ICIC Express Letters ICIC International ©2016 ISSN 1881-803X
Volume **10**, Number **10**, October **2016** pp. 2305–2312

COMBINED HIERARCHICAL BASED RECURSIVE LEAST SQUARES ALGORITHM AND KALMAN FILTER FOR HAMMERSTEIN STATE SPACE SYSTEMS

Hong Wang, Wei Li, Shuqing Zhang, Xusheng Wang and Chao Ge

Institute of Information Engineering
North China University of Science and Technology
No. 46, Xinhua Road, Tangshan 063009, P. R. China
gechao365@126.com

Received March 2016; accepted June 2016

Abstract. *This paper is concerned with the parameter estimation of nonlinear dynamic Hammerstein systems in state space form. More specifically, a hierarchical based recursive least squares algorithm is employed to estimate the parameters of the proposed Hammerstein systems and a Kalman filter is derived to estimate the states of the canonical state space subsystem. Simulation examples demonstrate the effectiveness of the proposed method.*

Keywords: Hammerstein system, Parameter estimation, Hierarchical based recursive least squares algorithm, Kalman filter

1. **Introduction.** System modeling and parameter estimation have received much attention in recent years [1]. All physical systems are nonlinear and can be modelled by block-oriented nonlinear models. System identification has been widely used in many areas [2]. This paper focuses on the identification of Hammerstein model which consists of a nonlinear static block followed by a linear dynamic subsystem. The linear dynamic block has been assumed to be ARMAX [3], state space model [4], and so on.

Many approaches have been proposed to estimate parameters of Hammerstein systems. The iterative method was first proposed to estimate Hammerstein system in Narendra and Gallman [5], which is a very simple and efficient algorithm. Over-parameterization method is one direct method for the identification of Hammerstein models, in which the output of Hammerstein system is linear on the parameter space. However, the resulting parameter vector contains cross-products between the parameters of the non-linear part and the linear part, which increases the dimension of the parameter vector and leads to many redundant parameter estimates [6]. Vörös used the key term separation technique to deal with identification problems of Hammerstein systems with discontinuous nonlinearities [7]. For the key term separation technique, the output of system can be expressed as the linear regressive of all parameters which can be directly estimated. The blind identification method is presented by Bai to identify the Hammerstein models. By using the blind approach, Bai discussed the the Hammerstein and Hammerstein-Wiener model identification problems [8]. Hierarchical identification is inspired by the decomposition-coordination principle based hierarchical control for large-scale systems, which uses subsystem decomposition in identification [9].

In the area of state space system identification, Wills et al. addressed the problem of estimating parameters in state space model from observed frequency domain data [10]. Mercère and Bako presented a new subspace-based identification method to estimate the multi-inputs multi-outputs (MIMO) canonical state space model directly from data, which does not require the so-called observability/controllability indices [11]. Ding discussed a recursive least squares algorithm for a canonical state space dynamic system, in which the states of the linear system are estimated through the Kalman filter using the estimated parameters [12].

This paper presents a canonical state space based Hammerstein system. Differing from the work in [9], we estimate states of the linear subsystem through the Kalman filter [12]. In addition, a hierarchical based recursive least squares algorithm is adopted to generate the parameter estimates. By using the hierarchical identification principle, dimensions of the covariance matrices become smaller and the computation is more efficient, when compared with the over-parameterizations method. We frame our study in the identification of state space systems. The basic idea is to use the iterative technique to deal with the identification problem and to present a hierarchical gradient based iterative algorithm and a hierarchical least squares based iterative algorithm for a state space model. The calculated amount can be decreased compared with the over-parameter method which includes a high dimension parameter vector. In addition, the Kalman filter is successfully applied to states estimation of the nonlinear Hammerstein system.

2. **Problem Statement and Preliminaries.** Consider a Hammerstein system depicted in Figure 1, which consists of a static nonlinear block followed by a linear state space block. For the system, $u(t) \in \mathbb{R}$ and $y(t) \in \mathbb{R}$ are the system input and output variables of the Hammerstein system at time instant t; $\bar{u}(t) \in \mathbb{R}$ is the internal variable; $v(t) \in \mathbb{R}$ is a white noise with zero mean; $f(\cdot)$ is the polynomial function; $\boldsymbol{x}(t) = [x_1(t), x_2(t), \cdots, x_n(t)]^T \in \mathbb{R}^n$ is the state vector of the state space subsystem. $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{b} \in \mathbb{R}^n$, $\boldsymbol{c} \in \mathbb{R}^{1\times n}$ and $\boldsymbol{h} \in \mathbb{R}^m$ are the system parameter matrices/vectors. Assume that $(\boldsymbol{c}, \boldsymbol{A})$ is observable and $u(t) = 0$, $y(t) = 0$ for $t \leqslant 0$. $\boldsymbol{A}$, $\boldsymbol{b}$ and $\boldsymbol{h}$ are the unknown parameters to be estimated from the input-output sequence $\{u(t), y(t)\}$. The mathematical description of the above Hammerstein state space system can be written as:

$$\bar{u}(t) = f(u(t)) = \sum_{i=1}^{m} h_i f_i[u(t)] = \boldsymbol{f}^T(t)\boldsymbol{h}, \tag{1}$$

$$\boldsymbol{x}(t+1) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{b}\bar{u}(t), \tag{2}$$

$$y(t) = \boldsymbol{c}\boldsymbol{x}(t) + v(t), \tag{3}$$

where

$$\boldsymbol{A} := \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{n-1} & 0 & 0 & \cdots & 1 \\ -a_n & 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n\times n}, \quad \boldsymbol{f}(t) := \begin{bmatrix} f_1[u(t)] \\ f_2[u(t)] \\ \vdots \\ f_m[u(t)] \end{bmatrix} \in \mathbb{R}^n, \tag{4}$$

$$\boldsymbol{h} := \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{bmatrix} \in \mathbb{R}^m, \quad \boldsymbol{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in \mathbb{R}^n, \quad \boldsymbol{c} := [1, 0, \cdots, 0] \in \mathbb{R}^{1\times n}. \tag{5}$$

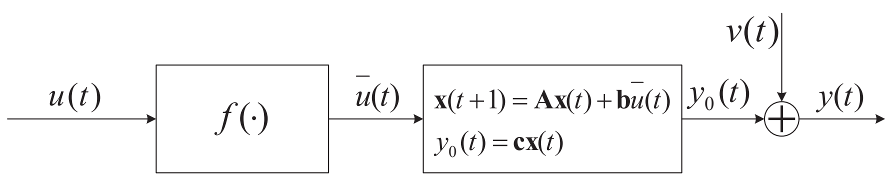

FIGURE 1. The Hammerstein state space system

From (2) and (3), we have

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \\ \vdots \\ x_{n-1}(t+1) \\ x_n(t+1) \end{bmatrix} = \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{n-1} & 0 & 0 & \cdots & 1 \\ -a_n & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix} \bar{u}(t), \quad (6)$$

$$y(t) = [1, 0, 0, \cdots, 0]\boldsymbol{x}(t) + v(t), \quad (7)$$

which can be written as

$$\begin{aligned} x_i(t+1) &= -a_i x_1(t) + x_{i+1}(t) + b_i \bar{u}(t), & (8) \\ x_n(t+1) &= -a_n x_1(t) + b_n \bar{u}(t), & (9) \\ y(t) &= x_1(t) + v(t). & (10) \end{aligned}$$

Multiplying (8) by z^{-i} and (9) by z^{-n} gives

$$\begin{aligned} x_i(t-i+1) &= -a_i x_1(t-i) + x_{i+1}(t-i) + b_i \bar{u}(t-i), & (11) \\ x_n(t-n+1) &= -a_n x_1(t-n) + b_n \bar{u}(t-n). & (12) \end{aligned}$$

Then summing for i from $i = 1$ to $i = n-1$ for (10) gives

$$x_1(t) = -\sum_{i=1}^{n-1} a_i x_1(t-i) + x_n(t-n+1) + \sum_{i=1}^{n-1} b_i \bar{u}(t-i). \quad (13)$$

Substituting (12) into (13) gives

$$\begin{aligned} x_1(t) &= -\sum_{i=1}^{n-1} a_i x_1(t-i) - a_n x_1(t-n) + b_n \bar{u}(t-n) + \sum_{i=1}^{n-1} b_i \bar{u}(t-i) \\ &= -\sum_{i=1}^{n} a_i x_1(t-i) + \sum_{i=1}^{n} b_i \bar{u}(t-i). & (14) \end{aligned}$$

From (10) and (14), we can get

$$y(t) = -\sum_{i=1}^{n} a_i x_1(t-i) + \sum_{i=1}^{n} b_i \bar{u}(t-i) + v(t). \quad (15)$$

According to [8], assume $\| \boldsymbol{b} \| = 1$ and the first nonzero entry of $\boldsymbol{b}$ is positive, i.e., $b_1 > 0$. Based on (1) and (15), the proposed Hammerstein nonlinear subsystem can be written as

$$y(t) = -\sum_{i=1}^{n} a_i x_1(t-i) + \sum_{j=1}^{n} b_j \sum_{k=1}^{m} h_k f_k[u(t-j)] + v(t) = \boldsymbol{\varphi}^T(t)\boldsymbol{a} + \boldsymbol{b}^T \boldsymbol{F}(t)\boldsymbol{h} + v(t), \quad (16)$$

where

$$\boldsymbol{\varphi}(t) = [-x_1(t-1), -x_1(t-2), \cdots, -x_1(t-n)]^T \in \mathbb{R}^n, \quad (17)$$

$$\boldsymbol{F}(t) := \begin{bmatrix} f_1[u(t-1)] & f_2[u(t-1)] & \cdots & f_m[u(t-1)] \\ f_1[u(t-2)] & f_2[u(t-2)] & \cdots & f_m[u(t-2)] \\ \vdots & \vdots & \vdots & \vdots \\ f_1[u(t-n)] & f_2[u(t-n)] & \cdots & f_m[u(t-n)] \end{bmatrix} \in \mathbb{R}^{n \times m}, \quad (18)$$

$$\boldsymbol{a} := [a_1, a_2, \cdots, a_n]^T \in \mathbb{R}^n. \quad (19)$$

This is the identification model of the Hammerstein state space model. In the next section, we would present a hierarchical based recursive least square identification algorithm and a Kalman filter to estimate the parameters and states of the system, which performs excellently.

3. **The Hierarchical Based Recursive Least Square Identification Algorithm.** The least squares method is a basic idea for system identification and parameter estimation. According to the input sequence $\{u(t)\}$ and output sequence $\{y(t)\}$ of a system which has white noise sequence $\{v(t)\}$ and parameter vector $\hat{\boldsymbol{\theta}}(t)$, the recursive least squares algorithm can be written as:

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \boldsymbol{L}(t)\left[y(t) - \boldsymbol{\varphi}^T(t)\hat{\boldsymbol{\theta}}(t-1)\right], \tag{20}$$

$$\boldsymbol{L}(t) = \boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)\left[1 + \boldsymbol{\varphi}^T(t)\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)\right]^{-1}, \tag{21}$$

$$\boldsymbol{P}(t) = \left[\boldsymbol{I} - \boldsymbol{L}(t)\boldsymbol{\varphi}^T(t)\right]\boldsymbol{P}(t-1), \quad \boldsymbol{P}(0) = p_0\boldsymbol{I}_n. \tag{22}$$

According to the least squares principle, the quadratic criterion function of the proposed Hammerstein system can be defined as

$$J(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{h}) := \sum_{i=1}^{t}\left[y(i) - \boldsymbol{\varphi}^T(i)\boldsymbol{a} + \boldsymbol{b}^T\boldsymbol{F}(i)\boldsymbol{h}\right]^2. \tag{23}$$

Since the information vector $\boldsymbol{\varphi}^T(t)$ contains the unknown state variables $x_n(t-i)$, here we replace the unknown state variables $x_n(t-i)$ in $\boldsymbol{\varphi}^T(t)$ with its estimates $\hat{x}_n(t-i)$, and replace $\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{x}(t)$, $\bar{u}(t)$ with their estimates $\hat{\boldsymbol{A}}$, $\hat{\boldsymbol{b}}$, $\hat{\boldsymbol{x}}(t)$, $\hat{\bar{u}}(t)$. For the bilinear parameter vectors in $J(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{h})$, we adopt the hierarchical based recursive least square identification method with decomposing the cost function into three linear $\left(J\left(\boldsymbol{a}, \hat{\boldsymbol{b}}(t-1), \hat{\boldsymbol{h}}(t-1)\right),\right.$ $J\left(\hat{\boldsymbol{a}}(t), \boldsymbol{b}, \hat{\boldsymbol{h}}(t-1)\right), \left.J\left(\hat{\boldsymbol{a}}(t), \hat{\boldsymbol{b}}(t), \boldsymbol{h}\right)\right)$ cost functions. The details can be described as [9]:

$$\hat{\boldsymbol{a}}(t) = \hat{\boldsymbol{a}}(t-1) + \boldsymbol{L}_1(t)\left[y(t) - \hat{\boldsymbol{\varphi}}^T(t)\hat{\boldsymbol{a}}(t-1) - \hat{\boldsymbol{b}}^T(t-1)\boldsymbol{F}(t)\hat{\boldsymbol{h}}(t-1)\right], \tag{24}$$

$$\boldsymbol{L}_1(t) = \boldsymbol{P}_1(t-1)\hat{\boldsymbol{\varphi}}(t)\left[1 + \hat{\boldsymbol{\varphi}}^T(t)\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\varphi}}(t)\right]^{-1}, \tag{25}$$

$$\boldsymbol{P}_1(t) = \left[\boldsymbol{I} - \boldsymbol{L}_1(t)\hat{\boldsymbol{\varphi}}^T(t)\right]\boldsymbol{P}_1(t-1), \quad \boldsymbol{P}_1(0) = p_0\boldsymbol{I}_n, \tag{26}$$

$$\hat{\boldsymbol{b}}(t) = \hat{\boldsymbol{b}}(t-1) + \boldsymbol{L}_2(t)\left[y(t) - \hat{\boldsymbol{\varphi}}^T(t)\hat{\boldsymbol{a}}(t) - \hat{\boldsymbol{b}}^T(t-1)\boldsymbol{F}(t)\hat{\boldsymbol{h}}(t-1)\right], \tag{27}$$

$$\boldsymbol{L}_2(t) = \boldsymbol{P}_2(t-1)\boldsymbol{F}(t)\hat{\boldsymbol{h}}(t-1)\left[1 + \left[\boldsymbol{F}(t)\hat{\boldsymbol{h}}(t-1)\right]^T\boldsymbol{P}_2(t-1)\boldsymbol{F}(t)\hat{\boldsymbol{h}}(t-1)\right]^{-1}, \tag{28}$$

$$\boldsymbol{P}_2(t) = \left[\boldsymbol{I} - \boldsymbol{L}_2(t)\left[\boldsymbol{F}(t)\hat{\boldsymbol{h}}(t-1)\right]^T\right]\boldsymbol{P}_2(t-1), \quad \boldsymbol{P}_2(0) = p_0\boldsymbol{I}_n, \tag{29}$$

$$\hat{\boldsymbol{h}}(t) = \hat{\boldsymbol{h}}(t-1) + \boldsymbol{L}_3(t)\left[y(t) - \hat{\boldsymbol{\varphi}}^T(t)\hat{\boldsymbol{a}}(t) - \hat{\boldsymbol{b}}^T(t)\boldsymbol{F}(t)\hat{\boldsymbol{h}}(t-1)\right], \tag{30}$$

$$\boldsymbol{L}_3(t) = \boldsymbol{P}_3(t-1)\left[\hat{\boldsymbol{b}}^T(t)\boldsymbol{F}(t)\right]^T\left[1 + \hat{\boldsymbol{b}}^T(t)\boldsymbol{F}(t)\boldsymbol{P}_3(t-1)\left[\hat{\boldsymbol{b}}^T(t)\boldsymbol{F}(t)\right]^T\right]^{-1}, \tag{31}$$

$$\boldsymbol{P}_3(t) = \left[\boldsymbol{I} - \boldsymbol{L}_3(t)\hat{\boldsymbol{b}}^T(t)\boldsymbol{F}(t)\right]\boldsymbol{P}_3(t-1), \quad \boldsymbol{P}_3(0) = p_0\boldsymbol{I}_n. \tag{32}$$

Here we estimate $\hat{\boldsymbol{x}}(t+1)$ by using the following Kalman filter which was introduced by Ding in [12], and the final form is shown as follows

$$\hat{\boldsymbol{x}}(t+1) = \hat{\boldsymbol{A}}(t)\hat{\boldsymbol{x}}(t) + \hat{\boldsymbol{b}}(t)\hat{\bar{u}}(t) + \boldsymbol{L}_0(t)[y(t) - \boldsymbol{c}\hat{\boldsymbol{x}}(t)], \quad \hat{\boldsymbol{x}}(1) = \mathbf{1}_n/p_0, \tag{33}$$

$$\boldsymbol{L}_0(t) = \hat{\boldsymbol{A}}(t)\boldsymbol{P}_0(t)\boldsymbol{c}^T\left[1 + \boldsymbol{c}\boldsymbol{P}_0(t)\boldsymbol{c}^T\right]^{-1}, \tag{34}$$

$$\boldsymbol{P}_0(t+1) = \hat{\boldsymbol{A}}(t)\boldsymbol{P}_0(t)\hat{\boldsymbol{A}}^T(t) - \boldsymbol{L}_0(t)\boldsymbol{c}\boldsymbol{P}_0(t)\hat{\boldsymbol{A}}^T(t), \quad \boldsymbol{P}_0(t) = \boldsymbol{I}_n, \tag{35}$$

where

$$\hat{\boldsymbol{A}}(t) := \begin{bmatrix} -\hat{a}_1(t) & 1 & 0 & \cdots & 0 \\ -\hat{a}_2(t) & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\hat{a}_{n-1}(t) & 0 & 0 & \cdots & 1 \\ -\hat{a}_n(t) & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \hat{\boldsymbol{b}}(t) := \begin{bmatrix} \hat{b}_1(t) \\ \hat{b}_2(t) \\ \vdots \\ \hat{b}_{n-1}(t) \\ \hat{b}_n(t) \end{bmatrix}, \quad \hat{u}(t) = \boldsymbol{f}^T(t)\hat{\boldsymbol{h}}(t). \tag{36}$$

The pseudo-code of computing the parameter estimation vectors $\hat{\boldsymbol{a}}(t)$, $\hat{\boldsymbol{b}}(t)$, $\hat{\boldsymbol{h}}(t)$ and the state estimation vector $\hat{\boldsymbol{x}}(t)$ is shown below in Algorithm 1.

Algorithm 1 The pseudo-code of computing the parameter estimation vectors $\hat{\boldsymbol{a}}(t)$, $\hat{\boldsymbol{b}}(t)$, $\hat{\boldsymbol{h}}(t)$ and the state estimation vector $\hat{\boldsymbol{x}}(t)$.

Initialize:

$\boldsymbol{a}_0 = \mathbf{1}_n/p_0$, $\boldsymbol{b}_0 = \mathbf{1}_n/p_0$, $\boldsymbol{h}_0 = \mathbf{1}_m/p_0$; $\boldsymbol{P}_1(0) = \boldsymbol{I}/p_0$, $\boldsymbol{P}_2(0) = \boldsymbol{I}/p_0$, $\boldsymbol{P}_3(0) = \boldsymbol{I}/p_0$; $p_0 = 10^6$; $\boldsymbol{x}(i) = 0$, $\boldsymbol{u}(i) = 0$, $\boldsymbol{y}(i) = 0$ for $i <= 0$; $\boldsymbol{x}(1)$ is a random vector.

Iterate:

for t=1:t **do**

Collect the input sequence $\{u(t)\}$ and output sequence $\{y(t)\}$, and form $\boldsymbol{\varphi}(t)$, $\boldsymbol{F}(t)$ according to (17), (18);

Based on the updated parameter vector $\hat{\boldsymbol{a}}(t-1)$, $\hat{\boldsymbol{b}}(t-1)$, and $\hat{\boldsymbol{h}}(t-1)$, compute the gain vector $\boldsymbol{L}_1(t)$ and the covariance matrix $\boldsymbol{P}_1(t)$ using (25) and (26), and update the parameter estimation vector $\hat{\boldsymbol{a}}(t)$ using (24);

Based on the updated parameter vector $\hat{\boldsymbol{a}}(t)$, $\hat{\boldsymbol{b}}(t-1)$, and $\hat{\boldsymbol{h}}(t-1)$, compute the gain vector $\boldsymbol{L}_2(t)$ and the covariance matrix $\boldsymbol{P}_2(t)$ using (28) and (29), and update the parameter estimation vector $\hat{\boldsymbol{b}}(t)$ using (27);

Based on the updated parameter vector $\hat{\boldsymbol{a}}(t)$, $\hat{\boldsymbol{b}}(t)$, and $\hat{\boldsymbol{h}}(t-1)$, compute the gain vector $\boldsymbol{L}_3(t)$ and the covariance matrix $\boldsymbol{P}_3(t)$ using (31) and (32), and update the parameter estimation vector $\hat{\boldsymbol{h}}(t)$ using (30);

Form $\hat{\boldsymbol{A}}(t)$ and $\hat{\boldsymbol{b}}(t)$ using (36);

Compute $\boldsymbol{L}_0(t)$ and $\boldsymbol{P}_0(t)$ and update the parameter estimation vector $\hat{\boldsymbol{x}}(t+1)$ using (33)-(35);

end for

4. **Example.** To further illustrate the performance of the proposed method, we consider a system as follows

$$x(t+1) = \begin{bmatrix} 0.340 & 1 \\ -0.429 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0.409 \\ 1.352 \end{bmatrix} u(t), \tag{37}$$

$$y(t) = [1, 0]x(t) + v(t), \quad u(t) = 1.288u(t) + 0.760u^2(t) + 0.100u^3(t). \tag{38}$$

Then the parameters vector can be defined as

$$\boldsymbol{\theta} = [a_1, a_2, b_1, b_2, h_1, h_2, h_3]^T = [0.340, -0.429, 0.409, 1.352, 1.288, 0.760, 0.100]^T. \tag{39}$$

The hierarchical based recursive parameter estimation error is defined by

$$\delta(t) := \sqrt{\frac{\| \hat{\boldsymbol{a}}(t) - \boldsymbol{a} \|^2 + \| \hat{\boldsymbol{b}}(t) - \boldsymbol{b} \|^2 + \| \hat{\boldsymbol{h}}(t) - \boldsymbol{h} \|^2}{\| \boldsymbol{a} \|^2 + \| \boldsymbol{b} \|^2 + \| \boldsymbol{h} \|^2}}. \tag{40}$$

To test the effectiveness of Algorithm 1, the input sequence is taken as an independent persistent excitation signal sequence with zero mean and unit variance, and the variance of the white noise sequence is set as $\sigma^2 = 0.1^2$ and $\sigma^2 = 0.5^2$, respectively. Then the parameter and state estimate errors are shown in Table 1, the parameter estimates of $\hat{\boldsymbol{\theta}}(t)$

versus t are shown in Figure 2 and Figure 3. For the parameters and states to be estimated, the convergence precision is high and the convergence speed is fast. The parameter

TABLE 1. The parameter estimates of $\hat{\boldsymbol{\theta}}(t)$ versus t

σ^2	t	a_1	a_2	b_1	b_2	h_1	h_2	h_3	$\delta(\%)$
0.1^2	500	0.31859	−0.35245	0.37393	1.38374	1.31552	0.72440	0.07515	4.96564
	1000	0.33296	−0.39074	0.39290	1.38132	1.29211	0.73115	0.08769	2.82863
	3000	0.33896	−0.41880	0.40642	1.38071	1.27356	0.73901	0.09668	1.87472
0.5^2	500	0.24813	−0.28556	0.35948	1.39329	1.34301	0.73802	0.07601	9.05393
	1000	0.28592	−0.34430	0.38112	1.37940	1.31537	0.74492	0.08245	5.33202
	3000	0.32315	−0.40365	0.39047	1.37757	1.28865	0.74732	0.09720	2.14670
	True values	0.340	−0.429	0.409	1.352	1.288	0.760	0.100	0

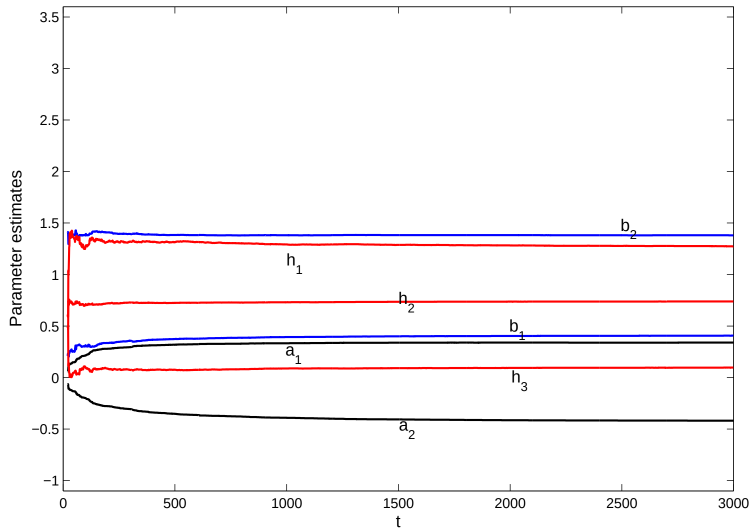

FIGURE 2. The parameter estimates of $\hat{\boldsymbol{\theta}}(t)$ versus t ($\sigma^2 = 0.1^2$)

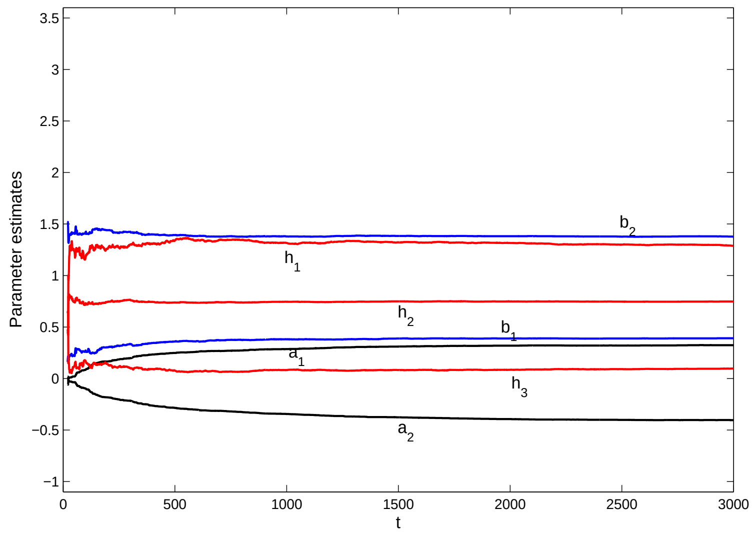

FIGURE 3. The parameter estimates of $\hat{\boldsymbol{\theta}}(t)$ versus t ($\sigma^2 = 0.5^2$)

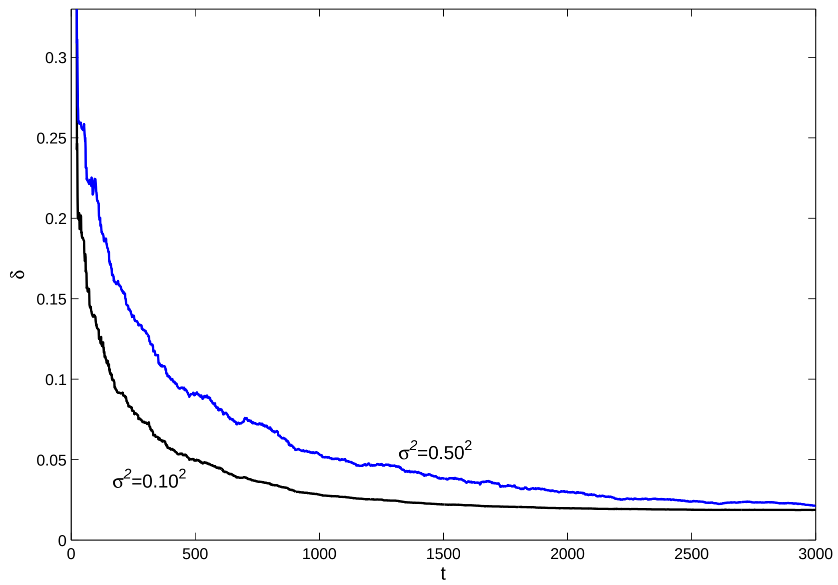

FIGURE 4. The parameter estimation errors δ versus t

estimation errors δ versus t are shown in Figure 4. From Figure 4, the estimation errors δ become smaller with the increasing of t. By using the hierarchical recursive least squares algorithm, the parameters and states of the Hammerstein system with canonical state space linear subsystem are estimated. Because the dimensions of the covariance matrices become smaller and the computation is more efficient when the hierarchical principle is used. The calculated amount can be decreased compared with the over-parameter method which includes a high dimension parameter vector. Therefore, we can conclude the proposed method is effective for nonlinear Hammerstein state space systems.

5. **Conclusion.** This paper proposes a Kalman filter based hierarchical recursive least squares algorithm to estimate the parameters and states of a Hammerstein system with canonical state space linear subsystem. By using the hierarchical identification principle, the calculated amount can be decreased compared with the over-parameter method. In addition, the Kalman filter is successfully applied to states estimation of the nonlinear Hammerstein system. The simulation results indicate that the proposed algorithm is effective. In further research, hierarchical recursive least squares algorithm can be used in the off line identification field, such as the waste water treatment process and other industry circle.

Acknowledgment. This work is partially supported by the National Natural Science Foundation of China (61503120), the Natural Science Foundation of Hebei Province (F2016209382). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] G. Bottegal and G. Pillonetto, Regularized spectrum estimation using stable spline kernels, *Automatica*, vol.49, no.11, pp.3199-3209, 2013.
[2] L. Wang, C. Li, G. Yin, L. Guo and C. Xu, State observability and observers of linear-time-invariant systems under irregular sampling and sensor limitations, *IEEE Trans. Automatic Control*, vol.56, no.11, pp.2639-2654, 2011.
[3] T. Baldacchino, S. Anderson and V. Kadirkamanathan, Computational system identification for Bayesian NARMAX modelling, *Automatica*, vol.49, no.9, pp.2641-2651, 2013.

[4] L. Bako, G. Mercère, S. Lecoeuche and M. Lovera, Recursive subspace identification of Hammerstein models based on least squares support vector machines, *IET Control Theory Applications*, vol.3, no.9, pp.1209-1216, 2009.

[5] K. Narendra and P. Gallman, Continuous time Hammerstein system identification, *IEEE Trans. Automatic Control*, vol.11, pp.546-550, 1966.

[6] F. Ding and T. Chen, Identification of Hammerstein nonlinear ARMAX systems, *Automatica*, vol.41, no.9, pp.1479-1489, 2005.

[7] J. Vörös, Recursive identification of Hammerstein systems with discontinuous nonlinearities containing dead-zones, *IEEE Trans. Automatic Control*, vol.48, no.12, pp.2203-2206, 2003.

[8] E. Bai, A blind approach to the Hammerstein-Wiener model identification, *Automatica*, vol.38, no.6, pp.967-979, 2002.

[9] D. Wang, F. Ding and X. Liu, Least squares algorithm for an input nonlinear system with a dynamic subspace state space model, *Nonlinear Dynamics*, vol.75, no.1, pp.49-61, 2014.

[10] A. Wills, B. Ninness and S. Gibson, Maximum likelihood estimation of state space models from frequency domain data, *IEEE Trans. Automatic Control*, vol.54, no.1, pp.19-33, 2009.

[11] G. Mercère and L. Bako, Parameterization and identification of multivariable state-space systems: A canonical approach, *Automatica*, vol.47, no.8, pp.1547-1555, 2011.

[12] F. Ding, Combined state and least squares parameter estimation algorithms for dynamic systems, *Applied Mathematical Modelling*, vol.38, no.1, pp.403-412, 2014.