilon}}$ and $\widehat{\mathbf{\Theta}}_k$ with $\widehat{\mathbf{\Theta}}_{\overline{\Lambda}_{k\varepsilon}} = 0$.

ALGORITHM 1: (IV-OMP algorithm).

5. Examples

Consider the following Hammerstein CARMA system

$$\begin{aligned}
A(z)y(t) &= B(z)x(t) + D(z)v(t), \\
x(t) &= -1.70u(t) + 0.80u^2(t), \\
A(z) &= 1 + 1.20z^{-1} + 0.60z^{-2}, \\
B(z) &= 1 + 0.12z^{-1} - 0.06z^{-2}, \\
D(z) &= 1 + 0.11z^{-1} - 0.10z^{-2}, \\
\mathbf{\Theta} &= \left[\underbrace{1.20, 0.60}_{n_a=2}, \underbrace{-1.70, 0.80, 0, 0, 0}_{L=5}, \underbrace{0.12, -0.06, 0, 0, 0}_{L}, \underbrace{0.11, -0.10, 0, 0, 0}_{L} \right]^T.
\end{aligned}$$

(37)

TABLE 1: the recursive parameter estimates and errors by using different algorithms.

| Algorithm	k	$\hat{a}_{1,k}$	$\hat{a}_{2,k}$	$\hat{c}_{1,k}$	$\hat{c}_{2,k}$	$\hat{c}_{3,k}$	$\hat{c}_{4,k}$	$\hat{c}_{5,k}$	$\hat{b}_{1,k}$	$\hat{b}_{2,k}$	$\hat{b}_{3,k}$	$\hat{b}_{4,k}$	$\hat{b}_{5,k}$	$\hat{d}_{1,k}$	$\hat{d}_{2,k}$	$\hat{d}_{3,k}$	$\hat{d}_{4,k}$	$\hat{d}_{5,k}$	δ (%)
	1	0.7717	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	87.52
	2	0.7100	0	0	0	0	0	−0.2365	0	0	0	0	0	0	0	0	0	0	88.71
	3	0.824	0	0	0	0	0	−0.2875	0	0	0	0	0	0	0	0.5002	0	0	90.55
	4	1.304	0.7615	0	0	0	0	−0.2439	0	0	0	0	0	0	0	0.0488	0	0	82.65
	5	1.331	0.8617	0	0	0	0.2315	−0.2468	0	0	0	0	0	0	0	−0.0846	0	0	83.87
	6	1.399	0.8219	0	0	0	0.1992	−0.2422	0	0.3025	0	0	0	0	0	−0.2355	0	0	85.66
	7	1.086	0.5736	−1.857	0	0	0.3747	0.0527	0	−0.0982	0	0	0	0	0	0.0080	0	0	40.00
IV-OMP	8	1.126	0.5786	−1.668	0.8567	0	−0.0220	0	0	−0.0314	0	0	0	0	0	0.0172	0	0	9.46
	9	1.124	0.5833	−1.693	0.8531	0	0	0	0	−0.0362	0	0	0	0	0	0.0126	0	0	9.27
	10	1.124	0.5855	−1.686	0.8181	0	0	0	0	−0.0370	0	0	0	0	0	0.0104	0	0	9.02
	11	1.125	0.5732	−1.697	0.8275	0	0	0	0	−0.0531	0	0	0	0	−0.0238	0.0126	0	0	8.58
	12	1.122	0.5754	−1.690	0.8295	0	0	0	0.0303	−0.0550	0	0	0	0	−0.0477	0.0038	0	0	7.54
	13	1.163	0.5926	−1.688	0.8159	0	0	0	0.0864	−0.0610	0	0	0	0.0593	−0.0686	−0.0090	0	0	3.48
	14	1.187	0.5913	−1.682	0.8096	0	0	0	0.0976	−0.0680	0	0	0	0.0858	−0.0756	−0.0016	0	0	2.18
	15	1.196	0.5973	−1.659	0.8035	0	0	0	0.1014	−0.0626	0	0	0	0.0929	−0.0726	0	0	0	2.46
True values		1.20	0.60	−1.70	0.80	0	0	0	0.12	−0.06	0	0	0	0.11	−0.10	0	0	0	0
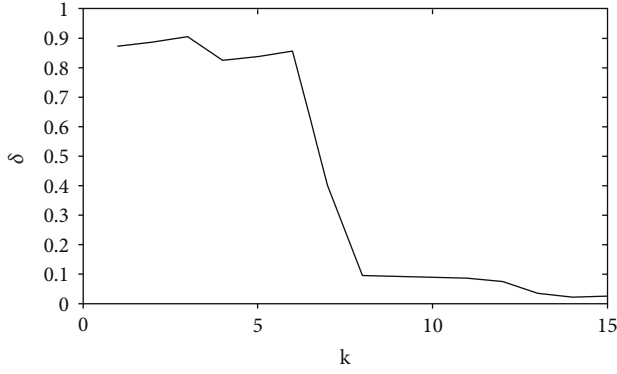
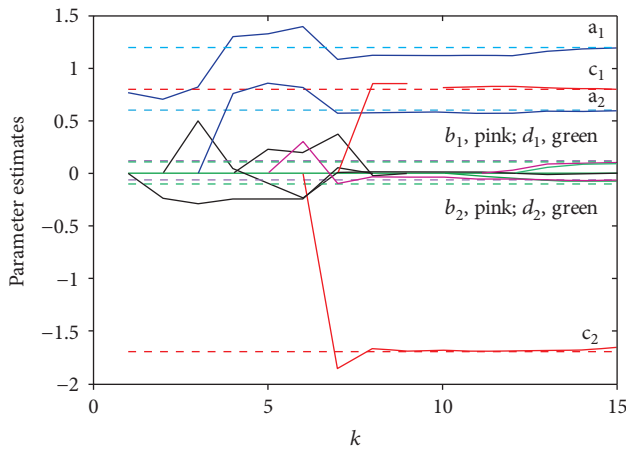Figure 2: The parameter estimation errors versus δ.

Figure 3: The parameter estimates.

In the simulation, the input $\{u(t)\}$ is taken as an uncorrelated persistently excited signal vector sequence with zero mean and unit variance, and $\{v(t)\}$ is taken as a white noise sequence with zero mean and variance $\sigma^2 = 0.10^2$. Applying the proposed IV-OMP algorithm to estimate the parameters and the orders of this system, the parameter estimates and their errors are shown in Table 1 and Figures 2 and 3. The relative error of the parameters is

$$\delta := \frac{\left\| \widehat{\Theta}_k - \Theta \right\|}{\left\| \Theta \right\|} \times 100\%. \qquad (38)$$

From Table 1 and Figures 2–3, we can get the following:

(1) The parameter estimation errors become (generally) smaller and smaller with the iteration k increasing.

(2) There exist 3 wrongly selected atoms corresponding to parameters $\widehat{c}_{4,k}$, $\widehat{c}_{5,k}$, and $\widehat{d}_{3,k}$ (black-colored lines in Figure 3). With the iteration k increasing, the wrongly chosen atoms are deleted, and the parameter estimation errors become smaller.

6. Conclusions

This paper is aimed at exploring the identification of both the parameters and orders of the Hammerstein CARMA system with limited sampling data. To solve the parameterizing difficulty caused by parameter coupling between the nonlinear part and the linear part in a Hammerstein system, firstly, by filtering the equation of the linear block with the coefficient function of the controlled term, we separate the parameter coupling between the linear block and the nonlinear block. Moreover, by using two instrumental variables, the Hammerstein system is parameterized into an autoregressive form. To achieve in simultaneously identifying parameters and orders and to promote the computational efficiency of the identification algorithm, an instrumental variable-based orthogonal matching pursuit (IV-OMP) optimization method of compressive sensing is extended to identify parameters and orders of the Hammerstein system. Simulation results illustrate that the investigated method is effective and has advantages of simplicity and efficiency. The proposed IV-OMP optimization method can be extended to the colored noise systems, the networked dynamic systems [42–46], and so on.

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.
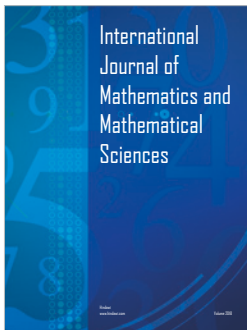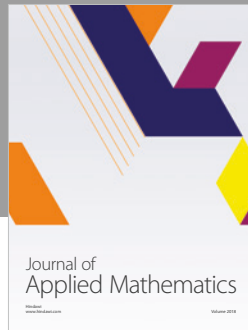
Acknowledgments

References

[1] G.-y. Chen, M. Gan, and G.-l. Chen, "Generalized exponential autoregressive models for nonlinear time series: stationarity, estimation and applications," *Information Sciences*, vol. 438, pp. 46–57, 2018.

[2] J. Na, J. Yang, X. Ren, and Y. Guo, "Robust adaptive estimation of nonlinear system with time-varying parameters," *International Journal of Adaptive Control and Signal Processing*, vol. 29, no. 8, pp. 1055–1072, 2015.

[3] A. Cordero, E. Gómez, and J. R. Torregrosa, "Efficient high-order iterative methods for solving nonlinear systems and their application on heat conduction problems," *Complexity*, vol. 2017, Article ID 6457532, 11 pages, 2017.

[4] M. Gan, C. L. P. Chen, G.-Y. Chen, and L. Chen, "On some separated algorithms for separable nonlinear least squares problems," *IEEE Transactions on Cybernetics*, vol. PP, pp. 1–9, 2017.

[5] J. Zambrano, J. Sanchis, J. M. Herrero, and M. Martínez, "WH-EA: an evolutionary algorithm for Wiener-Hammerstein system identification," *Complexity*, vol. 2018, Article ID 1753262, 17 pages, 2018.

[6] J. Na, J. Yang, X. Wu, and Y. Guo, "Robust adaptive parameter estimation of sinusoidal signals," *Automatica*, vol. 53, pp. 376–384, 2015.

[7] Y. Tang, Z. Han, Y. Wang, L. Zhang, and Q. Lian, "A changing forgetting factor RLS for online identification of nonlinear systems based on ELM-Hammerstein model," *Neural Computing and Applications*, vol. 28, Supplement 1, pp. 813–827, 2017.

[8] H. Salhi and S. Kamoun, "A recursive parametric estimation algorithm of multivariable nonlinear systems described by Hammerstein mathematical models," *Applied Mathematical Modelling*, vol. 39, no. 16, pp. 4951–4962, 2015.

[9] M. Jafari, M. Salimifard, and M. Dehghani, "Identification of multivariable nonlinear systems in the presence of colored noises using iterative hierarchical least squares algorithm," *ISA Transactions*, vol. 53, no. 4, pp. 1243–1252, 2014.

[10] J. Chen, Y. Liu, and X. Wang, "Recursive least squares algorithm for nonlinear dual-rate systems using missing-output estimation model," *Circuits, Systems, and Signal Processing*, vol. 36, no. 4, pp. 1406–1425, 2017.

[11] J. Zhang, F. Yu, and K. S. Chin, "Recursive identification of Hammerstein systems with dead-zone nonlinearity in the presence of bounded noise," *International Journal of Systems Science*, vol. 48, no. 11, pp. 2394–2404, 2017.

[12] P. Mattsson and T. Wigren, "Convergence analysis for recursive Hammerstein identification," *Automatica*, vol. 71, pp. 179–186, 2016.

[13] E. W. Bai and K. Li, "Convergence of the iterative algorithm for a general Hammerstein system identification," *Automatica*, vol. 46, no. 11, pp. 1891–1896, 2010.

[14] G. Li, C. Wen, W. X. Zheng, and G. Zhao, "Iterative identification of block-oriented nonlinear systems based on biconvex optimization," *Systems & Control Letters*, vol. 79, pp. 68–75, 2015.

[15] J. Vörös, "Identification of nonlinear cascade systems with output hysteresis based on the key term separation principle," *Applied Mathematical Modelling*, vol. 39, no. 18, pp. 5531–5539, 2015.

[16] J. Vörös, "Iterative identification of nonlinear dynamic systems with output backlash using three-block cascade models," *Nonlinear Dynamics*, vol. 79, no. 3, pp. 2187–2195, 2015.

[17] Y. Tan, R. Dong, and R. Li, "Recursive identification of sandwich systems with dead zone and application," *IEEE Transactions on Control Systems Technology*, vol. 17, no. 4, pp. 945–951, 2009.

[18] Z. Wang, Y. Wang, and Z. Ji, "Hierarchical recursive least squares algorithm for Hammerstein systems using the filtering method," *Nonlinear Dynamics*, vol. 77, no. 4, pp. 1773–1781, 2014.

[19] D. Wang, "Hierarchical parameter estimation for a class of MIMO Hammerstein systems based on the reframed models," *Applied Mathematics Letters*, vol. 57, no. 3, pp. 13–19, 2016.

[20] F. Ding, L. Xu, F. E. Alsaadi, and T. Hayat, "Iterative parameter identification for pseudo-linear systems with ARMA noise using the filtering technique," *IET Control Theory and Applications*, vol. 12, no. 7, pp. 892–899, 2018.

[21] 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.

[22] D. Wang, F. Ding, and Y. Chu, "Data filtering based recursive least squares algorithm for Hammerstein systems using the key-term separation principle," *Information Sciences*, vol. 222, no. 4, pp. 203–212, 2013.

[23] J. Li, W. X. Zheng, J. Gu, and L. Hua, "Parameter estimation algorithms for Hammerstein output error systems using Levenberg–Marquardt optimization method with varying interval measurements," *Journal of the Franklin Institute*, vol. 354, no. 1, pp. 316–331, 2017.

[24] D. Q. Wang, Z. Zhang, and J. Y. Yuan, "Maximum likelihood estimation method for dual-rate Hammerstein systems," *International Journal of Control, Automation and Systems*, vol. 15, no. 2, pp. 698–705, 2017.

[25] L. Ma and X. Liu, "Recursive maximum likelihood method for the identification of Hammerstein ARMAX system," *Applied Mathematical Modelling*, vol. 40, no. 13-14, pp. 6523–6535, 2016.

[26] S. Mete, S. Ozer, and H. Zorlu, "System identification using Hammerstein model optimized with differential evolution algorithm," *AEU - International Journal of Electronics and Communications*, vol. 70, no. 12, pp. 1667–1675, 2016.

[27] E. Cuevas, P. Díaz, O. Avalos, D. Zaldívar, and M. Pérez-Cisneros, "Nonlinear system identification based on ANFIS-Hammerstein model using gravitational search algorithm," *Applied Intelligence*, vol. 48, no. 1, pp. 182–203, 2018.

[28] M. Schoukens and K. Worden, "Evolutionary identification of block-structured systems," *Dynamics of Coupled Structures*, vol. 4, pp. 359–366, 2017.

[29] J. Chen, J. Li, and Y. Liu, "Gradient iterative algorithm for dual-rate nonlinear systems based on a novel particle filter," *Journal of the Franklin Institute*, vol. 354, no. 11, pp. 4425–4437, 2017.

[30] K. Krishnanathan, S. R. Anderson, S. A. Billings, and V. Kadirkamanathan, "Computational system identification of continuous-time nonlinear systems using approximate Bayesian computation," *International Journal of Systems Science*, vol. 47, no. 15, pp. 3537–3544, 2015.

[31] D. Wang, L. Li, Y. Ji, and Y. Yan, "Model recovery for Hammerstein systems using the auxiliary model based orthogonal matching pursuit method," *Applied Mathematical Modelling*, vol. 54, pp. 537–550, 2018.

[32] Y. Zhou, S. Kwong, H. Guo, X. Zhang, and Q. Zhang, "A two-phase evolutionary approach for compressive sensing reconstruction," *IEEE Transactions on Cybernetics*, vol. 47, no. 9, pp. 2651–2663, 2017.

[33] Q. Lin, B. Hu, Y. Tang et al., "A local search enhanced differential evolutionary algorithm for sparse recovery," *Applied Soft Computing*, vol. 57, pp. 144–163, 2017.

[34] M. Nasiri, B. Minaei, and Z. Sharifi, "Adjusting data sparsity problem using linear algebra and machine learning algorithm," *Applied Soft Computing*, vol. 61, pp. 1153–1159, 2017.

[35] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[36] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2210–2219, 2008.

[37] D. Q. Wang, Y. R. Yan, Y. J. Liu, and J. H. Ding, "Model recovery for Hammerstein systems using the hierarchical orthogonal matching pursuit method," *Journal of Computational and Applied Mathematics*, vol. 345, pp. 135–145, 2019.

[38] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[39] J. F. C. Mota, J. M. F. Xavier, P. M. Q. Aguiar, and M. Puschel, "Distributed basis pursuit," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1942–1956, 2012.

[40] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[41] Y. Mao, F. Ding, and Y. Liu, "Parameter estimation algorithms for Hammerstein time-delay systems based on the orthogonal matching pursuit scheme," *IET Signal Processing*, vol. 11, no. 3, pp. 265–274, 2017.

[42] X. Wang, F. Ding, A. Alsaedi, and T. Hayat, "Filtering based parameter estimation for observer canonical state space systems with colored noise," *Journal of the Franklin Institute*, vol. 354, no. 1, pp. 593–609, 2017.

[43] J. Ding, "The hierarchical iterative identification algorithm for multi-input-output-error systems with autoregressive noise," *Complexity*, vol. 2017, Article ID 5292894, 11 pages, 2017.

[44] J. Na, G. Herrmann, and K. Zhang, "Improving transient performance of adaptive control via a modified reference model and novel adaptation," *International Journal of Robust and Nonlinear Control*, vol. 27, no. 8, pp. 1351–1372, 2017.

[45] H. Liu and D. Wang, "Robust state estimation for wireless sensor networks with data-driven communication," *International Journal of Robust and Nonlinear Control*, vol. 27, no. 18, pp. 4622–4632, 2017.

[46] H. Liu and H. Yu, "Decentralized state estimation for a large-scale spatially interconnected system," *ISA Transactions*, vol. 74, pp. 67–76, 2018.