Applied Mathematics and Computation 413 (2022) 126620

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

System identification of hammerstein models by using backward shift algorithm☆

Wen Mi[a], Tao Qian[b]

[a] School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China
[b] Macau Institute of Systems Engineering, Macau University of Science and Technology, Macao, China

ARTICLE INFO

Article history:
Received 21 May 2021
Revised 31 July 2021
Accepted 16 August 2021
Available online 31 August 2021

Keywords:
Nonlinear system
Backward shift
Hammerstein system
System identification
Transfer function

ABSTRACT

In this paper, a new identification method for discrete-time Hammerstein systems is proposed. The method is a joint use of discrete Fourier transform, backward shift method, and the least squares method. The frequency responses are obtained with sampled input and output data in the time domain through discrete Fourier transform. It is followed by the backward shift algorithm that was originally developed for estimating poles of linear time-invariant systems. After poles of linear subsystem are estimated, coefficients of linear and nonlinear subsystems are respectively determined by using the least squares (LS) method. The robustness of the backward shift algorithm guarantees the effectiveness of the proposed algorithm. Simulation results show that the poles of linear subsystem are well located. Thus, it is practical to identify discrete Hammerstein systems.

1. Introduction

A special kind of block-oriented models are Hammerstein models, they are capable of modeling the kind of nonlinear systems whose nonlinear parts (static without memory) are followed by their linear parts (dynamic system with memory). The identification of Hammerstein models is a challenging topic and has been attracting more and more attentions [3,6,12,39]. Researchers have proposed many methods for identification of Hammerstein models. Among them, there are blind method [3], stochastic methods [13,30–32], recursive methods [10,28,34,35,39], key term separation technique [16], nonparametric approaches [26,33], Hierarchical parameter estimation algorithm [4,7–9,11,17,29,36–38] and so on. All methods have their advantages on some conditions and are efficient to some special kind of Hammerstein systems. However, almost all methods have one thing in common, that the parameters of linear parts are estimated simultaneously. For numerical algorithms, they would obtain more accurate numerical results when there are less unknown parameters to estimate. If the parameters of linear parts can be estimated separatedly, then it reduces the computational burden, hence it is helpful to the identification of Hammerstein systems.

Frequency-domain identification is one of the two branches (time-domain and frequency domain) in system identification. Its advantages have been attracting more and more researchers not only for linear systems [1,24] but also nonlinear systems [2,32]. In [2], by using the sinusoidal inputs, there are many advantages for identification of continuous-time Hammerstein systems. First of all, the output signals could be expressed in a Fourier series expansion, whose Fourier coefficients

☆ The research was supported by Department of Guangzhou Science and Technology under Grant 202002030472 and the Science and Technology Development Fund, Macau SAR (File no. FDCT079/2016/A2 and FDCT0123/2018/A3€ ©.

E-mail addresses: mathmw@uestc.edu.cn (W. Mi), tqian@must.edu.mo (T. Qian).

https://doi.org/10.1016/j.amc.2021.126620
0096-3003/

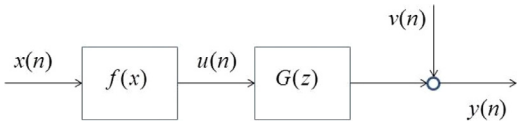

Fig. 1. A Hammerstein system.

are invariable to different frequencies of input sinusoid signals. Second, the frequency responses of linear subsystems can be estimated through a point estimate by using input-output signals of entire systems. Referring to these results, a similar property will be verified for discrete-time systems in this paper. After that, a new method to get estimations of Hammerstein systems is proposed, in which we look into the problem through estimating poles of linear subsystem. As a foundation, it is shown in [2] that a defined point estimate is identical with the sum of the true frequency response and an error, which is an average of finite time Fourier transform of the output noise. Then our basic idea is that we first get approximating frequency responses to linear subsystems with sampled data, then the poles of linear subsystem are estimated by using backward shift algorithm. This is a quite different framework from existing identification methods where the parameters of the linear part are not separated but generally estimated simultaneously. As a consequence, it will rise the numerical efficiency.

The poles of a linear system play a very important role in system identification. On one hand, the poles affect the performance of a system. One the other hand, the generalized rational orthornomal bases are useful tools in identification of linear systems [15,18,19,21,23], where the basis functions are used to constructed linear-in-parameters models. The convergence rate of constructed models are determined by their poles of basis functions. So information of poles is quite important to these problems. The backward shift algorithm (BSA), which was developed in [20], is a useful method in estimating true poles including their particular values and multiplicities. In the processes of backward shift algorithm, one may only need decades of measurements in frequency domain. In the present study of identifying Hammerstein systems, we pay our attention to the poles of linear subsystems at the beginning. The backward shift algorithm to estimate the poles will be also generalized. Next, different from [20], before the modified BSA is applied, we start with measurements sampled in the time domain. After the mission of finding poles is complete, models for linear subsystems are constructed by using the obtained poles for the basis functions. At last, coefficients of the linear and nonlinear parts are respectively determined by using the least squares (LS) method. The main contributions of this study include:

- The backward shift algorithm is modified and generalized. In particular, we release the equal-spacing restriction for sampling frequency spaces in the modified algorithm.
- The modified backward shift algorithm is applied to identification of Hammerstein systems. Poles of linear subsystems are efficiently estimated by using the modified backward shift algorithm.

The outline of this paper is given as follows. In Section 2, the problem settings are provided. We give a summary of the backward shift algorithm in Section 3. Section 4 is to propose the identification processes of a Hammerstein system by using our DFT-BSA-LS method. After that, an example is given in Section 5 to illustrate the proposed idea. Conclusions are drawn in the last section.

2. Problem formulation

A discrete-time Hammerstein system, shown in Fig. 1, consists of a dynamical linear subsystem $G(z)$ and a static nonlinear subsystem $f(x)$. $G(z)$ is the transfer function of its linear subsystem, $y(n)$ is the output signal with input $x(n)$, it is noised by $v(n)$. We set $y_0(n)$ as the true output without noise. $u(n)$ is the internal signal, it can be treated as either output of the static nonlinear subsystem $f(x)$ or input of the linear subsystem $G(z)$.

In engineering, the transfer function of a linear system $G(z)$ is defined by the series

$$G(z) = \sum_{n=-\infty}^{\infty} h(n) z^{-n}, \tag{1}$$

where $h(n)$ is the impulse response of system $G(z)$. It can also be defined by replacing z with z^{-1} in (1), then the transfer function of a stable and casual linear discrete system is analytic in the unit disc in the complex plane. In this case, the study is proceeded in the unit disc. Here are some assumptions about the considered Hammerstein system in this paper:

- **Assumption** 1. The discrete linear subsystem $G(z)$ is a rational function, that is to say, $G(z)$ is a stable and casual system and its poles are not in the closed unit disc.
- **Assumption** 2. The nonlinear static subsystem $f(x)$ is assumed to be a polynomial with known maximal order p.
- **Assumption** 3. The noise $v(n)$ is Gassian, its a stationary random process with zero mean and zero mean and finite variance σ, $0 < \sigma < +\infty$.

Generally speaking, a Hammerstein system $\{f(x), G(z)\}$ is unidentifiable, because either linear subsystem $G(z)$ or nonlinear subsystem $f(x)$ are not unique. For any different nonzero constant c, the pairs of $\{cf(x), G(z)/c\}$ give rise to the same

product under identical input signals. Some studies tried to estimate true linear and nonlinear subsystems, that do not make sense and nor are they necessary neither. Many ways can be taken to avoid this problem, such as estimating a block at first, fixing a norm of $f(x)$ or $G(z)$ [14], or setting the first order coefficient of the Fourier series of the output $f(A\cos(w_l t))$ being equal 1, etc[2].

For **Assumption** 1, $G(z)$ is a rational function, it also can be equally rewritten as

$$G(z) = \sum_{k=1}^{m_1} \frac{\lambda_{1_k}}{(1-a_1 z)^k} + \ldots + \sum_{k=1}^{m_\alpha} \frac{\lambda_{\alpha_k}}{(1-a_\alpha z)^k}, \tag{2}$$

where $|a_i| < 1$, $K = m_1 + \ldots + m_\alpha$ is the order of $G(z)$, (for simplicity a_k is just called poles of system in the left text). If a_i is the single pole, then the index m_i is 1.

The algorithm starts with sets of measurements in time domain, and the excitation signals are set to be

$$x^{<l>}(n) = A\cos(w_l n), \tag{3}$$

where $w_l (l = 1, 2, \ldots, L)$ are frequencies. In each signal, there are N sampled data. Because of the structure of Hammerstein model, the internal signal $u(n) = f(x(n))$ and output signal $y(n)$ are periodic under input signal $x(n)$. This study begins with the following sampled data set: there are L inputs

$$\{x^{<1>}(n)\}_{n=1}^N, \ldots, \{x^{<l>}(n)\}_{n=1}^N \ldots, \{x^{<L>}(n)\}_{n=1}^N$$

and L noised outputs

$$\{y^{<1>}(n)\}_{n=1}^N, \ldots, \{y^{<l>}(n)\}_{n=1}^N \ldots, \{y^{<L>}(n)\}_{n=1}^N$$

that are given by

$$y^{<l>}(n) = y_0^{<l>}(n) + v^{<l>}(n) \quad l = 1, 2, \ldots, L.$$

$y_0^{<l>}(n)$ stand for the true values of signals and $v^{<l>}(n)$ are the noises. For computation convenience, we choose that $N = mL$, where m is a positive integer. Other types of input could also be used, such as multisine [24].

The aim of this study is to identify a pair of $\{cG(z), \frac{f(x)}{c}\}$ based on the data sets, but not to find $G(z)$, $f(z)$ and c themselves, where c is a nonzero number. Since the coefficients of linear and nonlinear parts can not be uniquely fixed, while the poles of $G(z)$ are unique. $cG(z)$ and $G(z)$ share the same poles, then the poles of $G(z)$ plays a key role in identification of a Hammerstein system. This is why one should consider the poles at first. So we will start our identification process with estimating poles of $G(z)$, where backward shift algorithm is used. In the next section, we briefly introduce the backward shift algorithm.

3. The backward shift algorithm

The backward shift algorithm is developed based on a famous mathematical tool: the backward shift operator which is defined by

$$\mathbf{B}(g)(z) = \frac{g(z) - g(0)}{z}, \tag{4}$$

and $g(z)$ is a function in Hardy spaces. This operator is continuous for Hardy spaces functions $H^p(\mathbb{D})$ $p > 1$ [5]. For a closed subspace $\mathfrak{M}$ in $H^p(\mathbb{D})$, it is also a backward shit invariant subspace for $\mathbf{B}$, e.g., for any function $g(z) \in \mathfrak{M}$, $\mathbf{B}(g) \in \mathfrak{M}$.

Set $e_a(z) = \frac{1}{1-az}$, it is a rational wavelet basis function with parameter a in the unit disc. For a K-sequence $\{a_1, \ldots, a_K\}$, a_k could repeat, we have multiple analytic wavelets, $\{\tilde{e}_{a_1}, \ldots, \tilde{e}_{a_K}\}$, where

$$\tilde{e}_{a_k} = \frac{1}{(1-a_k z)^{l_k}},$$

and l_k is the repeating time of a_k in $(a_1, \ldots, a_k)$.

The set of multiple analytic wavelets $\{\tilde{e}_{a_1}, \ldots, \tilde{e}_{a_K}\}$ could generate a closed subspace of $H^p(\mathbb{D})$, in fact, it is shown in [15,25], that

$$span\{\mathbb{B}_1(z), \ldots, \mathbb{B}_K(z)\} = span\{\tilde{e}_{a_1}, \ldots, \tilde{e}_{a_K}\}$$

where the rational orthomormal basis functions $\mathbb{B}_k(z)$ are defined by

$$\mathbb{B}_k(z) = \frac{\sqrt{1-|a_k|^2}}{1-a_k z} \prod_{l=1}^{k-1} \frac{z-\overline{a_l}}{1-a_l z} \tag{5}$$

with $\mathbb{B}_1(z) = \frac{\sqrt{1-|a_1|^2}}{1-a_1 z}$, $\overline{a_k}$ is the conjugate.

Given a function in form (2), it can be reformed in linear combination of rational orthogonal basis functions (5), i.e.,

$$g(z) = \sum_{k=1}^{K} \vartheta_k \mathbb{B}_k(z). \tag{6}$$

We can notice that the space $X_{a_1,\ldots,a_K}$ spanned by $(\tilde{e}_{a_1},\ldots,\tilde{e}_{a_K})$ is a closed subspace in Hardy spaces, thus, for $g(z)$ in form of (2), we can see not only $g(z)$ belongs to $X_{a_1,\ldots,a_K}$, but also the function after taking backward shift algorithm $\mathbf{B}(g)(z)$ belongs to $X_{a_1,\ldots,a_K}$. In fact, we have

$$\begin{aligned}\mathbf{B}(e_{a_k})(z) &= \frac{e_{a_k}(z)-e_{a_k}(0)}{z}\\ &= a_k e_{a_k}(z).\end{aligned}$$

And we have a result shown in [20] that

Theorem 3.1. *For a rational function $g(z)$ given in form (2), suppose the order is K, after taking k times backward shift operator one obtain $\mathbf{B}^k(g)(z)$, then there is a unique sequence $\{\mu_1,\mu_2,\ldots,\mu_K\}$ such that*

$$\mathbf{B}^K(g)(z)+\mu_K\mathbf{B}^{K-1}(g)(z)+\ldots+\mu_1 g((z)=0, \tag{7}$$

where μ_k are not all zeros. Furthermore, zeros of corresponding algebraic equation

$$x^K+\mu_K x^{K-1}+\ldots+\mu_2 x+\mu_1=0, \tag{8}$$

are exactly the parameters $\{a_1,a_2,\ldots,a_K\}$, including the multiplicities.

Based on the relationship given in (7) and (8), the so called *Backward Shift Algorithm* (BSA) was developed in [20]. The direct purpose is to estimate parameters $\{a_k\}$ (including multiplicities) by using frequency-domain measurements $\{E_l\}$, where $E_l=g(e^{jw_l})+V_l$, each $g(e^{jw_l})$ is the true value at w_l and V_l is the noise. The algorithm can be briefly stepped as:

Algorithm 1

- Step 1. Collect data set of frequency responses $\{E_l\}$;
- Step 2. Take backward shift operator to $\{E_l\}$ and obtain data sets $\{\mathbf{B}(\{E_l\})\}$, $\{\mathbf{B}^2(\{E_l\})\}$, ..., $\{\mathbf{B}^M(\{E_l\})\}$;
- Step 3. Find $\mu_1,\mu_2,\ldots,\mu_K$ from equation (7);
- Step 4. Find $\{a_k\}$ by solving equation (8).

It is shown in [20] that only dozens of frequency-domain measurements are enough to get true poles in case of no noise.

After poles $\{a_k\}$ are obtained, then e_{a_k}(or $\mathbb{B}_k(z)$) are fixed, their coefficients can be estimated quickly by using the least squares method. Due to the particular structure of Hammerstein systems, it is reasonable to consider the application of the backward shift algorithm in their identification.

4. Identification of hammerstein systems with BSA

4.1. Modification of BSA

In the algorithm of BSA, given linear subsystem $G(z)$, an approximation to $G(0)$ is taken by using sampled data as follows:

$$\begin{aligned}\mathbf{B}(G)(e^{jw_l}) &= \frac{G(e^{jw_l})-G(0)}{e^{jw_l}} \\ &= \frac{G(e^{jw_l})-\frac{1}{2\pi}\int_0^{2\pi}G(e^{jw})dw}{e^{jw_l}} \\ &\approx \frac{G(e^{jw_l})-\frac{1}{L}\sum_{k=1}^{L}G(e^{jw_k})}{e^{jw_l}}.\end{aligned} \tag{9}$$

This approximation is simple, and truly, it is a special case of the following approximation,

$$\mathbf{B}(G)(e^{jw_l}) \approx \frac{G(e^{jw_l})-\frac{1}{2\pi}\int_0^{2\pi}\sum G(e^{jw_l})\chi(\cdot)dw}{e^{jw_l}}, \tag{10}$$

in which $\chi(\cdot)$ is the indicator function. So one can use the extended formula instead of the old approximating formula if no equal spacing is used. And this is more helpful for algorithm and experimental designing.

Remark 4.1. In fact, according to the theory of holomorphic functions, a function given in 2 can be uniquely determined by a set of function values on a convergent sequence in its analytic region [27].

Even though the estimation is replaced by (10), the backward shift algorithm (**Algorithm 1**) still keeps its efficiency and robustness. We have

Theorem 4.1. *Suppose we have a set of noised data $\{E_l\}_{l=1}^{L}$ (i.e, $E_l=G(e^{jw_l})+V_l$) and each V_l is bounded as $|V_l|<\varepsilon$. If $\mu^\star$ is the true solution and μ_L is the estimated solution of (7) by BSA with (10), then there is*

$$\mu_L\to\mu^\star \quad (L\to\infty,\ \varepsilon\to 0). \tag{11}$$

For the proof, it is analogous to the process of Theorem 4 in [20] by replacing (9) with (10), the detail is given in the Appendix Section. From formula (8), we can see when μ_L are getting true, and the estimated poles approach to the true values of $\{a_k\}$.

4.2. Estimation of linear subsystem

Now we are going to show the identification of Hammerstein systems, where the backward shift algorithm is mainly used in the identification of linear subsystem. On considering the initial condition starting from input and output measurements $\{x^{<l>}(n)\}_{n=1}^{N}$, $\{y^{<l>}(n)\}_{n=1}^{N}$ $(l = 1, 2 \dots L)$ in the time domain, while the BSA is a frequency-responses driven algorithm (**Algorithm 1**), then an estimation of frequency responses is necessary before proceeding BSA.

In this paper, discrete Fourier transform (DFT) is used to get the estimation of frequency responses of linear subsystem. By computation, there are:

$$X(e^{jw_l}) = \frac{A\sqrt{N}}{2} \tag{12}$$

and

$$Y(e^{jw_l}) = \frac{c\sqrt{N}G(e^{jw_l})}{2} + V(e^{jw_l}), \tag{13}$$

where c is a nonzero constant according to [2] and $V(e^{jw_l})$ is DFT of noise. From equations (12) and (13), one can define an estimator as

$$\widetilde{G}(e^{jw}) = A\frac{Y(e^{jw})}{X(e^{jw})}. \tag{14}$$

Consequently, we obtain estimates of frequency responses of the linear subsystem at each frequency $w = w_l$, i.e.,

$$\widetilde{G}(e^{jw_l}) = cG(e^{jw_l}) + \frac{2V(e^{jw_l})}{\sqrt{N}}. \tag{15}$$

A result about convergence of (15) guarantees this estimation:

Theorem 4.2. *Given a Hammerstein system in Fig. 1 with assumptions 1–3, then the estimation of frequency responses for the linear subsystem $G(z)$ by (15) converges to $cG(e^{jw_l})$ uniformly in l,viz.,*

$$\widetilde{G}(e^{jw_l}) \to cG(e^{jw_l})$$

uniformly in l in probability as $N \to \infty$.

Proof. Denote by $\mathbf{E}(\bullet)$ the expectation for $\bullet$. According to [24], there are

$$\mathbf{E}V(e^{jw_l}) = 0 \tag{16}$$

and

$$\mathbf{E}\{\frac{|V(e^{jw_l})|^2}{N}\} = O(\frac{1}{\sqrt{N}}). \tag{17}$$

Then the above equations lead to

$$\mathbf{E}(\widetilde{G}(e^{jw_l}) - cG(e^{jw_l})) = 0 \tag{18}$$

and

$$\mathbf{E}|\widetilde{G}(e^{jw_l}) - cG(e^{jw_l})|^2 = O(\frac{1}{\sqrt{N}}) \to 0, \ \ (N \to \infty) \tag{19}$$

uniformly in w_l, respectively. The proof is complete.

After these preparations, the backward shift algorithm could be now used to estimate poles $\{a_k\}$ and linear subsystem for the considered Hammerstein system. The complete algorithm is shown as follows. □

Algorithm 2

- Step 1. Take samples of input and output data $\{x^{<l>}(n)\}_{n=1}^{N}$, $\{y^{<l>}(n)\}_{n=1}^{N}$, under the input $x(n) = A\cos(w_l n)$, $(l = 1, 2 \dots L)$.
- Step 2. Estimate frequency responses with estimator $\widetilde{G}(e^{jw}) = A\frac{Y(e^{jw})}{X(e^{jw})}$ at the frequency w_l, and obtain $\{(w_l, \widetilde{G}(e^{jw_l}))\}$.
- Step 3. Get the approximated poles $\{a_k\}$ by applying BSA with the result in step 2, and.
- Step 4. Get approximation to $cG(z)$, say,

$$\widetilde{G}(z) = \sum_{k=1}^{K} \vartheta_k \mathbb{B}_k(z), \tag{20}$$

in which the generalized rational orthogonal basis functions $\mathbb{B}_k(z)$ are constructed by using $\{a_k\}$ and the coefficients ϑ_k's are estimated by the least squares method.

Note that the linear subsystem that one approaches is $cG(z)$ but not $G(z)$.

4.3. Estimation of nonlinear part

When the linear part is fixed, next mission is to estimate the nonlinear part. According to Assumption 2, the unknown nonlinearity is assumed to be a polynomial. Although this assumption is simple, it is reasonable for smooth and continuous nonlinearities. Then precisely, the task is to estimate the coefficients of

$$f(x) = \beta_0 + \beta_1 x + \ldots + \beta_p x^p, \tag{21}$$

where p is the order of the polynomial f, we are now to find $\{\beta_0, \beta_1, \ldots, \beta_p\}$ such that $f(x)$ matches the inputs and outputs of the Hammerstein system the best, on coincidence with the fixed linear part $\widetilde{G}(z)$.

Since the obtained linear subsystem $\widetilde{G}(z)$ is a rational function, without losing generality, it can be set as

$$\widetilde{G}(z) = \frac{b_{K-1}z^{K-1} + b_{K-2}z^{K-2} + \ldots + b_1 z + b_0}{z^K + d_{K-1}z^{K-1} + \ldots + d_1 z + d_0} \tag{22}$$

with order K, b_k and d_k being real coefficients.

For the lth input signal at frequency w_l, $x^{<l>}(n) = A\cos(w_l n)$, then the middle signal (output of $f(x)$) is

$$u^{<l>}(n) = \beta_0 + \beta_1 A\cos(w_{ln}) + \ldots + \beta_p[A\cos(w_{ln})]^p, \tag{23}$$

and the output $y^{<l>}(n)$ is given by

$$y^{<l>}(n) = (\widetilde{G} * u^{<l>})(n), \tag{24}$$

being the convolution of $\{u^{<l>}(n)\}$ and impulse response of $\widetilde{G}(z)$. Equation (24) also can be rewritten in the form of a differencial equation, that is

$$\begin{aligned}&(z^K + d_{K-1}z^{K-1} + \ldots + d_1 z + d_0)\cdot y^{<l>}(n)\\ &= (b_{K-1}z^{K-1} + b_{K-2}z^{K-2} + \ldots + b_1 z + b_0)\cdot u^{<l>}(n),\end{aligned}$$

where z is seen as a backward operator to the sampled series in time domain. The above equations leads to

$$\begin{aligned}&y^{<l>}(n-K)\ldots + d_1 y^{<l>}(n-1) + d_0 y^{<l>}(n)\\ &= b_{K-1}u^{<l>}(n-K+1)\ldots + \ldots + b_0 u^{<l>}(n)\\ &= b_{K-1}\sum_{i=0}^{p}\beta_i\{A\cos[w_l(n-K+1)]\}^i + \ldots +\\ &b_1\sum_{i=0}^{p}\beta_i\{A\cos[w_l(n-1)]\}^i + b_0\sum_{i=0}^{p}\beta_i\{A\cos[w_l(n)]\}^i\\ &= \beta_0\sum_{i=0}^{K-1}b_i + \beta_1\sum_{i=0}^{K-1}b_i\{A\cos[w_l(n-i)]\}\ldots\\ &+ \beta_p\sum_{i=0}^{K-1}b_i\{A\cos[w_l(n-i)]\}^p.\end{aligned} \tag{25}$$

Denote a parameter column vector $\beta = [\beta_0\beta_1\ldots\beta_p]'$, then above equations (25) can be written as a system of equations

$$\Phi\beta = \mathbb{Y}, \tag{26}$$

where Φ is a $(N-q)\times(p+1)$ matrix with the kth column given by

$$\Phi(\cdot,k) = \begin{pmatrix}\sum_{i=0}^{K-1} b_i A^{k-1}\cos^{k-1}[w_l(K+1-i)])\\ \sum_{i=0}^{K-1} b_i A^{k-1}\cos^{k-1}[w_l(K+2-i)])\\ \vdots\\ \sum_{i=0}^{K-1} b_i A^{k-1}\cos^{k-1}[w_l(N-i)])\end{pmatrix}, \tag{27}$$

and

$$\mathbb{Y} = \begin{pmatrix}\sum_{i=0}^{K} d_i y^{<l>}(K+1-i)\\ \sum_{i=0}^{K} d_i y^{<l>}(K+2-i)\\ \vdots\\ \sum_{i=0}^{K} d_i y^{<l>}(N-i)\end{pmatrix}. \tag{28}$$

Then the parameters β can be solved in the least-squares sense according to (26), that is

$$\beta^\star = (\Phi'\Phi)^{-1}\Phi'\mathbb{Y}. \tag{29}$$

Remark 4.2. The presented estimation of nonlinear subsystem is natural. Other algorithms can be also applied to estimate the nonlinear part, see [40].

Table 1
The poles obtained without noise to output $y(n)$.

Poles	*BSA*	*IBSA*	*Nano's*[22]
a_1	0.50001215	0.50000080	$-0.8761 + 0.1062j$
a_2	0.20240560	0.20233191	$0.5643 - 0.0221j$

Table 2
The poles obtained with noise $SNR = 20dB$ to output $y(n)$.

Poles	*BSA*	*IBSA*	*Nano's*[22]
a_1	0.48261800	0.50001566	$0.8743 - 0.4325j$
a_2	0.18260904	0.20231243	$-0.6116 + 0.6413j$

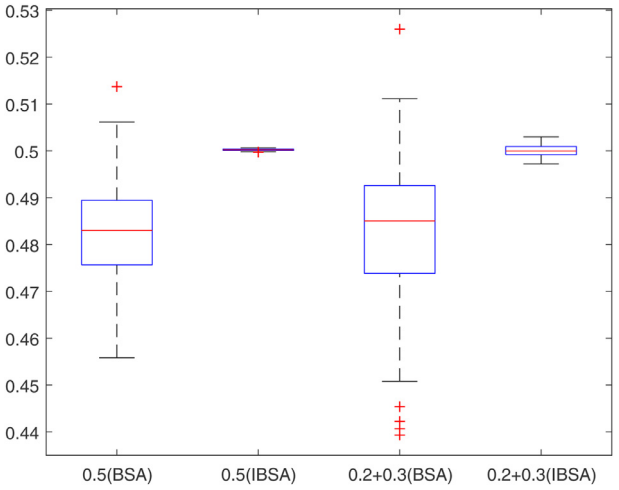

Fig. 2. Box chart of 100 estimated poles for linear subsystem obtained by backward shift algorithm, noises are given randomly with $SNR = 20dB$.

5. Example

In this section, an example is presented to illustrate the proposed algorithm. Consider a Hammerstein system, whose linear subsystem is given by

$$G(z) = \frac{9 - 1.05z}{(1 - 0.5z)(1 - 0.2z)} \tag{30}$$

and nonlinear subsystem is

$$f(x) = x + 0.8x^2 + 0.7x^3, \tag{31}$$

respectively. The linear part is a second order system with two distinguished poles, while its nonlinearity is a polynomial with order 3.

In the simulation, there are $L = 16$ input signals with frequency w_l $(l = 1, 2, \ldots, L)$, where w_l are arranged in the open interval $(0.1, \pi - 0.1)$ and the points 0 and π are avoided. Frequency responses of the linear subsystem in $(\pi, 2\pi)$ are set to be the conjugated values of estimated frequency responses in $(0, \pi)$. There are 2000 sampled data in each signal. The noise level is set to be *SNR* (*sigal to noise ratio*) $= 20dB$. By using BSA, we get the linear part $\widetilde{G}(z)$. The corresponding nonlinear part is approximated according to (29). We repeated the algorithm 100 times with random noises in $20dB$. The symbols IBSA and BSA are used representing the general spacing and equal spacing algorithm in the results, respectively. In the non-noise case, Table 1 shows the estimated results in different algorithms, including: BSA, IBSA and Nara-Ando's algorithm [22]. While the results of noised case are shown in Table 2.

Furthermore, we show the box chart of all data in 100 tests in Fig. 2. '•(IBSA)' and '•(BSA)' stand for the estimated results for pole • by non equally spaced BSA and equally spaced BSA, respectively. In order to combine the results in one figure for two poles 0.5 and 0.2, the results for 0.2 are all added up with 0.3. The estimated coefficients for $e_{a_1}(z)$ and $e_{a_2}(z)$ are $\lambda_1 = 0.1153, \lambda_2 = -0.0265$, respectively, associated with estimated nonlinear subsystem

$$\widetilde{f}(x) = -0.3068 + 101.4124x + 81.0854x^2 + 70.8824x^3. \tag{32}$$

In order to compare with other methods and the original system, simulation results are modified. The linear part are set to be $100\widetilde{G}(z)$ and nonlinear part is changed to $\widetilde{f}(x)/100$. This would not affect the behavior of entire Hammerstein system.

Table 3
Simulation results (λ_1, λ_2) and modified results for BSA and IBSA.

	λ_1	λ_2
Original	11.5	−2.5
IBSA	0.1153	−0.0265
Mod. IBSA	11.5261	−2.6470
BSA	0.1134	−0.0393
Mod. BSA	11.3416	−3.9312

Table 4
Simulation results and modified results for BSA and IBSA.

	β_0	β_1	β_2	β_3
Original	0	1	0.8	0.7
IBSA	−0.3068	101.4124	81.0854	70.8824
Mod. IBSA	−0.0031	1.0141	0.8108	0.7088
BSA	−3.9270	121.4482	97.2002	85.0324
Mod. BSA	−0.0393	1.2145	0.9720	0.8503

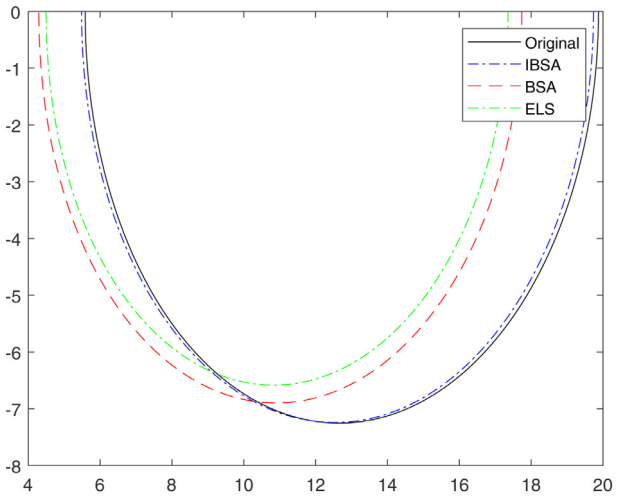

Fig. 3. The black line, blue dash-dot line, red dash line and green line are original linear subsystem, linear subsystem obtained by IBSA,BSA and ELS, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

After this modification, the modified results are shown in Table 3 and 4, where 'Mod.' stands for modified results. The obtained results are more clear.

In Fig. 5 and 5, identification results are compared with the extended least squares method (ELS). In both figures, the black lines, blue dash-dot lines, red dash lines and green lines are original linear (nonlinear) subsystem and results obtained by using backward shift algorithms and ELS, respectively. From the tables and figures, it can be seen that the backward shift algorithm can get effective estimations of poles, thus it can lead to efficient identification of entire Hammerstein systems.

6. Conclusion

In this paper, the backward shift algorithm is applied in identification of Hammerstein systems. A modification on spacing of frequencies not only generalizes the original algorithm, but also allows users to design their experiments more freely. From the tables and figures, it can be seen that the backward shift algorithm is more powerful in estimating poles of linear subsystems, and hence the entire Hammerstein systems. The backward shift algorithm is a new identification framework, then there would be more applications in identification problems.

Appendix A. Appendix

Proof of Theorem 4.1.. Without loss of generality, let $c = 1$, the frequency estimation for linear subsystem $cG(z)$ at w_l becomes

$$\widetilde{G}(e^{jw_l}) = G(e^{jw_l}) + V_l, \quad (l = 1, 2, \ldots L)$$

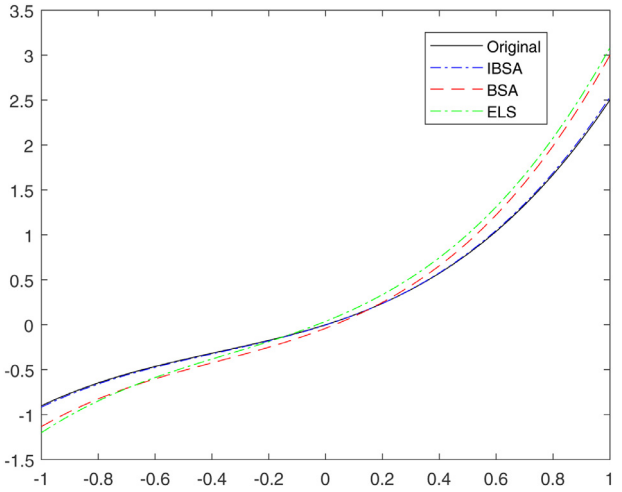

Fig. 4. Comparison of nonlinear parts. The black line, blue dash-dot line, red dash line and green line are original nonlinear subsystem, nonlinear subsystem obtained by IBSA, BSA and ELS, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

where $V_l = \frac{2V(e^{jw_l})}{\sqrt{N}}$ are the errors between true values $G(e^{jw_l})$ and estimated values shown in (15). Take backward operator to data set $\{\widetilde{G}(e^{jw_l})\}$, there is

$$\begin{aligned}&\mathbf{B}(\widetilde{G})(e^{jw_l})\\&= \frac{\widetilde{G}(e^{jw_l}) - \frac{1}{2\pi}\sum_{k=1}^{L}\widetilde{G}(e^{jw_l})(w_{k+1}-w_k)}{e^{jw_l}}\\&= \mathbf{B}(G)(e^{jw_l}) + \widetilde{\mathbf{B}}(O)(V_l),\end{aligned}$$

where $\widetilde{\mathbf{B}}(O)(V_l)$ is

$$\widetilde{\mathbf{B}}(O)(V_l) = \frac{V_l - \frac{1}{2\pi}\sum_{k=1}^{L}V_k(w_{k+1}-w_k) + T_0}{e^{jw_l}},$$

and T_0 is the error given by

$$T_0 = G(0) - \frac{1}{2\pi}\sum_{k=1}^{L}G(e^{jw_k})(w_{k+1}-w_k).$$

Continue to the second step: taking backward shift to $\mathfrak{B}(\widetilde{G})(e^{jw_l})$,

$$\mathbf{B}^2(\widetilde{G})(e^{jw_l}) = \mathbf{B}^2(G)(e^{jw_l}) + \widetilde{\mathbf{B}}^2(O)(V_l),$$

$\mathbf{B}^2(G)(e^{jw_l})$, $\widetilde{\mathbf{B}}^2(O)(V_l)$ are given respectively by

$$\mathbf{B}^2(G)(e^{jw_l}) = \frac{\mathbf{B}(G)(e^{jw_l}) - \mathbf{B}(G)(0)}{e^{jw_l}},$$

$$\begin{aligned}&\widetilde{\mathbf{B}}^2(O)(V_l)\\&= \frac{\widetilde{\mathbf{B}}(O)(V_l) - \frac{1}{2\pi}\sum_{k=1}^{L}\widetilde{\mathbf{B}}(O)(V_k)(w_{k+1}-w_k) + T_1}{e^{jw_l}}\end{aligned}$$

with

$$T_1 = \mathbf{B}(G)(0) - \frac{1}{2\pi}\sum_{k=1}^{L}\mathbf{B}(G)(e^{jw_k})(w_{k+1}-w_k).$$

Repeating to the m-th time, one would obtain data sets $\{\mathfrak{B}^m(G)(e^{jw_l})\}_{l=1}^{L}(m = 1, \ldots, K)$ as

$$\mathbf{B}^m(\widetilde{G})(e^{jw_l}) = \mathbf{B}^m(G)(e^{jw_l}) + \widetilde{\mathbf{B}}^m(O)(V_l),$$

and $\mathbf{B}^m(G)(e^{jw_l})$, $\widetilde{\mathbf{B}}^m(O)(V_l)$ are given respectively by

$$\mathbf{B}^m(G)(e^{jw_l}) = \frac{\mathbf{B}^{m-1}(G)(e^{jw_l}) - \mathbf{B}^{m-1}(G)(0)}{e^{jw_l}},$$

$$\widetilde{\mathbf{B}}^m(O)(V_l) \tag{A.1}$$
$$= \frac{\widetilde{\mathbf{B}}^{m-1}(O)(V_l) - \frac{1}{2\pi}\sum_{k=1}^{L}\widetilde{\mathbf{B}}^{m-1}(O)(V_k)(w_{k+1}-w_k)}{e^{jw_l}}$$
$$+ \frac{T_{m-1}}{e^{jw_l}},$$

and T_m are errors given by

$$T_m = \mathbf{B}^m(G)(0) - \frac{1}{2\pi}\sum_{k=1}^{L}\mathbf{B}^m(G)(e^{jw_l})(w_{k+1}-w_k).$$

Finally, one gets a system of equations

$$\Psi\mu = b, \tag{A.2}$$

where Ψ is an $L\times K$ matrix with elements $\Psi_{LK}(l,m) = \mathfrak{B}^{m-1}(\widetilde{G})(e^{jw_l})$, μ is the column vector of parameters $[\mu_1\mu_2\ldots\mu_{K-1}]'$ and b is a column whose lth element is $\mathfrak{B}^M(\widetilde{G})(e^{jw_l})$.

For the new approximation, we can see

$$\lim_{\substack{\epsilon\to 0,\\ L\to\infty}} \frac{1}{2\pi}\int_0^{2\pi}\sum G(e^{jw_l})\chi(\cdot)dw = G(0), \tag{A.3}$$

consequently, for all l

$$\lim_{\substack{\epsilon\to 0,\\ L\to\infty}} \mathbf{B}(\widetilde{G})(e^{jw_l}) = \mathfrak{B}(G)(e^{jw_l}), \tag{A.4}$$

and thus

$$\lim_{\substack{\epsilon\to 0,\\ L\to\infty}} \mathbf{B}^m(\widetilde{G})(e^{jw_l}) = \mathbf{B}^m(G)(e^{jw_l}). \tag{A.5}$$

Then according to equations (A.2), the left goes similarly with the proof in [20], we can get

$$\lim_{\substack{\epsilon\to 0,\\ L\to\infty}} \mu_L = \mu^\star, \tag{A.6}$$

where $\mu^\star$ is the true solution of (7), μ_L is the obtained solution by using *BSA* with $\{E_l\}_{l=1}^{L}$. The proof is complete. □

References

[1] H. Akçay, B. Ninness, Rational basis functions for robust identification from frequency and time domain measurements, Automatica 34 (1998) 1101–1117.
[2] E.W. Bai, Frequency domain identification of Hammerstein models, IEEE trans, Autom. Control 48 (4) (2003) 530–542.
[3] E.W. Bai, A blind approach to the Hammerstein-wiener model identification, Automatica 38 (2002) 967–979.
[4] H.B. H. B. Chen, Y.S. Xiao, F. Ding, Hierarchical gradient parameter estimation algorithm for Hammerstein nonlinear systems using the key term separation principle,, Appl. Math. Comput. 247 (2014) 1202–1210.
[5] J.A. Cima, W.T. Ross, The backward shift on the hardy space, Providence, RI: American Mathematical Society (2000).
[6] F. Ding, X.P.P. Liu, G.J. Liu, Identification methods for Hammerstein nonlinear systems, Digit. Signal Processing 21 (2) (2011) 215–238.
[7] F. Ding, X.G. Liu, J. Chu, Gradient-based and least-squares-based iterative algorithms for Hammerstein systems using the hierarchical identification principle, IET Control Theory Appl. 7 (2) (2013) 176–184.
[8] F. Ding, Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling,, Appl. Math. Model. 37 (4) (2013) 1694–1704.
[9] F. Ding, H.B. Chen, L. Xu, J.Y. Dai, Q.S. Li, T. Hayat, A hierarchical least squares identification algorithm for Hammerstein nonlinear systems using the key term separation, J. Frankl. Inst. 355 (8) (2018) 3737–3752.
[10] S.J. Dong, L. Yu, W.A. Zhang, B. Chen, Robust extended recursive least squares identification algorithm for Hammerstein systems with dynamic disturbances,, Digit. Signal Processing 101 (2020) 102716.
[11] S.J. Dong, L. Yu, W.A. Zhang, B. Chen, Robust hierarchical identification of Wiener systems in the presence of dynamic disturbance, J. Frankl. Inst. 357 (6) (2020) 3809–3834.
[12] F. Giri, E.W. Bai, Block-oriented nonlinear system identification, Springer, 2010.
[13] W. Greblichi, Continuous-time Hammerstein system identification, IEEE Trans. Autom. Control 45 (2000) 1232–1236.
[14] J.C. Gómez, E. Baeyens, Identification of block-oriented nonlinear systems using orthonormal bases, J. Process Contr. 14 (2004) 685–697.
[15] P.S.C. Heuberger, P.M.J.V.d. Hof, B. B. Wahlberg, Modelling and identification with rational orthogonal basis functions, eds., Springer, London, 2005.
[16] Y. Ji, C. Zhang, Z. Kang, Y. T, Parameter estimation for block-oriented nonlinear systems using the key term separation, Int. J. Robust Nonlinear Control 30 (9) (2020) 3727–3752.
[17] Y. Ji, X.K. Jiang, L.J. Wan, Hierarchical least squares parameter estimation algorithm for two-input Hammerstein finite impulse response systems, J. Frankl. Inst. 357 (8) (2020) 5019–5032.
[18] W. Mi, T. Qian, F. Wan, A fast adaptive model reduction method based on Takenaka-Malmquist systems, Syst. Control Lett. 61 (2012) 223–230.
[19] W. Mi, T. Qian, Frequency domain identification: an adaptive algorithm based on rational orthogonal system, Automatica 48 (6) (2012) 1154–1162.
[20] W. Mi, T. Qian, On backward shift algorithm for estimating poles of systems, Automatica 50 (2014) 1603–1610.
[21] W. Mi, T. Qian, S. Li, Basis pursuit for frequency domain identification, Math. Methods Appl. Sci. 39 (3) (2016) 498–507.
[22] T. Nara, S. Ando, Direct localization of poles of a meromorphic function from measurements on an incomplete boundary, Inverse Probl. 26 (2010) 015011.

[23] T.O.e. Silva, Optimality conditions for truncated kautz networks with two periodically repeating complex conjugate poles, IEEE Trans. Autom. Control. 40 (2) (1995) 342–346.
[24] R. Pintelon, J. Schoukens, System identification: A Frequency domain approach, second edition, John Wiley & Sons, Inc., Hoboken, New Jersey, 2012.
[25] T. Qian, E. Wegert, Optimal approximation by blaschke forms, Complex Variables and Elliptic Equations 58 (1) (2013) 123–133.
[26] R.S. Risuleo, G. Bottegal, H. Hjalmarsson, A nonparametric kernel-based approach to Hammerstein system identification, Automatica 85 (2017) 234–247.
[27] E.M. Stein, R. Shakarchi, Complex analysis, Princeton University Press, Princeton and Oxford, 2003.
[28] D.Q. Wang, L.W. Li, Y. Ji, Y.R. Yan, Model recovery for Hammerstein systems using the auxiliary model based orthogonal matching pursuit method, Appl. Math. Model. 54 (2018) 537–550.
[29] D.Q. Wang, Y.R. Yan, Y.J. Liu, J.H. Ding, Model recovery for Hammerstein systems using the hierarchical orthogonal matching pursuit method, J. Comput. Appl. Math. 345 (2019) 135–145.
[30] D.Q. Wang, S. Zhang, M. Gan, J.L. Qiu, A novel EM identification method for Hammerstein systems with missing output data, IEEE T. Ind. Inform. 16 (4) (2020) 2500–2508.
[31] D.Q. Wang, Q.H. Fan, Y. Ma, An interactive maximum likelihood estimation method for multivariable Hammerstein systems, J. Frankl. Inst. 357 (2020) 12986–13005.
[32] A. Wills, T.B. Schön, L. Ljung, B. Ninness, Identification of Hammerstein-wiener models, Automatica 49 (2013) 70–81.
[33] L. Xu, A proportional differential control method for a time-delay system using the taylor expansion approximation, Appl. Math. Comput. 236 (2014) 391–399.
[34] L. Xu, L. Chen, W.L. Xiong, Parameter estimation and controller design for dynamic systems from the step responses based on the newton iteration, Nonlinear Dyn. 79 (3) (2015) 2155–2163.
[35] L. Xu, The damping iterative parameter identification method for dynamical systems based on the sine signal measurement, Signal Process. 120 (2016) 660–667.
[36] L. Xu, F. Ding, Q.M. Zhu, Hierarchical newton and least squares iterative estimation algorithm for dynamic systems by transfer functions based on the impulse responses, Int. J. Syst. Sci. 50 (1) (2019) 141–151.
[37] L. Xu, F. Ding, L.J. Wan, J. Sheng, Separable multi-innovation stochastic gradient estimation algorithm for the nonlinear dynamic responses of systems, Int. J. Adapt Control 34 (7) (2020) 937–954.
[38] L. Xu, F. Ding, Q.M. Zhu, Decomposition strategy-based hierarchical least mean square algorithm for control systems from the impulse responses, Int. J. Syst. Sci. 52 (9) (2021) 1806–1821.
[39] W.X. Zhao, H.F. Chen, Adaptive tracking and recursive identification for Hammerstein systems, Automatica 45 (2019) 2773–2783.
[40] W.X. Zhao, H.F. Chen, R. Tempo, Recursive nonparametric identification of nonlinear systems with adaptive binary sensors, IEEE Trans. Autom. Control 62 (8) (2017) 3959–3971.