Applied Mathematics and Computation 361 (2019) 354–364

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Identification of discrete Hammerstein systems by using adaptive finite rational orthogonal basis functions☆

Wen Mi[a,*], Hangmei Rao[a], Tao Qian[b], Shouming Zhong[a]

[a] School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China
[b] Department of Mathematics, Faculty of Science and Technology, University of Macau, Macao, China

ARTICLE INFO

Keywords:
Discrete systems
Hammerstein model
Rational functions
Rational approximation
Adaptive identification

ABSTRACT

In this paper, an adaptive identification method is presented, it is developed for discrete Hammerstein systems. We adopt a frequency-domain technique by using sampled input–output data. The excitation signals are chosen to be the fundamental harmonics, by which one could get approximating estimate of frequency responses to linear subsystem. By using adaptive rational orthogonal basis functions, we obtain approximations of linear subsystem, it is carried out through selecting poles of the basis functions under some criterion. Meanwhile, efficient estimates of both the nonlinear part and linear part could be obtained in the presence of noise.

1. Introduction

Identification of nonlinear systems has been an interesting research topic. This could probably own to that many systems in nature and engineering are nonlinear. At presence of the nonlinearity in the systems, the identification problem becomes more difficult than linear systems. One efficient way is to use block-oriented models for some kind of nonlinear systems, these models construct systems as series subsystems of linear parts and nonlinear parts. The classical models include Hammerstein model, Wiener model and Sandwich model (Hammerstein–Wiener and Wiener–Hammerstein model). The Hammerstein model is powerful in presenting the nonlinear systems in which the static (memoryless) nonlinear subsystem is followed by a dynamic (with memory) linear subsystem, and it has many applications in engineering such as in [15]. Because of this special structure, many efficient identification methods of linear systems could be referred in the identification problem of Hammerstein systems both in frequency domain and time domain [2,3,5,15,19,26,28].

People have been interested in identification of Hammaerstein model for a long time since [26], many identification methods have been developed, including the iterative method [26,31,35] in which an appropriate parameterization of the linear and nonlinear part is chosen, the estimation is carried out by minimizing one of the separated parameter set, then the other is fixed. The risk of the iterative method is that it could be divergent [32]; without knowing the form of the nonlinear part, the stochastic method [14] still works on identification of Hammerstein model under a condition that the input signal is white; on some restrictive conditions, the nonlinear least-squares method can be efficient [5]; in the blind method [3], the idea is to use the blind identification method to estimate the linear part no matter what structure of the nonlinearity is; the subspace method [12,34] is a nonparametric estimation by using periodic input signals; the over-parameterization

☆ The research was partially supported by National Natural Science Foundation of China (NSFC) under Grants 11401079 and 11571083.
* Corresponding author.
E-mail addresses: mathmw@uestc.edu.cn (W. Mi), yqr@std.uestc.edu.cn (H. Rao), fsttq@umac.mo (T. Qian), zhongsm@uestc.edu.cn (S. Zhong).

https://doi.org/10.1016/j.amc.2019.05.051
0096-3003/

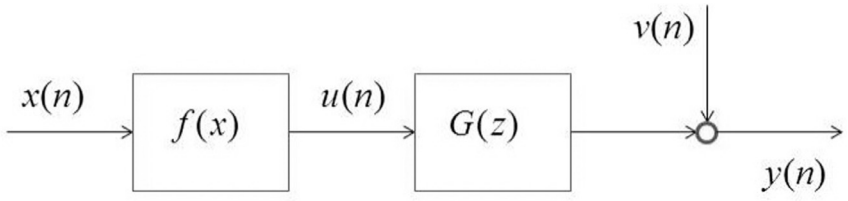

Fig. 1. Discrete Hammerstein model.

method [7] is the over-parameterization of the Hammerstein systems, unknown parameters could be expressed as linear form by over-parameterization method so that the problem can be solved by linear estimation, however, the difficulty is the dimension of unknown parameters is usually large, which results in problems in analysis of convergence and robustness; frequency domain identification methods were also studied in [2,6,40] for Hammerstein model.

In this paper, we will study the identification problem of Hammerstein systems in frequency domain. By using sinusoidal input signals, the frequency domain methods have many advantages [19,28,33]. For Hammerstein systems, specially, the periodicity of inputs will result to interesting phenomenon: first, signals inside the Hammerstein systems consist of frequencies that are integer multiples of the input frequencies; second, the output of the nonlinear part could be expressed as a Fourier series [2]. We will use the same idea of input in our approach as [2], but finite rational orthogonal basis functions (FROBFs) will be used in our method. In fact, the rational orthogonal bases is not its first time applied in identification of Hammerstein systems, authors in [15] studied identification of Hammerstein systems with orthogonal bases in process control, it is a parameterization identification method with two sets of parameters of linear subsystem and nonlinear subsystem, the parameters including the coefficients of basis functions' are estimated by using matrix decomposition. In this paper, we use the rational orthogonal basis functions by taking consideration of their poles' selection, because obviously the poles of basis functions play a key role in approximation convergence of these bases, if they were selected at will, it may lead to low convergence rate. In order to avoid this defect, we will carry out our identification of Hammerstein systems by focusing on the poles selections of basis functions.

The outline of this paper is given as follows. In Section 2, we state the identification problem and introduce the estimate of frequency responses of linear subsystem by using input and output data. The estimate method of linear part is given in Section 3. After that, identification of the nonlinear part is discussed in Section 4. An example is given to illustrate the efficiency of the proposed idea in Section 5. Some conclusions are presented in the last section.

2. Problem formulation

This paper is concerned with identification problem of discrete Hammerstein model as shown in Fig. 1. Both input and output are single. In the Hammerstein model, a linear dynamic subsystem $G(z)$ follows a nonlinear static subsystem $f(x)$. $x(n)$, $y(n)$ and $v(n)$ are discrete input signal, output signal and noise, respectively. Particularly, we use $y^{\star}(n)$ as measured output signal (with noise) in our text, while $y(n)$ stands for true output signal (without noise). $u(n) = f(x(n))$ is the unknown middle signal which is not only the output of nonlinear subsystem $f(x)$ but also the input of the linear part $G(z)$.

The idea of frequency domain identification is to apply a series of following discrete input signals that are:

$$x(n) = E\cos(\omega_k n) \tag{1}$$

with frequency $\omega_k = \frac{2\pi k}{K}(k = 1, 2, \ldots, K)$. The same as linear systems, the proposed method need the experiments to be repeated at each frequency ω_k, that is, to each input signal in (1) at ω_k, there is an input–output sampled data set $\{x^k(n), y^{\star k}(n)\}_{n=1}^{N}$.

Throughout this paper, we have some assumptions as following:

- The linear subsystem $G(z)$ is assumed to be stable and casual. It is assumed that $G(z) \in H_2(\mathbb{D})$. Let $\mathbb{D}$ denote the unit disc in the complex plane and $H_2(\mathbb{D})$ denote the Hardy-2 space in $\mathbb{D}$.
- The nonlinear part is a static system, and $u = f(\cdot)$ is a polynomial with known order range from p_1 to p_2.

The assumption about nonlinear part is common, it is like [6,13,15] in that the nonlinearity is assumed as a polynomial with known order, but in fact, it is not easy to know the precise order of polynomial. Note that the second assumption above could be released, such as $f(x)$ is a continuous function in $[-E, E]$, then according to approximation theory, $f(x)$ certainly could be well approximated by a polynomial. Under signals given by (1) which is even with periodic $\frac{2\pi}{\omega_n}$, it is clear to notice $u(n) = f(x)$ is also even and periodic. With continuity assumption of f, it has a decomposition in Fourier series as

$$u(n) = \sum_{i=0}^{\infty} r_i \cos(i\omega_k n), \tag{2}$$

where r_i is the ith term coefficients of Fourier series. In [2], it shows

Lemma 2.1. *Let $x(n) = E\cos(\omega_k n)$ be input signals, then each Fourier coefficients r_i is invariant to any input frequency ω_k.*

The above result means the Fourier expansion (2) has identical coefficients for different input frequencies. On the other hand, since the internal signal has an expansion of sinusoid signals (2), when $x(n)$ is treated as input of linear part $G(z)$, the output of $G(z)$ could be written as

$$y(n) = \sum_{i=0}^{\infty} r_i |G(e^{j\omega_k i})| \cos(i\omega_k n + \varphi_i) \tag{3}$$

where j is the imaginary unit and

$$\varphi_i = \angle G(e^{jw_k i}). \tag{4}$$

Those relationships are parts of our research foundations. Our task is to find a nonlinear subsystem $\widehat{f}(x)$ and a linear part $\widehat{G}(z)$ based on the sampled input and output data sets $\{x^k(n), y^{\star k}(n)\}_{n=1}^N$. Furthermore, we require $\widehat{f}, \widehat{G}$ have the same input–output behavior as $\{f(x), G(z)\}$ does, viz,

$$(\widehat{f}(x), \widehat{G}(z)) \to (f(x), G(z)). \tag{5}$$

Note that the focus is the whole input–output behavior of the pair $(\widehat{f}(x), \widehat{G}(z))$, neither $\widehat{f}(x)$ necessarily approximate to $f(x)$ nor $\widehat{G}(z)$ necessarily approximate to $G(z)$ completely. They may have a difference in multiple. This is also a special characteristic of Hammerstein model. Because any pair of $(\alpha f(x), \frac{1}{\alpha} G(z))$ $(\alpha \neq 0)$ would carry out identical input and output signals, thus any efforts to find true linear or nonlinear subsystems are in vain under current assumptions, so Hammerstein systems are unidentifiable. But some kind of frequency responses of linear part could be estimated through input and output measurements of the entire Hammerstein system as shown in the following subsection.

2.1. Estimate of frequency responses

In order to actualize the frequency-domain identification, frequency responses are estimated from input and output measurements $\{x^k(n), y^{\star k}(n)\}_{n=1}^N$. The discrete time Fourier transform of $u(n), y(n), v(n)$ are given by

$$U_N(\omega) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} u(n) e^{(-j\omega n)}, \tag{6}$$

$$Y_N(\omega) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} y(n) e^{(-j\omega n)}, \tag{7}$$

$$V_N(\omega) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} v(n) e^{(-j\omega n)}, \tag{8}$$

respectively. Similar to frequency responses of linear systems, following estimates are considered in frequency domain method of Hammerstein model:

$$\widetilde{G}(e^{j\omega_k}) = E \frac{Y_N(\omega_k)}{U_N(\omega_k)}, \tag{9}$$

where $Y_N(\omega_k)$ and $U_N(\omega_k)$ are the discrete time Fourier transforms as defined in (7) and (6). By simple computation we have

Lemma 2.2. *Given $Y_N(\omega_k)$ and $U_N(\omega_k)$, the discrete time Fourier transforms defined in (7) and (6), there holds*

$$\frac{1}{\sqrt{N}} U_N(\omega_k) = \frac{E}{2} \tag{10}$$

and

$$\frac{1}{\sqrt{N}} Y_N(\omega_k) = \frac{r_1}{2} G(e^{j\omega_k}) + \frac{1}{\sqrt{N}} V_N(\omega_k). \tag{11}$$

Proof. Mathematical computation gives rise to this result. The detail is as follows: consider the interval $n = 1, 2, \ldots, N$, N is a multiple of L: $N = sL$ writing by

$$\cos(\omega_k n) = \frac{1}{2}(e^{j\omega_k n} + e^{-j\omega_k n}), \tag{12}$$

which gives

$$U_N(\omega) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N} \frac{E}{2} (e^{j(\omega_k - \omega)n} + e^{-j(\omega_k - \omega)n}). \tag{13}$$

According to the following result

$$\frac{1}{N}\sum_{n=0}^{N} e^{\frac{2\pi jrn}{N}} = \begin{cases} 1, & r = 0, \\ 0, & 1 \le r < N, \end{cases} \tag{14}$$

then it is followed by

$$|U_N(\omega)|^2 = \begin{cases} N * \frac{E^2}{4}, & if \quad \omega = \pm\omega_k = \frac{2\pi}{P} = \frac{2\pi s}{N}, \\ 0, & if \quad \omega = \frac{2\pi k}{N}, k \neq s. \end{cases} \tag{15}$$

This leads to $\frac{1}{\sqrt{N}}U_N(\omega) = \frac{E}{2}$ [19].

In (7), substitute $y(n)$ by

$$y^{\star}(n) = \sum_{i=0}^{\infty} r_i |G(e^{jw_k i}| \cos(iw_k n + \varphi_i) + v(n), \tag{16}$$

we have

$$\begin{aligned} \frac{1}{\sqrt{N}}Y_N(\omega_k) &= \sum_{n=0}^{N}\sum_{i=0}^{\infty} r_i |G(e^{j\omega_k i})| \cos(i\omega_k n + \varphi_i) e^{-j\omega_k n} + \frac{1}{\sqrt{N}}V_N(\omega_k) \\ &= \sum_{i=0}^{\infty} r_i |G(e^{j\omega_k i})| \frac{1}{2}\sum_{n=1}^{N} (e^{j(i-1)\omega_k n} e^{j\varphi_i} + e^{-j(i+1)\omega_k n} e^{-j\varphi_i}) + \frac{1}{\sqrt{N}}V_N(\omega_k) \\ &= \frac{r_1}{2}|G(e^{j\omega_k 1})| e^{j\varphi_1} + \frac{1}{\sqrt{N}}V_N(\omega_k) \\ &= \frac{r_1}{2}G(e^{j\omega_k}) + \frac{1}{\sqrt{N}}V_N(\omega_k). \end{aligned} \tag{17}$$

The proof is complete. □

From the calculation of $Y_N(\omega_k)$ and $U_N(\omega_k)$, it is followed by

$$\widetilde{G}(e^{j\omega_k}) = E\frac{Y_N(\omega_k)}{U_N(\omega_k)} = r_1 G(e^{j\omega_k}) + \frac{2}{\sqrt{N}}V_N(\omega_k), \tag{18}$$

where $V_N(\omega_k)$ is the DFT of $v(n)$, then we get the error of estimate $\widetilde{G}(e^{j\omega_k})$ as follows:

$$\widetilde{G}(e^{j\omega_k}) - r_1 G(e^{j\omega_k}) = 2V_N(\omega_k). \tag{19}$$

Note that (18) gives estimates of $r_1 G(e^{j\omega})$ but no $G(e^{j\omega})$, fortunately, thanks to its unidentifiability, we still can give an identical Hammerstein model $(\frac{f(x)}{r_1}, r_1 G(z))$ no matter what r_1 is (except $r_1 = 0$). Furthermore, for our estimates there is one more assumption about nonlinearity $f(x)$ that r_1 is nonzero.

With sufficient many estimated frequency responses $\{\omega_k, \widetilde{G}(e^{jw_k})\}$, the first proceed is the identification of linear part $r_1 G(z)$. Our identification technique is based on an adaptive rational approximation.

3. Identification of linear subsystem by best rational approximation

In this section, identification of the linear part $r_1 G(z)$ will be carried out by using finite rational orthogonal basis functions (FROBFs). These bases are well studied in identification of linear systems, they are fundamental tools in system identification of linear systems both in frequency domain and time domain [16]. In the unit disc $\mathbb{D}$ case, the generalized rational orthogonal basis functions are defined by

$$\mathcal{B}_k(z) \triangleq \frac{\sqrt{1-|\zeta_k|^2}}{1-\overline{\zeta}_k z}\prod_{l=1}^{k-1}\frac{z-\zeta_l}{1-\overline{\zeta}_l z}, \tag{20}$$

where $\zeta_k \in \mathbb{D}$ and $k = 1, 2, \ldots$. These bases are complete in the Hardy spaces $H_p(\mathbb{D}), 1 \le p < \infty$ and Disc Algebra $A(\mathbb{D})$ with a miner condition [27]

$$\sum_{k=1}^{\infty}(1-|\zeta_k|) = \infty. \tag{21}$$

There are only finite basis functions used in practice, these basis functions span a subspace, in this point of view, it could be treated as finding a projection in this spanned subspace. In the existing method of using FROBFs, researchers often construct

linear-in-parameters models in identification of linear systems, as follows,

$$\widetilde{G}(z) = \sum_{m=1}^{M} \theta_m \mathcal{B}_m(z), \tag{22}$$

where $\{\mathcal{B}_k(z)\}$ are rational orthogonal basis functions, $\{\theta_l\}$ are the parameters to be determined. This model structure (22) is a generalization of classical models, including the classical FIR model when all ζ_k are fixed as 0; the Laguerre model [36] while all ζ_k are a same real-valued nonzero number; two-parameter Kautz model [37] when all ζ_k are fixed as a same complex-valued number. In this construction, researches can use Least-square method to get parameters quickly, analogous computation algorithms are widely used in applied mathematics and control theory [17,18,20,21,39]. The general model (22) is well studied since 1990s [1,8–10,22,24,25,27] when only condition (21) is required. This requirement is easy to satisfy, because the number of basis functions in use is finite, then the rest of sequence $\{\zeta_k\}$ can be viewed as 0 and finite $\{\mathcal{B}_k(z)\}$ generate a closed subspace of $H_2(\mathbb{D})$. An advantage of the generalized model, compared to classical models, is variability of the poles. Especially, when finite basis functions are used, suitable ζ_ks would lead to best approximations to the original linear system. However, finding true poles of systems or selecting poles for best approximations are still problems [16] that prevent maximum efficiency of ROBF.

In this stage, we will use finite M rational orthogonal basis functions $\{\mathcal{B}_1(z), \mathcal{B}_2(z), \ldots, \mathcal{B}_M(z)\}$ to get approximation of $r_1 G(z)$ in the energy sense, this set is also called *M Blaschke form* [30]. The key point is to select $\zeta = \{\zeta_m^\star\}_{m=1}^M$ in the unit disc such that

$$\{\zeta_m^\star\}_{m=1}^M = \arg\min_{\zeta\in\mathbb{D}} \| r_1 G(z) - \sum_{m=1}^{M} \theta_m \mathcal{B}_m(z) \|_2^2, \tag{23}$$

or equivalently

$$\{\zeta_m^\star\}_{m=1}^M = \arg\max_{\zeta\in\mathbb{D}} \| \sum_{m=1}^{M} \theta_m \mathcal{B}_m(z) \|_2^2. \tag{24}$$

These problems will result to the classical *M-best rational approximation* problem whose existence has already been proven [22,25,38]. But the algorithm to find the M-best rational approximation are still not satisfying. Because of the orthonormality, problem (24) could be further written as

$$\begin{aligned}
\{\zeta_m^\star\}_{m=1}^M &= \arg\max_{\zeta\in\mathbb{D}} \| \sum_{m=1}^{M} \theta_m \mathcal{B}_m(z) \|_2^2 \\
&= \arg\max_{\zeta\in\mathbb{D}} \| \sum_{m=1}^{M} \langle r_1 G, \mathcal{B}_m \rangle \mathcal{B}_m(z) \|_2^2 \\
&= \arg\max_{\zeta\in\mathbb{D}} \sum_{m=1}^{M} \| \langle r_1 G, \mathcal{B}_m \rangle \|^2,
\end{aligned} \tag{25}$$

the above formula means it convert problem (24) to a nonlinear optimization problem with multiple real-valued variables. Problem (25) could theoretically be solved by using modern numerical methods, unfortunately when M increases the problem get harder or even has no efficient methods to find the optimal solution, in this case an alternative flexible algorithm is adopted to get suboptimal solutions, the algorithm is briefly stated in the next subsection.

3.1. Adaptive Fourier decomposition algorithm

In this subsection, we provide a summery to the adaptive Fourier decomposition algorithm (AFD) for $H_2(\mathbb{D})$ functions, the detail could be found in [29]. An evaluator is denoted by $e_{\{\zeta\}}(z) = \frac{\sqrt{1-|\zeta_k|^2}}{1-\bar{\zeta}_k z}$, for $h, g \in H_2(\mathbb{D})$ the inner product in $H_2(\mathbb{D})$ is defined by

$$\langle h, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} h(e^{j\omega}) \overline{g(e^{j\omega})} d\omega. \tag{26}$$

Then by Cauchy integral formula, there holds

$$\begin{aligned}
\langle h, e_{\{\zeta\}} \rangle &= \frac{\sqrt{1-|\zeta_k|^2}}{2\pi} \int_0^{2\pi} h(e^{j\omega}) \overline{\frac{1}{1-\bar{\zeta} e^{j\omega}}} d\omega \\
&= \sqrt{1-|\zeta_k|^2} h(\zeta).
\end{aligned} \tag{27}$$

Based on this truth, the algorithm for $h(z) \in H_2(\mathbb{D})$ could be summarized as follows:

Step 1: Let $h_1 = g_1 = h$.

$$h(z) = g_2(z)\frac{z-\zeta_1}{1-\overline{\zeta_1}z} + \langle g_1, e_{\{\zeta_1\}}\rangle \mathcal{B}_{\{\zeta_1\}}, \tag{28}$$

where $g_2(z) \in H_2(\mathbb{D})$ and ζ_1 is selected by

$$\zeta_1 = \arg\max\{|\langle g_1, e_{\{\zeta\}}\rangle|^2 : \zeta \in \mathbb{D}\}.$$

Step 2: For $g_2(z)$, in the same way it has the following decomposition

$$g_2(z) = g_3(z)\frac{z-\zeta_2}{1-\overline{\zeta_2}z} + \langle g_2, e_{\{\zeta_2\}}\rangle \mathcal{B}_{\{\zeta_2\}}, \tag{29}$$

which gives rise to

$$h(z) = g_3(z)\prod_{l=1}^{2}\frac{z-\zeta_l}{1-\overline{\zeta_l}z} + \langle g_2, e_{\{\zeta_2\}}\rangle \mathcal{B}_{\{\zeta_1,\zeta_2\}} + \langle g_1, e_{\{\zeta_1\}}\rangle \mathcal{B}_{\{\zeta_1\}}, \tag{30}$$

in which $g_2(z) \in H_2(\mathbb{D})$, too.

Step 3: Continue to kth time, it gives finite k terms approximations

$$h_k(z) = \sum_{l=1}^{k}\langle g_l, e_{\{\zeta_1\}}\rangle \mathcal{B}_{\zeta_1,\zeta_2,\ldots,\zeta_l}(z), \tag{31}$$

where all $g_l(z)$s belong to $H_2(\mathbb{D})$, and they are recursively obtained by

$$g_l(z) = (g_{l-1}(z) - (1-|\zeta_{l-1}|^2)\frac{g_{l-1}(\zeta_{l-1})}{1-\overline{\zeta_{l-1}(z)}})\frac{1-\overline{\zeta_l}z}{z-\zeta_{l-1}}. \tag{32}$$

Meanwhile, each ζ_l is selected in maximal criterion

$$\zeta_l = \arg\max\{|\langle g_l, e_{\{\zeta\}}\rangle|^2 : \zeta \in \mathbb{D}\}. \tag{33}$$

This method also results to rational approximations (31), but it avoids nonlinear optimization problem with multiple variables, it is also a type of greedy algorithm, some convergence result could be referred in [11]. In each step there is only one complex-valued variable ζ in the maximal selection criterion, in fact, in the module sense it gives $|\langle g_1, e_{\{\zeta_1\}}\rangle|^2 = (1-|\zeta|^2)|g_k(\zeta)|^2$, which converts the complex-valued problem to a real-valued nonlinear problem with two variables. This real nonlinear optimization problem has at most two variables, it is easier than the original one (25), and it can be solved by using modern numerical algorithms. We will modify this algorithm by selecting two poles at each time in our identification problem.

3.2. Adaptive identification of $r_1G(z)$

The adaptive algorithm (AFD) is directly developed for functions belonging to $H_2(\mathbb{D})$, while in our problem we have a set of estimated frequency domain measurements of $r_1G(z)$, $\{\omega_k, r_1G(e^{j\omega_k})\}_{k=1}^{K}$. In order to use the adaptive algorithm, we first make up a function $\widetilde{h}(z)$ by using Cauchy integral formula based on the data set,

$$\widetilde{h}(z) = \frac{1}{2\pi j}\int_0^{2\pi}\frac{\sum_k r_1G(z)\chi_{k(\ldots)}(w)}{e^{jw}}de^{jw}, \tag{34}$$

where $\chi_{(.)}$ is the indication function. $\widetilde{h}(z)$ is an $H_2(\mathbb{D})$ function and approximates to $r_1G(z)$ with an assumption, that is, in L_2 norm sense,

$$\sum_k r_1G(z)\chi_{k(\ldots)}(w) = \sum_{k=1}^{\frac{N}{2}} r_1G(z)\chi_{(w_k,w_{k+1})}(w) + \sum_{k=\frac{N}{2}+2}^{N+1} r_1G(z)\chi_{(w_k,w_{k-1})}(w) \tag{35}$$

$$\underset{N\to\infty}{\longrightarrow} r_1G(e^{jw}). \tag{35}$$

Note that this assumption can be satisfied when the noise is at a low level.

Therefore, the identification algorithm of linear part is given as follows:

(1) Collect frequency responses estimates of $r_1G(z)$ which is $\{\omega_k, \widetilde{G}(e^{j\omega_k})\}_{k=1}^{K}$ by (18). Denote $E_k = \widetilde{G}(e^{j\omega_k})$.

(2) Construct a function $\widetilde{h}(z) \in H_2(\mathbb{D})$ as the first approximation

$$\widetilde{h}(z) = \frac{1}{2\pi j}\int_0^{2\pi}\frac{\sum_k E_k\chi_{k(\ldots)}(\omega)}{e^{j\omega}}de^{j\omega}. \tag{36}$$

(3) If only 2nd order approximation is needed, we select the poles simultaneously by

$$\begin{aligned}\{\zeta_1^\star, \zeta_2^\star\} &= \max_{\zeta_1,\zeta_2\in\mathbb{D}} |\langle g_1, e_{\zeta_1}\rangle|^2 + |\langle g_2, e_{\zeta_2}\rangle|^2 \\ &= \max_{\zeta_1,\zeta_2\in\mathbb{D}} (1-|\zeta_1|^2)|g_1(z)|^2 + (1-|\zeta_2|^2)|g_2(z)|^2\end{aligned}$$

where $g_1(z) = \widetilde{h}(z)$ and $g_2(z)$ is given by the recursive formula (32), thus a second order approximation could be obtained

$$\widehat{G}(z) = \sum_{l=1}^{2} \langle \widetilde{h}, \mathcal{B}_l\rangle \mathcal{B}_l(z), \tag{37}$$

which is the *2-best rational approximation* of $\widetilde{h}(z)$.

(4) If higher order (more than 2) is requested, AFD algorithm is applied to get the other poles from $k = 3$, thus an arbitrary kth order approximation can be given by

$$\widehat{G}(z) = \sum_{l=1}^{k} \langle \widetilde{h}, \mathcal{B}_l\rangle \mathcal{B}_l(z). \tag{38}$$

Remark 3.1. Note that in step 4, one also can choose to select 2 poles at a time, this will improve the approximating efficiency but increase a little computation complexity of numerical algorithms.

Under maximal selection criterion, a few term can achieve better approximation than Fourier method including Laguerre and Kautz models [23]. As the order k increases, the approximation would become better, theoretically, this method could give approximations in any order. In this algorithm, the selected poles ζ_ls depend on the frequency responses of system itself, for different systems, there are different poles for the basis functions leading to approximations to themselves. Thus this is an adaptive algorithm.

4. Identification of $f(\cdot)/r_1$

Though a Hammerstein system is unidentifiable, once linear subsystem $r_1G(z)$ is fixed, the nonlinear subsystem becomes identifiable under the assumption. We continue to estimate the nonlinear part by reversing the linear subsystem. If $\widehat{G}(z)^{-1}$ is not minimum phase, then inversion is a problem, some methods are developed for this issue in [4,26,31,35]. If $\widehat{G}(z)^{-1}$ is invertible, then the middle signal $\widehat{f}(n)$ could be obtained via

$$\widehat{f}(n) = \widehat{G}^{-1}(z)y^\star(n). \tag{39}$$

As the finite rational basis functions automatically results in a rational function, saying

$$\widehat{G}(z) = \frac{\beta_{m-1}z^{m-1} + \beta_{m-2}z^{m-2} + \cdots + \beta_0}{z^m + \alpha_{m-1}z^{m-1} + \cdots + \alpha_0} (\beta_{m-1} \neq 0),$$

$\widehat{f}(n)$ in time domain is given by

$$\begin{aligned}\widehat{f}(n)(n+m-1) = \beta_{m-1}^{-1}[&-\beta_{m-2}\widehat{f}(n+m-2) - \cdots - \beta_0\widehat{f}(n) \\ &+y^\star(n+m) + \alpha_{m-1}y^\star(n+m-1) + \cdots + \alpha_0 y^\star(n)].\end{aligned} \tag{40}$$

In the case $\{x(n)\}$ and $\{\widehat{f}(n)\}$ are available, parameters of polynomial $\widehat{f}(.)$ can be obtained by using the least-square method. Because the rang of polynomial order is known, we can make use of the maximal order p_2, that is

$$\widehat{f}(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{p_2}x^{p_2} \tag{41}$$

in which there are $p_2 + 1$ parameters $\{\alpha_i\}_{i=0}^{p_2}$. Generally speaking, the data number of $\{x(n)\}$ and $\{\widehat{f}(n)\}$ are great more than the number of parameters, thus it would make the estimate of least square method efficient. However, we shall use all of sampled data sets to repeat the estimate Kth times (K is the number of data sets of input(output) measurements). There is a result about the estimate when $p = 1$.

Lemma 4.1. *Let $y^\star(n) = y(n) + v(n)$ be the measured output. If $G(z)$ and $\widehat{G}(z)$ are invertible, output $y(n)$ is absolute summable, and each data of noise is bounded, i.e, $|v(n)| < \epsilon < \infty$, then we have*

$$\lim_{\substack{K,I\to\infty \\ \epsilon\to 0}} \|f(n) - \widehat{f}(n)\|_1 = 0, \tag{42}$$

where K, I are indexes of frequencies number and approximating terms of basis functions, respectively.

Proof. This result can be proved by using triangle inequality.

$$\begin{aligned}\|f(n)-\widehat{f}(n)\|_1 &= \|G^{-1}(z)y(n)-\widehat{G}(z)^{-1}y^{\star}(n)\|_1\\ &= \|G^{-1}(z)y(n)-\widehat{G}^{-1}(z)y(n)+\widehat{G}^{-1}(z)y(n)-\widehat{G}(z)^{-1}y^{\star}(n)\|_1\\ &\le \|G^{-1}(z)y(n)-\widehat{G}^{-1}(z)y(n)\|_1+\|\widehat{G}^{-1}(z)y(n)-\widehat{G}(z)^{-1}y^{\star}(n)\|_1\\ &\le \|G^{-1}-\widehat{G}^{-1}\|_1\|y(n)\|_1+\|\widehat{G}^{-1}\|\|y(n)-y^{\star}(n)\|_1.\end{aligned}$$

On the right side of the above inequality, as K and I tend to infinity, the first term tends to zero, while ϵ tends to zero, the second term tends to zero too. The proof is complete. □

Above all, the complete algorithm for estimate of nonlinear part under condition can be stated as follows:

(1) Recover middle data sets $\{\widehat{f}^k(n)\}_{k=1}^K$ by $\widehat{f}^k(n)=\widehat{G}(z)^{-1}y^{\star k}(n)$ with $\{y^{\star k}(n)\}$.
(2) Estimate the parameters $\{\alpha_i^1\}_{i=0}^{p_2}$ by using least-square method with $\{x^1(n),\widehat{f}^1(n)\}$.
(3) Repeat this estimate $k=2,\dots,K$ to get $\{\alpha_i^k\}_{i=0}^{p_2}$ by using $\{x^k(n),\widehat{f}^k(n)\}$.
(4) The final estimate of parameters is given by $\{\alpha_i\}_{i=0}^{p_2}$ where

$$\alpha_i=\frac{\alpha_i^1+\alpha_i^2+\cdots+\alpha_i^K}{K}(i=0,1,2,\dots,p_2). \tag{43}$$

Then we obtain a Hammerstein system $(\widehat{f}(x),\widehat{G}(z))$ which has approximating input–output behavior as the original system.

Theorem 4.1. *Let $G(z)\in H_2(\mathbb{D})$ and $f(x)$ be a finite order polynomial. Given an absolute summable discrete input $x(n)$, if $\widehat{G}(z)$ is invertible, there is a result between the output signal $y(n)$ of original Hammerstein system $(f(x),G(z))$ and output signal $y^{\star}(k)$ of identified Hammerstein system $(\widehat{f}(x),\widehat{G}(z))$:*

$$\lim_{\substack{K,I\to\infty\\ N\to\infty}}\|y(n)-y^{\star}(n)\|_2=0. \tag{44}$$

Proof. Denote $\widehat{G}(z)=r_1\widetilde{G}_I(z)$ (I is the order of approximations) and $\widehat{f}(x)=\widetilde{f}(x)/r_1$, where $\widetilde{G}_I(z)$ and $\widetilde{f}(x)$ approach to linear part $G(z)$ and nonlinear parts $f(x)$, respectively. Then we have

$$\begin{aligned}\|y(n)-y^{\star}(n)\|_2 &= \|G(z)\cdot f(x)-\widehat{G}(z)\cdot\widehat{f}(x)\|_2\\ &= \|G(z)\cdot f(x)-r_1\widetilde{G}_I(z)\cdot\widetilde{f}(x)/r_1\|_2\\ &= \|G(z)\cdot f(x)-\widetilde{G}_I(z)\cdot\widetilde{f}(x)\|_2\\ &= \|G(z)\cdot f(x)-\widetilde{G}_I(z)\cdot f(x)+\widetilde{G}_I(z)\cdot f(x)-\widetilde{G}_I(z)\cdot\widetilde{f}(x)\|_2\\ &\le \|G(z)\cdot f(x)-\widetilde{G}_I(z)\cdot f(x)\|_2+\|\widetilde{G}_I(z)\cdot f(x)-\widetilde{G}_I(z)\cdot\widetilde{f}(x)\|_2\\ &= \|(G-\widetilde{G}_I)\cdot f(x)\|+\|\widetilde{G}_I(z)\cdot(f(x)-\widetilde{f}(x)\|_2\\ &\le \|G-\widetilde{G}_I\|_2\|f\|_1+\|\widetilde{G}_I\|_2\|f(x)-\widetilde{f}(x)\|_1.\end{aligned} \tag{45}$$

In the right side of inequality (45), the first term tends to zero as K, I tend to infinity and the second term becomes zero when N rises to infinity, too. The proof is complete. □

Above all, the final algorithm for identification of entire Hammerstein system can be stated as follows;

Step 1: Collect input and output data $\{x^k(n)\}$, $\{y^{\star k}(n)\}$ for each frequency ω_k.
Step 2: Estimate the frequency responses of linear part $\{\omega_k,\widetilde{G}(e^{j\omega_k})\}$.
Step 3: Estimate the linear part $r_1G(z)$ with $\{\omega_k,\widetilde{G}(e^{j\omega_k})\}$ by using the rational orthogonal basis functions.
Step 4: Estimate the nonlinearity $f(x)/r_1$.

A simulation of this algorithm is given in the next section.

5. Example

In this section, we give an example to illustrate the proposed algorithm. The considered single-input and single-output system consists of a linear subsystem $G(z)$ [27]

$$G(z)=\frac{(0.0247z^4+0.0355z^3)}{((1-0.9048z)(1-0.3679z))}, \tag{46}$$

and a nonlinear subsystem

$$f(x)=x+x^2 \tag{47}$$

with known range from $p_1=2$ to $p_2=3$.

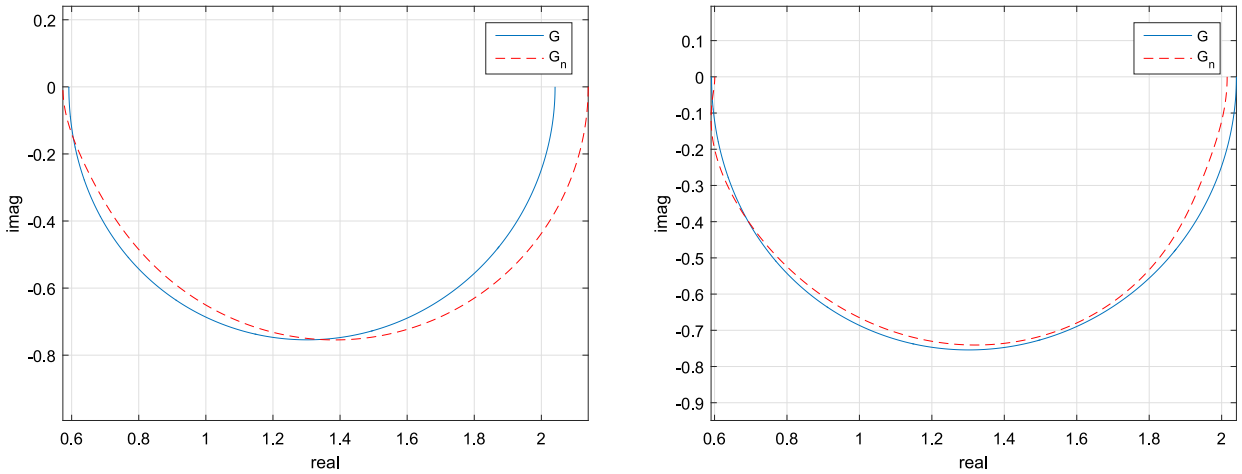

Fig. 2. The 2nd and 4th order approximation of linear part $\widehat{G}/r_1$.

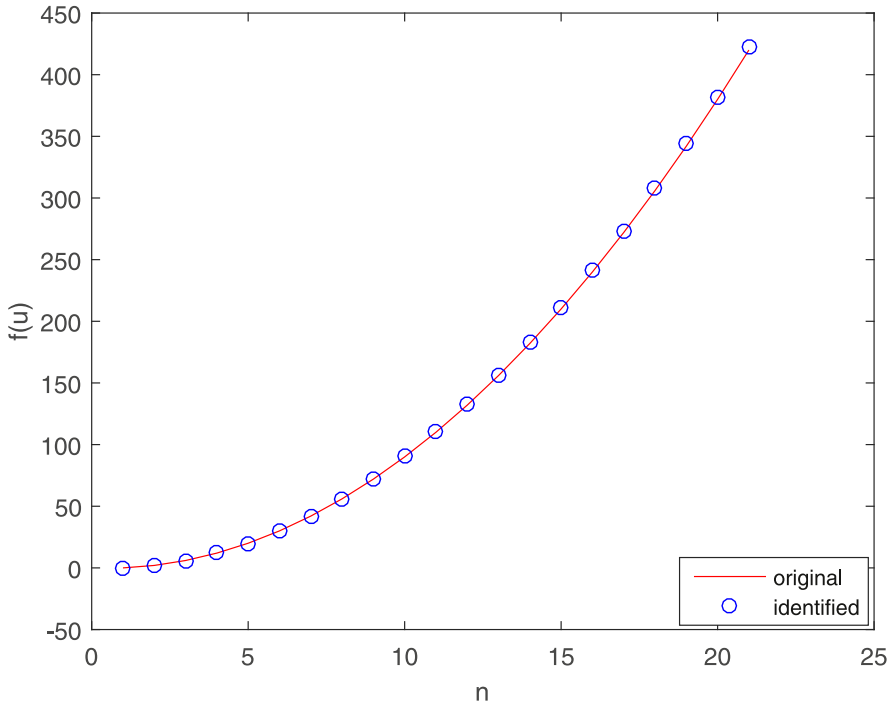

Fig. 3. Nonlinearity.

The input discrete signal is

$$x(n) = 2\cos(\omega_k n), \tag{48}$$

Then by computation, there is $r_1 = 2$.

In the simulation, the data number is set to $K = 500, N = 1000$ and $\omega_k \in [0, \pi)$. By using the estimates of frequency responses of $r_1 G(z)$ with (18), we obtain a set of frequency responses in the half circle $(0, \pi)$, the remaining ones of frequency estimates of $r_1 G(z)$ in the interval $(\pi, 2\pi)$ can be obtained by using the conjugate symmetry of the frequency responses. The presented adaptive algorithm results to an Ith partial sum $\widehat{G}(z)$ with finite I basis functions. In order to give a clear comparison, the figures of $\widehat{G}(z)/r_1$ is considered. Fig. 2 shows the frequency responses of $\widehat{G}(z)/r_1$ at 2nd order and 4th order,

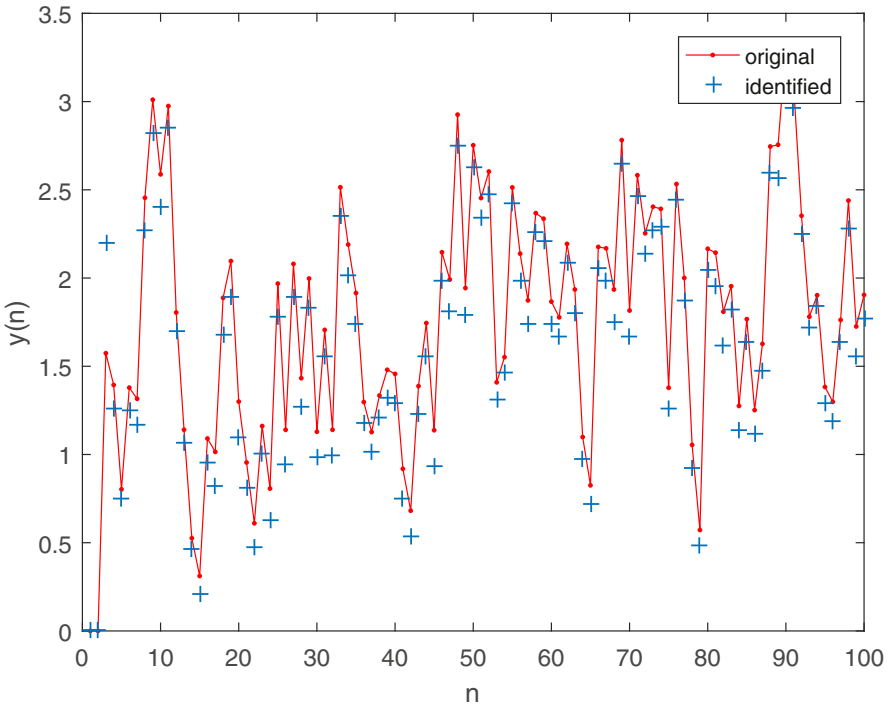

Fig. 4. The entire output behavior at a same random input.

respectively. In particular, the 2nd order approximation is shown as follows:

$$\widehat{G}(z) = \frac{-0.4775z - 0.9927}{0.2144z^2 + 0.1012z - 1}, \tag{49}$$

its zeros are outside the unit circle, then it can be used to recover data sets $\{\widehat{f}^k(n)\}$, which is the output of nonlinear part under input $2\cos(\omega_k n)$. If we use the minimal order of $f(x)$, then a polynomial with 2nd order is obtained,

$$\widehat{f}(x) = 1.0064x^2 + 1.0064x - 0.0439. \tag{50}$$

Fig. 3 shows the curves of $r_1\widehat{f}(x)$ and true $f(x)$, it can be seen that the lines of $r_1\widehat{f}(x)$ and true $f(x)$ are very close to each other.

The entire input–output behavior of original Hammerstein system and identified one is also compared. In Fig. 4, it shows the outputs of $(f(x), G(z))$ and $(\widehat{f}(x), \widehat{G}(z))$ under a same random input signal, it can be seen that the proposed algorithm is efficient.

6. Conclusion

A novel algorithm for identification of Hammerstein systems in frequency domain is proposed in this paper. In the proposed algorithm, finite rational orthogonal basis functions are used, the novelty of this algorithm is that the poles for the basis functions are selected adaptively, then the corresponding used finite rational orthogonal basis functions are selected adaptively. For different systems, this algorithm leads to adaptive approximations. Once approaches to the linear part are obtained, the estimate of left nonlinear part will be done. From the figures and numerical results of simulation, we can see this algorithm is efficient.

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] E.W. Bai, Decoupling the linear and nonlinear parts in Hammerstein model identification, Automatica 40 (2004) 671–676.
[5] D. Bauer, B. Ninness, Asymptotic properties of Hammerstein model estimates, in: Proceedings of the IEEE Conference on Decision and Control, Sydney, 2000, pp. 2855–2860.
[6] S. Baumgartner, W. Rugh, Complete identification of a class of nonlinear systems from steady state frequency response, IEEE Trans. Circuites Syst. CAS-22 (1975) 753–759.

[7] M. Boutayeb, H. Rafaralahy, M. Darouach, A robust and recursive identification method for Hammerstein model, in: Proceedings of the IFAC World Congress, CA, San Francisco, 1996, pp. 447–452.
[8] A. Bultheel, P. González-Vera, E. Hendriksen, O. Njaatad, Orthogonal rational functions, Cambridge Monographs on Applied and Computational Mathematics, 5, Cambrige University Press, 1999.
[9] Q. Chen, W. Mai, L. Zhang, W. Mi, System identification by discrete rational atoms, Automatica 56 (2015) 53–59.
[10] P.V. Gucht, A. Bultheel, Orthonormal rational functions for system identification: numerical aspects, IEEE Trans. Automat. Control 48 (2003) 705–709.
[11] R.A. Devore, V.N. Temlyakov, Some remarks on greedy algorithm, Adv. Comput. Math. 5 (1996) 173–187.
[12] W. Favoreel, B. De Moor, Subspace identification of bilinear systems, IEEE Trans. Autom. Control 44 (1999) 1157–1165.
[13] A. Gardiner, Frequency domain identification of nonlinear systems, in: Proceedings of the Third IFAC Symposium on Identification and System Parameter Estimation, Hague, The Netherlands, 1973, pp. 831–834.
[14] W. Greblichi, Continuous-time Hammerstein system identification, IEEE Trans. Autom. Control 45 (2000) 1232–1236.
[15] J.C. Gacuteomez, E. Baeyens, Identification of block-oriented nonlinear systems using orthonormal bases, J. Process Control 14 (2004) 685–697.
[16] P.S.C. Heuberger, P.M.J. Van den Hof, B. Wahlberg, Modelling and Identification with Rational Orthogonal Basis Functions, Springer, London, 2005.
[17] W. Kang, S. Zhong, K. Shi, Triple integral approach to reachable set bounding for linear singular systems with time-varying delay, Math. Methods Appl. Sci. 40 (2017) 2949–2960.
[18] J. Luo, W. Tian, S. Zhong, K. Shi, H. Chen, X.M. Gu, Non-fragile asynchronous H_∞ control for uncertain stochastic memory systems with bernoulli distribution, Appl. Math. Comput. 312 (2017) 109–128.
[19] L. Ljung, System Identification: Theory for the Users, second ed., Prentice-Hall, Upper Saddle River, NJ, 1999.
[20] W. Liu, S. Zhong, Web malware spread modelling and optimal control strategies, Sci. Rep. 7 (2017) 42308–42327.
[21] S. Long, S. Zhong, Improved results for stochastic stabilization of a class of discrete-time singular Markovian jump systems with time-varying delay, Nonlinear Anal. Hybrid Syst. 23 (2017) 11–26.
[22] W. Mi, T. Qian, F. Wan, A fast adaptive model reduction method based on Takenaka–Malmquist systems, Syst. Control Lett. 61 (2012) 223–230.
[23] W. Mi, T. Qian, Frequency domain identification: an adaptive algorithm based on rational orthogonal system, Automatica 48 (6) (2012) 1154–1162.
[24] W. Mi, T. Qian, On backward shift algorithm for estimating poles of systems, Automatica 50 (2014) 1603–1610.
[25] W. Mi, T. Qian, S. Li, Basis pursuit for frequency-domain identificaiton, Math. Methods Appl. Sci. 39(3).
[26] K.S. Narendra, P. G. Gallman, In iterative method for the identification of nonlinear systems using a Hammerstein model, IEEE Trans. Autom. Control 11 (7) (1966) 546–550.
[27] B. Ninness, F. Gustafsson, A unifying construction of orthonormal bases for system identification, IEEE Trans. Autom. Control 42 (1997) 512–515.
[28] R. Pintelon, J. Schoukens, System Identification: A Frequency Domain Approach, Second Edition, John Wiley & Sons, Inc., Hoboken, New Jersey, 2012.
[29] T. Qian, Y.B. Wang, Adaptive decomposition into basic signals of non-negative instantaneous frequencies – a variation and realization of greedy algorithm, Adv. Comput. Math. 34 (3) (2010) 279–293.
[30] T. Qian, E. Wegert, Optimal approximation by Blaschke forms, Complex Var. Elliptic Equ. 58 (1) (2013) 123–133.
[31] S. Rangan, G. Wolodkin, K. Poolla, Identification methods for Hammerstein systems, in: Proceedings of the IEEE Conference on Decision Control, New Orleans, LA, 1995, pp. 697–702.
[32] P. Stoica, On the convergence of an iterative algorithm used for Hammerstein system identification, IEEE Trans. Autom. Control AC–26 (1981) 967–969.
[33] J. Schoukens, R. Pintelon, Y. Rolain, Time domain identification, frequency domain identification. Equivalencies! Differences? Proceedings of the 2004 American Control Conference, Massachusetts, Boston, 2004, pp. 661–666.
[34] G. Vandersteen, Y. Rolain, J. Schoukens, Nonparametric estimation of the frequency response function of the linear blocks of a Wiener–Hammerstein models, Automatica 33 (1997) 1151–1155.
[35] J. Voros, Parameter identification of discontinuous Hammerstein systems, Automatica 33 (6) (1997) 1141–1146.
[36] B. Wahlberg, System identification using Laguerre models, IEEE Trans. Autom. Control 36 (5) (1991) 551–562.
[37] B. Wahlberg, System identification using Kautz models, IEEE Trans. Autom. Control 39 (6) (1994) 1276–1282.
[38] J.L. Walsh, Interpolation and Approximation by Rational Functions in the Complex Domain, fourth ed., American Mathematical Society, 1965.
[39] Q. Wang, N.R. He, X.J. Chen, Replicator dynamics for public goods game with resource allocation in large populations, Appl. Math. Comput. 328 (2018) 162–170.
[40] L. Zadeh, On the identification problem, IRE Trans. Circuit Theory 3 (1956) 277–281.