algorithms

MDPI

Article

Auxiliary Model Based Multi-Innovation Stochastic Gradient Identification Algorithm for Periodically Non-Uniformly Sampled-Data Hammerstein Systems

Li Xie [1,2,*] and Huizhong Yang [1,2]

1 Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122, China; yhz@jiangnan.edu.cn

2 School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China

* Correspondence: xieli@jiangnan.edu.cn; Tel.: +86-134-8505-3600

Received: 12 May 2017; Accepted: 19 July 2017; Published: 31 July 2017

Abstract: Due to the lack of powerful model description methods, the identification of Hammerstein systems based on the non-uniform input-output dataset remains a challenging problem. This paper introduces a time-varying backward shift operator to describe periodically non-uniformly sampled-data Hammerstein systems, which can simplify the structure of the lifted models using the traditional lifting technique. Furthermore, an auxiliary model-based multi-innovation stochastic gradient algorithm is presented to estimate the parameters involved in the linear and nonlinear blocks. The simulation results confirm that the proposed algorithm is effective and can achieve a high estimation performance.

Keywords: non-uniform sampling; Hammerstein system; parameter estimation; multi-innovation theory; stochastic gradient algorithm

1. Introduction

The dynamics of most practical systems are inherently nonlinear due to complex physical, chemical and biological mechanisms. The modeling of nonlinear systems is challenging and has become an active research area in both academia and industry [1,2]. To simplify the modeling problem, block-oriented models, which are composed of linear dynamic blocks in combination with nonlinear memoryless blocks, have been widely utilized to describe nonlinear systems. The state of the art in designing, analyzing and implementing identification algorithms for block-oriented nonlinear systems were well summarized in a recent book by Giri and Bai [3]. Depending on the location of the static nonlinear component, block-oriented models can be classified into the Hammerstein model, the Wiener model and the Hammerstein–Wiener model [4–6]. The Hammerstein model represents a class of input nonlinear systems, where the nonlinear block is prior to the linear one. It can flexibly approximate various input nonlinearities, such as saturation, dead zone, backlash and hysteresis, thus having been extensively employed in realistic applications [7–11].

For decades, the identification of Hammerstein nonlinear systems has attracted much attention, and numerous methods have been reported in the literature. For example, Pouliquen et al. studied the parameter estimation of Hammerstein systems where the linear part is described by an output error model and presented an iterative algorithm based on the optimal bounding ellipsoid criterion [12]. Ding et al. applied the auxiliary model identification principle to deal with unmeasurable noise-free outputs in Hammerstein output error systems, presented a recursive least squares (RLS) algorithm and investigated its convergence properties [13]. Filipovic derived a robust extended RLS algorithm to estimate the parameters of Hammerstein systems interfered by non-Gaussian disturbance [14]. Gao et al. proposed a blind identification algorithm for Hammerstein systems with hysteresis nonlinearity and further developed a composite control strategy to track the reference input [15].

Among various new methodologies in the identification area, the multi-innovation theory has been considered as a useful way to improve estimation precision and convergence rate. The basic idea of the multi-innovation theory is innovation expanding, which helps to update the parameter estimates at each recursion using the data over a moving and fixed-size window [16,17]. Typically, the multi-innovation theory is incorporated with the RLS algorithm, the stochastic gradient (SG) algorithm, the stochastic Newton recursive algorithm, etc., to address identification problems [18]. For instance, a multi-innovation RLS algorithm was developed for Hammerstein AutoRegressive eXogenous (ARX) systems with backlash nonlinearity [19]. Furthermore, a multi-innovation SG algorithm [20] and an auxiliary model-based multi-innovation generalized extended SG algorithm [21] were proposed to estimate the parameters of Hammerstein nonlinear ARX and Box–Jenkins systems, respectively. Compared with the multi-innovation RLS algorithm, the multi-innovation SG algorithm is more efficient in computation because it avoids performing large matrix inversion [22].

The above-mentioned Hammerstein systems all belong to single-rate systems, the inputs and outputs of which are uniformly sampled at the same rate. However, non-uniform sampling can be encountered in practice due to hardware limitations or economic considerations [23–25]. For example, influenced by transmission delays and packet losses, the input-output data in networked control systems might be available at non-uniformly spaced time instants [26,27]. The non-uniform sampling includes the uniform sampling as its special case, which can always preserve controllability and observability in discretization [28]. Furthermore, it can overcome the restriction of the Nyquist limit and enable a much lower average sampling frequency. Therefore, intentional non-uniform sampling has the potential to reduce the hardware cost in control applications [29]. Due to the complexity of arbitrary non-uniform sampling, most of the literature works have focused on periodically non-uniformly sampled-data systems [30,31]. For periodically non-uniformly sampled-data Hammerstein systems, Li et al. derived the lifted transfer function model by means of the lifting technique and presented a least squares-based iterative algorithm for parameter estimation [32]. The lifting technique is a benchmark tool to deal with multirate and non-uniformly sampled-data systems [33,34]. However, the corresponding lifted models are complex and involve a large number of parameters, which brings a great challenge for identification. To simplify the model structure and reduce the identification complexity, Xie et al. put forward a novel input-output representation of linear systems with non-uniform sampling by introducing a time-varying backward shift operator δ^{-1} [35]. On the basis of that work, this paper aims to propose a δ^{-1}-based model to describe periodically non-uniformly sampled-data Hammerstein systems and presents an auxiliary model-based multi-innovation SG algorithm to estimate the model parameters.

The rest of this paper is organized as follows. Section 2 formulates the identification problem of periodically non-uniformly sampled-data Hammerstein systems. Identification algorithms are proposed in Section 3, and an example is provided in Section 4 to examine their estimation performance. Finally, concluding remarks are given in Section 5.

2. Problem Description

Consider a periodically non-uniformly sampled-data Hammerstein system as depicted in Figure 1, in which H_τ is a periodic non-uniform zero-order hold, converting a discrete-time input sequence $\{u(kT+t_i)\}$ to a continuous-time input $u(t)$, i.e.,

$$u(t)=\begin{cases} u(kT+t_0), & kT+t_0\le t<kT+t_1\ (t_0=0),\\ u(kT+t_1), & kT+t_1\le t<kT+t_2,\\ \vdots & \\ u(kT+t_{q-1}), & kT+t_{q-1}\le t<kT+t_q\ (t_q=T),\end{cases}$$

where $k = 0, 1, \cdots$; T is the frame period spaced by q non-uniform sampling instants t_i $(i = 0, 1, \cdots, q-1)$.

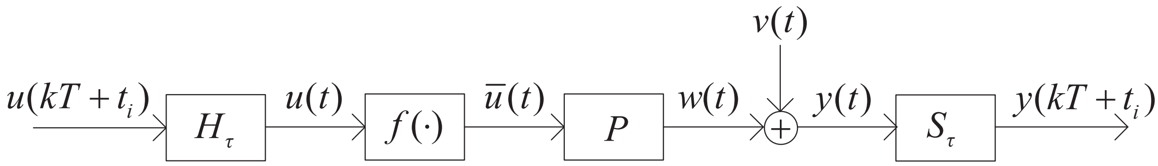

Figure 1. Periodically non-uniformly sampled-data Hammerstein system.

By passing through a nonlinear static block $f(\cdot)$, $u(t)$ is transformed into an unmeasurable inner input $\overline{u}(t)$ to a linear dynamic process P with order n, which can be expressed as:

$$\overline{u}(t) = \sum_{m=1}^{n_c} c_m f_m[u(t)], \tag{1}$$

where $f_m[u(t)]$ are known nonlinear basis functions and c_m are unknown coefficients to be estimated.

The noise-free output $w(t)$ of the process P is corrupted by a white noise $v(t)$, generating a measurable output $y(t)$. The non-uniform sampler S_τ has a synchronous sampling pattern with H_τ; thus, the discrete-time output sequence $\{y(kT + t_i)\}$ is obtained at sampling instants $t = kT + t_i$ $(i = 0, 1, \cdots, q-1)$.

For the notational simplicity, in the following the data $s(kT + t_i)$, at non-uniform sampling time $t = kT + t_i$, is denoted by $s_i(k)$. By using a time-varying backward shift operator δ^{-1} $[\delta^{-1}s_i(k) = s_{i-1}(k)]$ proposed in [35], the mapping relationship between the inner input $\overline{u}_i(k)$ and the noise-free output $w_i(k)$ of the periodically non-uniformly sampled-data Hammerstein system can be represented as:

$$w_i(k) = \frac{B_i(\delta)}{A_i(\delta)}\overline{u}_i(k),\ i = 0, 1, 2, \cdots, q-1, \tag{2}$$

where:

$$A_i(\delta) := 1 + a_{i1}\delta^{-1} + a_{i2}\delta^{-2} + \cdots + a_{in}\delta^{-n},$$
$$B_i(\delta) := b_{i0} + b_{i1}\delta^{-1} + b_{i2}\delta^{-2} + \cdots + b_{in}\delta^{-n}.$$

From the system schematic diagram in Figure 1, we have:

$$y_i(k) = w_i(k) + v_i(k). \tag{3}$$

Given the non-uniformly-sampled input-output data $\{u_i(k), y_i(k), k = 0, 1, 2, \cdots, i = 0, 1, 2, \cdots, q-1\}$, the objective of this paper is to estimate the parameters of the nonlinear block in (1):

$$\boldsymbol{c} := [c_1, c_2, \cdots, c_{n_c}]^{\mathrm{T}} \in \mathbb{R}^{n_c},$$

and the parameters of the linear block in (2):

$$\boldsymbol{a}_i := [a_{i1}, a_{i2}, \cdots, a_{in}]^{\mathrm{T}} \in \mathbb{R}^{n},$$
$$\boldsymbol{b}_i := [b_{i0}, b_{i1}, b_{i2}, \cdots, b_{in}]^{\mathrm{T}} \in \mathbb{R}^{n+1}.$$

3. Identification Algorithms

3.1. The AM-SG Algorithm

According to the over-parameterized linear regression approach [36], define the information vector $\boldsymbol{\varphi}_i(k)$ and the parameter vector $\boldsymbol{\theta}_i$ as:

$$\boldsymbol{\varphi}_i(k) := \begin{bmatrix} \boldsymbol{\varphi}_{iw}(k) \\ \boldsymbol{\varphi}_{iu}(k) \end{bmatrix} \in \mathbb{R}^{n_0}, \quad \boldsymbol{\theta}_i := \begin{bmatrix} \boldsymbol{a}_i \\ \boldsymbol{\theta}_{iu} \end{bmatrix} \in \mathbb{R}^{n_0},\ n_0 = n + (n+1)n_c,$$

$$\boldsymbol{\varphi}_{iw}(k) := \begin{bmatrix} -w_{i-1}(k) \\ -w_{i-2}(k) \\ \vdots \\ -w_{i-n}(k) \end{bmatrix} \in \mathbb{R}^{n}, \quad \boldsymbol{\phi}_{ij}(k) := \begin{bmatrix} f_1[u_{i-j}(k)] \\ f_2[u_{i-j}(k)] \\ \vdots \\ f_{n_c}[u_{i-j}(k)] \end{bmatrix} \in \mathbb{R}^{n_c},\ j = 0,\ 1,\ 2, \cdots,\ n,$$

$$\boldsymbol{\varphi}_{iu}(k) := \begin{bmatrix} \boldsymbol{\phi}_{i0}(k) \\ \boldsymbol{\phi}_{i1}(k) \\ \boldsymbol{\phi}_{i2}(k) \\ \vdots \\ \boldsymbol{\phi}_{in}(k) \end{bmatrix} \in \mathbb{R}^{(n+1)n_c}, \quad \boldsymbol{\theta}_{iu} := \begin{bmatrix} b_{i0}\boldsymbol{c} \\ b_{i1}\boldsymbol{c} \\ b_{i2}\boldsymbol{c} \\ \vdots \\ b_{in}\boldsymbol{c} \end{bmatrix} \in \mathbb{R}^{(n+1)n_c}.$$

Using Equation (1), Equation (2) can be written in the following vector form:

$$\begin{aligned} w_i(k) &= -\sum_{j=1}^{n} a_{ij} w_{i-j}(k) + \sum_{j=0}^{n} b_{ij} \left[\sum_{m=1}^{n_c} c_m f_m[u_{i-j}(k)] \right] \\ &= \boldsymbol{\varphi}_i^{\mathrm{T}}(k)\boldsymbol{\theta}_i. \end{aligned} \tag{4}$$

Substituting Equation (4) into Equation (3), we have:

$$y_i(k) = \boldsymbol{\varphi}_i^{\mathrm{T}}(k)\boldsymbol{\theta}_i + v_i(k). \tag{5}$$

Equation (5) is the identification model of the periodically non-uniformly sampled-data Hammerstein system, in which the parameter vector $\boldsymbol{\theta}_i$ includes the products of parameters b_{ij} and c_m. To guarantee a unique parametrization, the first coefficient c_1 of the nonlinear function is assumed to be one [14,37]. Furthermore, the information vector $\boldsymbol{\varphi}_i(k)$ contains unmeasurable noise-free outputs $w_{i-j}(k)$. A solution to this difficulty is to replace $w_{i-j}(k)$ with their estimates $\hat{w}_{i-j}(k)$ based on the auxiliary model identification idea [38–40]. Accordingly, define the estimate of $\boldsymbol{\varphi}_i(k)$ as:

$$\begin{aligned} \hat{\boldsymbol{\varphi}}_i(k) &:= \begin{bmatrix} \hat{\boldsymbol{\varphi}}_{iw}(k) \\ \boldsymbol{\varphi}_{iu}(k) \end{bmatrix}, \\ \hat{\boldsymbol{\varphi}}_{iw}(k) &:= [-\hat{w}_{i-1}(k),\ -\hat{w}_{i-2}(k),\ \cdots,\ -\hat{w}_{i-n}(k)]^{\mathrm{T}}. \end{aligned}$$

Using the estimates $\hat{\boldsymbol{\varphi}}_i(k)$ and $\hat{\boldsymbol{\theta}}_i(k)$ to replace $\boldsymbol{\varphi}_i(k)$ and $\boldsymbol{\theta}_i$ in (4), respectively, yields

$$\hat{w}_i(k) = \hat{\boldsymbol{\varphi}}_i^{\mathrm{T}}(k)\hat{\boldsymbol{\theta}}_i(k). \tag{6}$$

Applying the stochastic gradient search principle to minimize the following criterion function:

$$J_1(\boldsymbol{\theta}_i) = \mathrm{E}\left\{ [y_i(k) - \hat{\boldsymbol{\varphi}}_i^{\mathrm{T}}(k)\boldsymbol{\theta}_i]^2 \right\},$$

an auxiliary model-based stochastic gradient (AM-SG) algorithm can be derived for identifying the parameter vector $\boldsymbol{\theta}_i$ in (5):

$$\hat{\boldsymbol{\theta}}_i(k) = \hat{\boldsymbol{\theta}}_i(k-1) + \frac{\hat{\boldsymbol{\varphi}}_i(k)}{r_i(k)}[y_i(k) - \hat{\boldsymbol{\varphi}}_i^{\mathrm{T}}(k)\hat{\boldsymbol{\theta}}_i(k-1)], \tag{7}$$

$$r_i(k) = r_i(k-1) + \|\hat{\boldsymbol{\varphi}}_i(k)\|^2,\ r_i(0) = 1. \tag{8}$$

3.2. The AM-MISG Algorithm

The AM-SG algorithm only uses the current dataset to update the parameter estimates. Therefore, it has a slow convergence rate and low estimation accuracy. To improve the identification performance of the AM-SG algorithm, an innovation length p is introduced to derive an auxiliary model-based multi-innovation stochastic gradient (AM-MISG) algorithm.

Considering the most recent p sets of input-output data, define the stacked output vector $\boldsymbol{Y}_i(p,k)$, the stacked noise vector $\boldsymbol{V}_i(p,k)$ and the stacked information matrix $\boldsymbol{\Psi}_i(p,k)$ as:

$$\begin{aligned}
\boldsymbol{Y}_i(p,k) &= [y_i(k),\ y_i(k-1),\ \cdots,\ y_i(k-p+1)]^{\mathrm{T}} \in \mathbb{R}^p,\\
\boldsymbol{V}_i(p,k) &= [v_i(k),\ v_i(k-1),\ \cdots,\ v_i(k-p+1)]^{\mathrm{T}} \in \mathbb{R}^p,\\
\boldsymbol{\Psi}_i(p,k) &= [\boldsymbol{\varphi}_i(k),\ \boldsymbol{\varphi}_i(k-1),\ \cdots,\ \boldsymbol{\varphi}_i(k-p+1)] \in \mathbb{R}^{n_0\times p}.
\end{aligned}$$

Equation (5) can be expanded into the following matrix form:

$$\boldsymbol{Y}_i(p,k) = \boldsymbol{\Psi}_i^{\mathrm{T}}(p,k)\boldsymbol{\theta}_i + \boldsymbol{V}_i(p,k). \tag{9}$$

However, the information vectors $\boldsymbol{\varphi}_i(k-l),\ l = 0,\ 1,\cdots,p-1$ in $\boldsymbol{\Psi}_i(p,k)$ include unknown noise-free outputs. Let $\hat{\boldsymbol{\varphi}}_i(k-l)$ be their estimates; the estimate of $\boldsymbol{\Psi}_i(p,k)$ can be defined as:

$$\hat{\boldsymbol{\Psi}}_i(p,k) = [\hat{\boldsymbol{\varphi}}_i(k),\ \hat{\boldsymbol{\varphi}}_i(k-1),\ \cdots,\ \hat{\boldsymbol{\varphi}}_i(k-p+1)] \in \mathbb{R}^{n_0\times p}.$$

Define the following criterion function:

$$J_2(\boldsymbol{\theta}_i) = \|\boldsymbol{V}_i(p,k)\|^2 = \left\|\boldsymbol{Y}_i(p,k) - \hat{\boldsymbol{\Psi}}_i^{\mathrm{T}}(p,k)\boldsymbol{\theta}_i\right\|^2, \tag{10}$$

where $\|\boldsymbol{X}\|^2 = \mathrm{tr}[\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}]$ represents the norm of the matrix $\boldsymbol{X}$. The gradient of $J_2(\boldsymbol{\theta}_i)$ with respect to $\boldsymbol{\theta}_i$ is given by:

$$\mathrm{grad}[J_2(\boldsymbol{\theta}_i)] = \frac{\partial J_2(\boldsymbol{\theta}_i)}{\partial \boldsymbol{\theta}_i} = -2\hat{\boldsymbol{\Psi}}_i(p,k)[\boldsymbol{Y}_i(p,k) - \hat{\boldsymbol{\Psi}}_i^{\mathrm{T}}(p,k)\boldsymbol{\theta}_i].$$

Applying the stochastic gradient search principle to minimize the criterion function in (10), we have:

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_i(k) &= \hat{\boldsymbol{\theta}}_i(k-1) - \mu_i(k)\mathrm{grad}J_2[\hat{\boldsymbol{\theta}}_i(k-1)]\\
&= \hat{\boldsymbol{\theta}}_i(k-1) + 2\mu_i(k)\hat{\boldsymbol{\Psi}}_i(p,k)[\boldsymbol{Y}_i(p,k) - \hat{\boldsymbol{\Psi}}_i^{\mathrm{T}}(p,k)\hat{\boldsymbol{\theta}}_i(k-1)],
\end{aligned} \tag{11}$$

where $\mu_i(k) > 0$ is called the step size or the convergence factor. For the convenience of formula derivation, let $\mu_i(k) = \frac{1}{2r_i(k)}$; Equation (11) can be rewritten into:

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_i(k) &= \hat{\boldsymbol{\theta}}_i(k-1) + \frac{1}{r_i(k)}\hat{\boldsymbol{\Psi}}_i(p,k)[\boldsymbol{Y}_i(p,k) - \hat{\boldsymbol{\Psi}}_i^{\mathrm{T}}(p,k)\hat{\boldsymbol{\theta}}_i(k-1)]\\
&= \left[\boldsymbol{I} - \frac{1}{r_i(k)}\hat{\boldsymbol{\Psi}}_i(p,k)\hat{\boldsymbol{\Psi}}_i^{\mathrm{T}}(p,k)\right]\hat{\boldsymbol{\theta}}_i(k-1) + \frac{1}{r_i(k)}\hat{\boldsymbol{\Psi}}_i(p,k)\boldsymbol{Y}_i(p,k).
\end{aligned} \tag{12}$$

To guarantee the convergence of this recursive algorithm, all eigenvalues of the matrix $\left[\boldsymbol{I}-\frac{1}{r_i(k)}\hat{\boldsymbol{\Upsilon}}_i(p,k)\hat{\boldsymbol{\Upsilon}}_i^{\mathrm{T}}(p,k)\right]$ should be located inside the unit circle. Therefore, a conservative choice of $\frac{1}{r_i(k)}$ is:

$$0<\frac{1}{r_i(k)}\leqslant\frac{1}{\lambda_{\max}[\hat{\boldsymbol{\Upsilon}}_i(p,k)\hat{\boldsymbol{\Upsilon}}_i^{\mathrm{T}}(p,k)]}.$$

In this paper, we take a common choice:

$$r_i(k)=\lambda_i r_i(k-1)+\|\hat{\boldsymbol{\varphi}}_i(k)\|^2,\ r_i(0)=1, \tag{13}$$

where $\lambda_i\in(0,\ 1]$ is the forgetting factor.

Substituting Equation (13) into Equation (12), the auxiliary model-based multi-innovation stochastic gradient (AM-MISG) algorithm can be derived:

$$\hat{\boldsymbol{\theta}}_i(k)=\hat{\boldsymbol{\theta}}_i(k-1)+\frac{\hat{\boldsymbol{\Upsilon}}_i(p,k)}{r_i(k)}[\boldsymbol{Y}_i(p,k)-\hat{\boldsymbol{\Upsilon}}_i^{\mathrm{T}}(p,k)\hat{\boldsymbol{\theta}}_i(k-1)], \tag{14}$$

$$r_i(k)=\lambda_i r_i(k-1)+\|\hat{\boldsymbol{\varphi}}_i(k)\|^2,\ r_i(0)=1, \tag{15}$$

$$\boldsymbol{Y}_i(p,k)=[y_i(k),\ y_i(k-1),\ \cdots,\ y_i(k-p+1)]^{\mathrm{T}}, \tag{16}$$

$$\hat{\boldsymbol{\Upsilon}}_i(p,k)=[\hat{\boldsymbol{\varphi}}_i(k),\ \hat{\boldsymbol{\varphi}}_i(k-1),\ \cdots,\ \hat{\boldsymbol{\varphi}}_i(k-p+1)], \tag{17}$$

$$\hat{\boldsymbol{\varphi}}_i(k)=\begin{bmatrix}\hat{\boldsymbol{\varphi}}_{iw}(k)\\ \boldsymbol{\varphi}_{iu}(k)\end{bmatrix}, \tag{18}$$

$$\hat{\boldsymbol{\varphi}}_{iw}(k)=[-\hat{w}_{i-1}(k),\ -\hat{w}_{i-2}(k),\ \cdots,\ -\hat{w}_{i-n}(k)]^{\mathrm{T}}, \tag{19}$$

$$\boldsymbol{\varphi}_{iu}(k)=[\boldsymbol{\phi}_{i0}^{\mathrm{T}}(k),\ \boldsymbol{\phi}_{i1}^{\mathrm{T}}(k),\ \boldsymbol{\phi}_{i2}^{\mathrm{T}}(k),\ \cdots,\ \boldsymbol{\phi}_{in}^{\mathrm{T}}(k)]^{\mathrm{T}}, \tag{20}$$

$$\boldsymbol{\phi}_{ij}(k)=[f_1[u_{i-j}(k)],\ f_2[u_{i-j}(k)],\ \cdots,\ f_{n_c}[u_{i-j}(k)]]^{\mathrm{T}},\ j=0,\ 1,\ 2,\cdots,\ n, \tag{21}$$

$$\hat{w}_i(k)=\hat{\boldsymbol{\varphi}}_i^{\mathrm{T}}(k)\hat{\boldsymbol{\theta}}_i(k). \tag{22}$$

Since $c_1=1$, the estimates of a_{ij} and b_{ij} can be directly read from $\hat{\boldsymbol{\theta}}_i(k)$,

$$\hat{a}_{ij}(k)=\hat{\boldsymbol{\theta}}_{i,j}(k),\ j=1,\ 2,\ \cdots,\ n, \tag{23}$$

$$\hat{b}_{ij}(k)=\hat{\boldsymbol{\theta}}_{i,n+n_cj+1}(k),\ j=0,\ 1,\ 2,\cdots,\ n, \tag{24}$$

where $\hat{\boldsymbol{\theta}}_{i,j}(k)$ represents the j-th element of $\hat{\boldsymbol{\theta}}_i(k)$. Note that c_m $(m=2,\ 3,\ \cdots,\ n_c)$ has been estimated $n+1$ times at each non-uniform sampling instant $kT+t_i$ $(i=0,\ 1,\ 2,\cdots,\ q-1)$. Therefore, we can simply take their average like in [36] as the estimate of c_m over the k-th frame period, i.e.,

$$\hat{c}_m(k)=\frac{1}{q(n+1)}\sum_{i=0}^{q-1}\sum_{j=0}^{n}\frac{\hat{\boldsymbol{\theta}}_{i,n+n_cj+m}(k)}{\hat{b}_{ij}(k)},\ m=2,\ 3,\ \cdots,\ n_c. \tag{25}$$

Furthermore, a more numerically-sound SVD-based approach proposed by Bai [41] can be applied to obtain the estimates of b_{ij} and c_m.

The flowchart of the AM-MISG algorithm in (14)–(25) for computing the parameter estimates of periodically non-uniformly sampled-data Hammerstein systems can be illustrated in Figure 2, and the detailed implementation steps are summarized as follows:

1. Initialization: Choose the data length L, the innovation length p and the forgetting factor λ_i; give the nonlinear basis functions $\{f_m(\cdot),\ m=1,\ 2,\ \cdots,\ n_c\}$; set $u_i(k)=0$, $y_i(k)=0$, $\hat{w}_i(k)=0$ for $k\leqslant 0$ and $i=0,\ 1,\ 2,\ \cdots,\ q-1$; take the initial values to be $\hat{\boldsymbol{\theta}}_i(0)=\mathbf{1}_{n_0}/p_0$, where $p_0=10^6$ and $\mathbf{1}_{n_0}$ is a column vector of ones; let $k=1$ and $i=0$.
2. Collect the non-uniformly sampled input-output data $u_i(k)$ and $y_i(k)$.

3. Calculate $f_m[u_i(k)]$ based on $u_i(k)$; form $\hat{\boldsymbol{\varphi}}_{iw}(k)$, $\boldsymbol{\varphi}_{iu}(k)$, $\boldsymbol{\phi}_{ij}(k)$ by (19)–(21); and construct $\hat{\boldsymbol{\varphi}}_i(k)$ by (18).
4. Form the stacked output vector $\mathbf{Y}_i(p,k)$ and the stacked information matrix $\hat{\boldsymbol{\Psi}}_i(p,k)$ by (16) and (17), respectively.
5. Compute the step size $r_i(k)$ by (15) and update the parameter estimate $\hat{\boldsymbol{\theta}}_i(k)$ by (14); calculate $\hat{w}_i(k)$ by (22); obtain $\hat{a}_{ij}(k)$ and $\hat{b}_{ij}(k)$ based on (23) and (24), respectively.
6. If $i < q-1$, then increase i by one, and go to Step 2; otherwise, compute $\hat{c}_m(k)$, $m = 2, 3, \cdots, n_c$ by (25); let $i = 0$ and go to the next step.
7. If $k < L$, then increase k by one, and go to Step 2; otherwise, terminate the computing process.

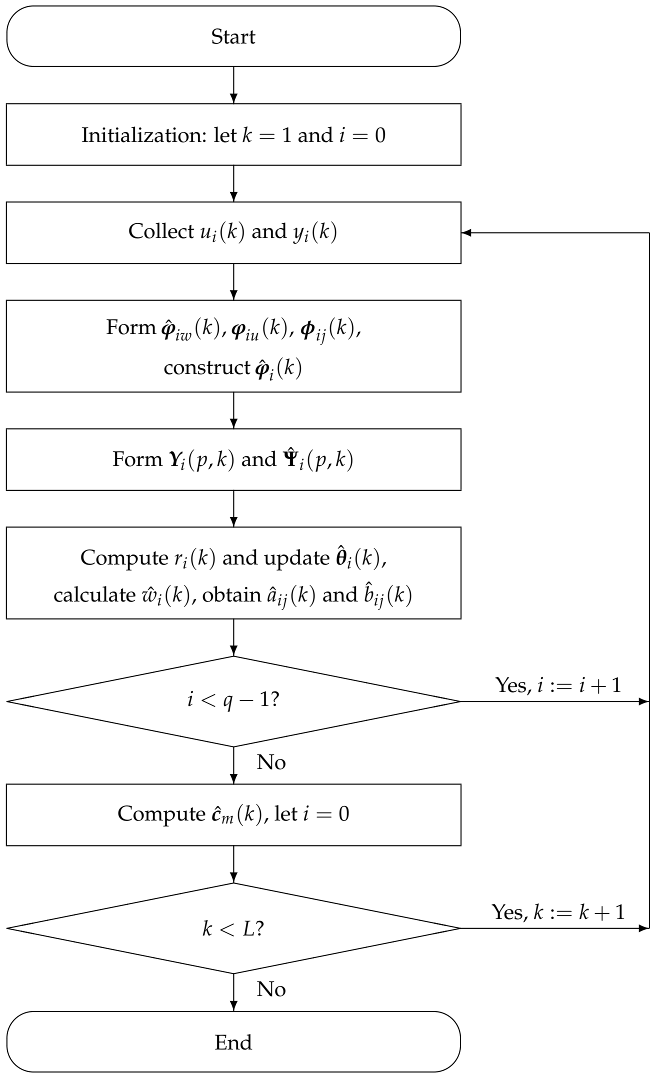

Figure 2. The flowchart of computing the parameter estimate.

3.3. The Main Convergence Result

The main convergence result of the proposed AM-MISG algorithm for periodically non-uniformly sampled-data Hammerstein systems is given in the following theorem.

Theorem 1. *Assume that the noise sequences* $v_i(k)$ $(i = 0, 1, 2, \cdots, q-1)$ *satisfy:*

(A1) $\mathrm{E}[v_i(k)] = 0$, $\mathrm{E}[v_i^2(k)] \leqslant \sigma^2 < \infty$, $\mathrm{E}[v_i(k)v_i(j)] = 0, k \neq j$,

and there exist a positive constant α and an integer N such that the following persistent excitation condition holds:

(A2) $\sum_{j=0}^{N}\sum_{l=0}^{p-1}\frac{\hat{\boldsymbol{\varphi}}_i(k+j-l)\hat{\boldsymbol{\varphi}}_i^{\mathrm{T}}(k+j-l)}{r_i(k+j)} \geqslant \alpha \boldsymbol{I}$.

Then, the parameter estimation vector $\hat{\boldsymbol{\theta}}_i(k)$ given by the AM-MISG algorithm consistently converges to the true parameter vector $\boldsymbol{\theta}_i$ in the mean-square sense.

Theorem 1 can be proven in a similar way to [42]. Therefore, its detailed proof is omitted here.

4. Simulation Example

Assume that the nonlinear block in Figure 1 is described by:

$$\overline{u}(t) = u(t) + 0.5u^2(t) + 0.25u^3(t),$$

and the linear continuous process P is described by:

$$P(s) = \frac{1}{4s^2 + 2s + 1}.$$

Over a frame period of $T = 2\sqrt{2} + 1$ s, the input-output data are non-uniformly-sampled twice (i.e., $q = 2$) at $t_0 = 0$ s and $t_1 = \sqrt{2}$ s. According to Theorem 1 in [35], the following δ^{-1}-based transfer function model is derived:

$$w_0(k) = \frac{1 - 0.61623\delta^{-1} + 0.50931\delta^{-2}}{1 - 0.8086\delta^{-1} + 0.25514\delta^{-2}}\overline{u}_0(k),$$
$$w_1(k) = \frac{1 - 0.49522\delta^{-1} + 0.75553\delta^{-2}}{1 - 0.94779\delta^{-1} + 0.57794\delta^{-2}}\overline{u}_1(k).$$

Therefore, the parameters of this periodically non-uniformly sampled-data Hammerstein system are:

$$\boldsymbol{a}_0 = [-0.8086,\ 0.25514]^{\mathrm{T}}, \quad \boldsymbol{b}_0 = [1,\ -0.61623,\ 0.50931]^{\mathrm{T}},$$
$$\boldsymbol{a}_1 = [-0.94779,\ 0.57794]^{\mathrm{T}}, \quad \boldsymbol{b}_1 = [1,\ -0.49522,\ 0.75553]^{\mathrm{T}},$$
$$\boldsymbol{c} = [1,\ 0.5,\ 0.25]^{\mathrm{T}}.$$

In the simulation, take the non-uniform inputs $\{u_0(k)\}$ and $\{u_1(k)\}$ as two uncorrelated random sequences with zero mean and unit variance, $\{v_0(k)\}$ and $\{v_1(k)\}$ as two white noise sequences with zero mean and variance $\sigma^2 = 0.10^2$. Based on 5000 non-uniform input-output dataset, applying the AM-MISG algorithm in (14)–(25) with $\lambda_0 = \lambda_1 = 0.95$ to estimate the system parameters, the results for $p = 1$, $p = 5$ and $p = 12$ are shown in Tables 1–3, respectively, where ε is the estimation error defined as:

$$\varepsilon = \sqrt{\frac{\|\hat{\boldsymbol{a}}_0 - \boldsymbol{a}_0\|^2 + \|\hat{\boldsymbol{b}}_0 - \boldsymbol{b}_0\|^2 + \|\hat{\boldsymbol{a}}_1 - \boldsymbol{a}_1\|^2 + \|\hat{\boldsymbol{b}}_1 - \boldsymbol{b}_1\|^2 + \|\hat{\boldsymbol{c}} - \boldsymbol{c}\|^2}{\|\boldsymbol{a}_0\|^2 + \|\boldsymbol{b}_0\|^2 + \|\boldsymbol{a}_1\|^2 + \|\boldsymbol{b}_1\|^2 + \|\boldsymbol{c}\|^2}} \times 100\%.$$

Meanwhile, the estimation errors ε versus k are shown in Figure 3.

Table 1. The auxiliary model-based multi-innovation stochastic gradient (AM-MISG) parameter estimates and errors ($p = 1$).

k	100	200	500	1000	2000	3000	4000	5000	True Values
a_{01}	−0.23798	−0.23467	−0.24155	−0.23676	−0.32003	−0.37711	−0.42344	−0.46120	−0.80860
a_{02}	−0.17253	−0.15925	−0.14095	−0.10136	−0.07608	−0.04580	−0.00265	0.02211	0.25514
b_{00}	0.21247	0.28325	0.36820	0.43435	0.54777	0.63347	0.70653	0.76997	1.00000
b_{01}	0.03019	0.01933	0.01196	0.00916	0.00769	−0.01410	−0.04420	−0.08780	−0.61623
b_{02}	0.08793	0.10742	0.14003	0.18442	0.23272	0.28553	0.31263	0.32151	0.50931
a_{11}	−0.14144	−0.17751	−0.20109	−0.25406	−0.32583	−0.39976	−0.44355	−0.49299	−0.94779
a_{12}	−0.26469	−0.24184	−0.17359	−0.12188	−0.07576	−0.02374	0.04034	0.07391	0.57794
b_{10}	0.24088	0.32232	0.39463	0.45898	0.57929	0.66877	0.73108	0.77370	1.00000
b_{11}	0.02779	0.05257	0.07272	0.09047	0.11098	0.11615	0.10811	0.09453	−0.49522
b_{12}	0.04644	0.04343	0.04713	0.07480	0.12538	0.16100	0.18916	0.22188	0.75553
c_2	1.05391	0.96668	0.98187	0.93895	0.80087	0.72297	0.66712	0.63075	0.50000
c_3	1.79633	1.68065	1.44046	1.16009	0.82335	0.65456	0.52589	0.45922	0.25000
ε (%)	103.04890	97.41924	89.40996	80.57154	68.94369	61.44114	55.52206	51.00555	

Table 2. The AM-MISG parameter estimates and errors ($p = 5$).

k	100	200	500	1000	2000	3000	4000	5000	True Values
a_{01}	−0.17179	−0.18574	−0.31748	−0.41405	−0.61777	−0.71905	−0.76959	−0.80135	−0.80860
a_{02}	−0.17111	−0.11012	−0.05496	0.01077	0.12643	0.17775	0.21975	0.24397	0.25514
b_{00}	0.37344	0.45616	0.62647	0.76016	0.91687	0.95693	0.98483	0.99692	1.00000
b_{01}	0.06673	0.04317	0.02695	−0.05431	−0.24467	−0.39466	−0.50874	−0.56748	−0.61623
b_{02}	0.18052	0.22395	0.27192	0.33856	0.35650	0.39392	0.41745	0.43668	0.50931
a_{11}	−0.34108	−0.41100	−0.47313	−0.56512	−0.68070	−0.77319	−0.82088	−0.85964	−0.94779
a_{12}	−0.01887	−0.03044	0.07773	0.16053	0.30682	0.37436	0.44459	0.47782	0.57794
b_{10}	0.38009	0.49836	0.59120	0.75693	0.92872	0.97376	0.98098	0.98525	1.00000
b_{11}	0.02215	0.04295	0.06078	0.02968	−0.07356	−0.20574	−0.29673	−0.36568	−0.49522
b_{12}	0.10985	0.13310	0.18300	0.25133	0.33733	0.43426	0.52067	0.59753	0.75553
c_2	0.97825	0.90595	0.80559	0.63455	0.51789	0.50729	0.50178	0.50388	0.50000
c_3	1.35667	1.13733	0.73326	0.47388	0.31951	0.27787	0.26181	0.25438	0.25000
ε (%)	84.52995	75.91772	61.64633	48.73054	32.94076	22.45332	15.03823	10.03271	

Table 3. The AM-MISG parameter estimates and errors ($p = 12$).

k	100	200	500	1000	2000	3000	4000	5000	True Values
a_{01}	−0.21097	−0.29207	−0.48337	−0.62718	−0.78913	−0.82242	−0.82167	−0.81394	−0.80860
a_{02}	−0.09616	−0.00999	0.04310	0.14720	0.23542	0.25631	0.25755	0.25552	0.25514
b_{00}	0.46831	0.60206	0.82476	0.92527	0.99463	0.99562	0.99966	1.00088	1.00000
b_{01}	0.05982	0.04225	−0.06860	−0.27395	−0.53707	−0.61472	−0.62645	−0.62093	−0.61623
b_{02}	0.27987	0.32848	0.34589	0.38704	0.42283	0.47864	0.49652	0.50327	0.50931
a_{11}	−0.42997	−0.51235	−0.60352	−0.72537	−0.84006	−0.90571	−0.93254	−0.94364	−0.94779
a_{12}	−0.00011	0.04270	0.21788	0.32975	0.47520	0.53405	0.56383	0.57217	0.57794
b_{10}	0.44256	0.61696	0.76526	0.92836	1.00028	0.99858	0.99669	1.00004	1.00000
b_{11}	0.01971	0.05871	0.01162	−0.14602	−0.33453	−0.42806	−0.46968	−0.48801	−0.49522
b_{12}	0.11625	0.16567	0.24897	0.36557	0.55820	0.68047	0.73284	0.75125	0.75553
c_2	0.88043	0.76552	0.64129	0.52932	0.49418	0.49855	0.50087	0.49915	0.50000
c_3	1.04485	0.79819	0.45534	0.30318	0.25542	0.25328	0.25163	0.25057	0.25000
ε (%)	73.56646	62.33004	45.43354	29.28325	12.29280	4.72652	1.75023	0.55900	

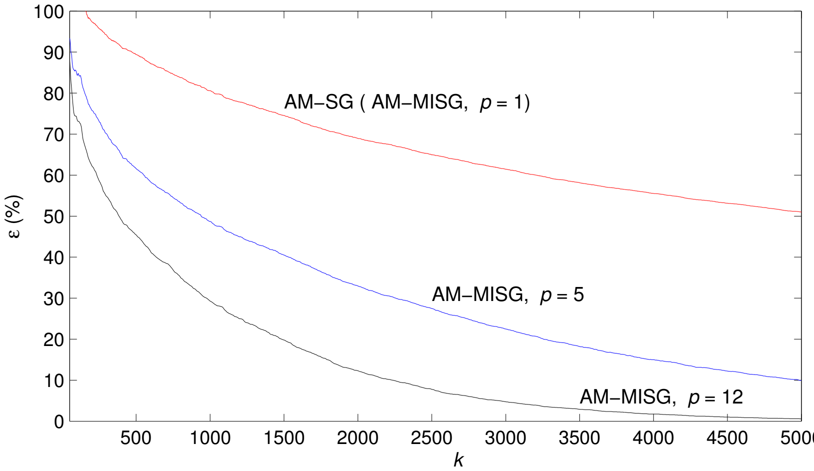

Figure 3. The AM-MISG estimation errors ε versus k with $p = 1, 5, 12$.

From Tables 1–3 and Figure 3, we can see that the parameter estimation error gradually decreases as the data length k increases, demonstrating the effectiveness of the proposed AM-MISG algorithm. Furthermore, the AM-MISG algorithm with a larger innovation length p can result in higher identification accuracy and faster convergence to the true parameters.

To study the identification performance of the proposed AM-MISG algorithm against the output noise, 50 Monte Carlo simulations for the noise variance being $\sigma^2 = 0.10^2$ and $\sigma^2 = 0.50^2$ have been conducted, respectively. In each simulation run, a new dataset with length $L = 5000$ is generated to estimate the model parameters. For the innovation length $p = 1$, $p = 5$ and $p = 12$, the mean values and the standard deviations of the parameter estimates are listed in Tables 4 and 5. From the simulation results, it can be observed that the estimation accuracy of the AM-MISG algorithm is higher when the noise variance is smaller. For a larger noise variance, increasing the innovation length p can help to obtain a satisfactory identification result.

Considering the noise variance $\sigma^2 = 0.10^2$ and $\sigma^2 = 0.50^2$, a separate dataset with length 30 has been generated for model validation, respectively. Using the mean values of the parameter estimates listed in Tables 4 and 5 for $p = 12$ to predict the outputs of the periodically non-uniformly sampled-data Hammerstein system, the results are shown in Figure 4, where the solid line and the x-marks represent the true measured outputs and the model predicted outputs, respectively. From Figure 4, it is clear that the model predictions can catch the trend of the true outputs well, and the prediction performance becomes better if a noise with smaller variance is introduced.

Table 4. The AM-MISG parameter estimates with standard deviations ($\sigma^2 = 0.10^2$).

Parameters	$p = 1$	$p = 5$	$p = 12$	True Values
a_{01}	-0.4618 ± 0.0241	-0.7999 ± 0.0067	-0.8156 ± 0.0016	-0.80860
a_{02}	0.0190 ± 0.0145	0.2390 ± 0.0050	0.2575 ± 0.0012	0.25514
b_{00}	0.7923 ± 0.0148	0.9979 ± 0.0038	1.0001 ± 0.0010	1.00000
b_{01}	-0.0783 ± 0.0165	-0.5640 ± 0.0081	-0.6243 ± 0.0020	-0.61623
b_{02}	0.3478 ± 0.0153	0.4419 ± 0.0054	0.5030 ± 0.0019	0.50931
a_{11}	-0.5010 ± 0.0187	-0.8457 ± 0.0105	-0.9396 ± 0.0024	-0.94779
a_{12}	0.0944 ± 0.0165	0.4700 ± 0.0099	0.5707 ± 0.0021	0.57794
b_{10}	0.7879 ± 0.0164	0.9951 ± 0.0049	0.9999 ± 0.0014	1.00000
b_{11}	0.1057 ± 0.0137	-0.3463 ± 0.0136	-0.4830 ± 0.0029	-0.49522
b_{12}	0.2357 ± 0.0127	0.5779 ± 0.0138	0.7465 ± 0.0033	0.75553
c_2	0.6136 ± 0.0106	0.4982 ± 0.0035	0.4998 ± 0.0010	0.50000
c_3	0.4436 ± 0.0153	0.2516 ± 0.0033	0.2499 ± 0.0009	0.25000

Table 5. The AM-MISG parameter estimates with standard deviations ($\sigma^2 = 0.50^2$).

Parameters	$p = 1$	$p = 5$	$p = 12$	True Values
a_{01}	-0.4613 ± 0.0261	-0.7870 ± 0.0217	-0.7921 ± 0.0249	-0.80860
a_{02}	0.0186 ± 0.0168	0.2275 ± 0.0173	0.2376 ± 0.0268	0.25514
b_{00}	0.7931 ± 0.0161	1.0017 ± 0.0170	1.0071 ± 0.0231	1.00000
b_{01}	-0.0777 ± 0.0191	-0.5491 ± 0.0217	-0.5954 ± 0.0350	-0.61623
b_{02}	0.3489 ± 0.0178	0.4452 ± 0.0203	0.5015 ± 0.0307	0.50931
a_{11}	-0.5009 ± 0.0225	-0.8320 ± 0.0203	-0.9036 ± 0.0269	-0.94779
a_{12}	0.0951 ± 0.0195	0.4552 ± 0.0203	0.5323 ± 0.0260	0.57794
b_{10}	0.7872 ± 0.0197	0.9938 ± 0.0203	1.0007 ± 0.0295	1.00000
b_{11}	0.1037 ± 0.0168	-0.3375 ± 0.0244	-0.4518 ± 0.0327	-0.49522
b_{12}	0.2375 ± 0.0171	0.5744 ± 0.0256	0.7177 ± 0.0312	0.75553
c_2	0.6130 ± 0.0121	0.4961 ± 0.0141	0.4949 ± 0.0207	0.50000
c_3	0.4446 ± 0.0185	0.2510 ± 0.0138	0.2462 ± 0.0166	0.25000

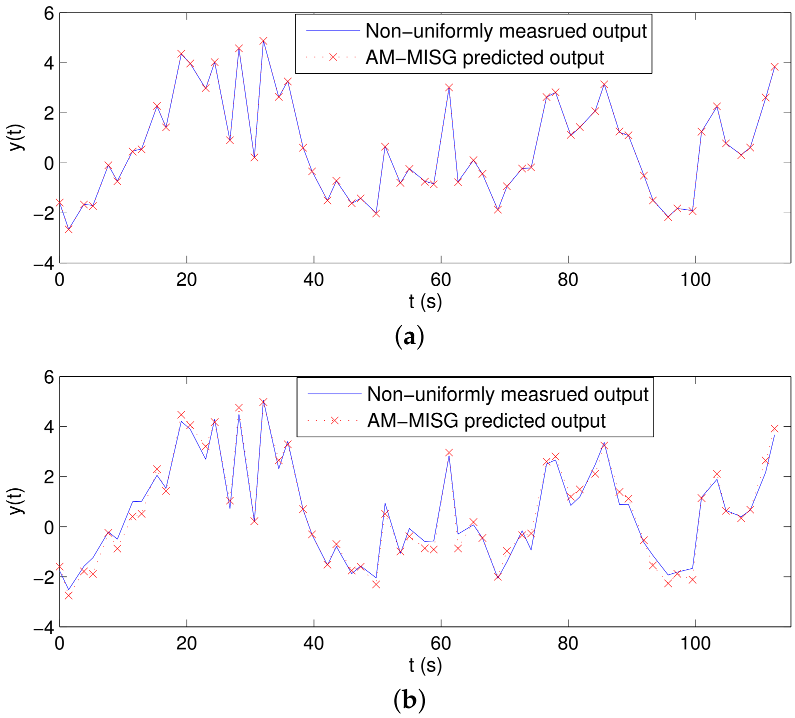

Figure 4. The predicted outputs and the measured outputs. (**a**) For the noise variance $\sigma^2 = 0.10^2$. (**b**) For the noise variance $\sigma^2 = 0.50^2$.

5. Conclusions

Based on the non-uniform input-output dataset, an auxiliary model-based stochastic gradient (AM-SG) algorithm is developed in this paper to estimate the parameters of Hammerstein systems. To improve the identification performance of the AM-SG algorithm, an auxiliary model-based multi-innovation stochastic gradient (AM-MISG) algorithm is proposed by introducing an innovation length p. The simulation results illustrate that the AM-MISG algorithm with a larger p can provide more accurate parameter estimates and a faster convergence rate. In addition, the proposed algorithm can be extended to identify more complex nonlinear systems with non-uniform sampling.

Acknowledgments: This paper is supported by the National Natural Science Foundation of China (No. 61403166), the Natural Science Foundation of Jiangsu Province (China; BK20140164) and Fundamental Research Funds for the Central Universities (JUSRP11561, JUSRP51510).

Author Contributions: Li Xie derived the identification algorithm and was in charge of the overall research. Huizhong Yang helped with preparing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, Y.; Wang, H.; Yu, J.; Li, P. Selective recursive kernel learning for online identification of nonlinear systems with NARX form. *J. Process Control* **2010**, *20*, 181–194.
2. Tang, Y.; Li, Z.; Guan, X. Identification of nonlinear system using extreme learning machine based Hammerstein model. *Commun. Nonlinear Sci. Numer. Simul.* **2014**, *19*, 3171–3183.
3. Giri, F.; Bai, E.W. *Block-Oriented Nonlinear System Identification*; Springer: London, UK, 2010.
4. Wills, A.; Schön, T.B.; Ljung, L.; Ninness, B. Identification of Hammerstein-Wiener models. *Automatica* **2013**, *49*, 70–81.
5. Lawrynczuk, M. Nonlinear predictive control for Hammerstein-Wiener systems. *ISA Trans.* **2014**, *55*, 49–62.
6. Wang, Y.Y.; Wang, X.D.; Wang, D.Q. Identification of dual-rate sampled Hammerstein systems with a piecewise-linear nonlinearity using the key variable separation technique. *Algorithms* **2015**, *8*, 366–379.
7. Chen, J.; Wang, X.; Ding, R. Gradient based estimation algorithm for Hammerstein systems with saturation and dead-zone nonlinearities. *Appl. Math. Model.* **2012**, *36*, 238–243.
8. Lv, X.; Ren, X. Non-iterative identification and model following control of Hammerstein systems with asymmetric dead-zone non-linearities. *IET Control Theory Appl.* **2012**, *6*, 84–89.
9. Giri, F.; Rochdi, Y.; Brouri, A.; Chaoui, F.Z. Parameter identification of Hammerstein systems containing backlash operators with arbitrary-shape parametric borders. *Automatica* **2011**, *47*, 1827–1833.
10. Fang, L.; Wang, J.; Zhang, Q. Identification of extended Hammerstein systems with hysteresis-type input nonlinearities described by Preisach model. *Nonlinear Dyn.* **2015**, *79*, 1257–1273.
11. Zhang, Q.; Wang, Q.; Li, G. Nonlinear modeling and predictive functional control of Hammerstein system with application to the turntable servo system. *Mech. Syst. Signal Process.* **2016**, *72*, 383–394.
12. Pouliquen, M.; Pigeon, E.; Gehan, O. Identification scheme for Hammerstein output error models with bounded noise. *IEEE Trans. Autom. Control* **2016**, *61*, 550–555.
13. Ding, F.; Shi, Y.; Chen, T. Auxiliary model-based least-squares identification methods for Hammerstein output-error systems. *Syst. Control Lett.* **2007**, *56*, 373–380.
14. Filipovic, V.Z. Consistency of the robust recursive Hammerstein model identification algorithm. *J. Frankl. Inst.* **2015**, *352*, 1932–1945.
15. Gao, X.; Ren, X.; Zhu, C.; Zhang, C. Identification and control for Hammerstein systems with hysteresis non-linearity. *IET Control Theory Appl.* **2015**, *9*, 1935–1947.
16. Cao, P.; Luo, X. Performance analysis of multi-innovation stochastic Newton recursive algorithms. *Digit. Signal Process.* **2016**, *56*, 15–23.
17. Ding, F. Complexity, convergence and computational efficiency for system identification algorithms. *Control Decis.* **2016**, *31*, 1729–1741.
18. Ding, F. Several multi-innovation identification methods. *Digit. Signal Process.* **2010**, *20*, 1027–1039.
19. Shi, Z.; Wang, Y.; Ji, Z. A multi-innovation recursive least squares algorithm with a forgetting factor for Hammerstein CAR systems with backlash. *Circuits Syst. Signal Process.* **2016**, *35*, 4271–4289.
20. Xiao, Y.; Song, G.; Liao, Y.; Ding, R. Multi-innovation stochastic gradient parameter estimation for input nonlinear controlled autoregressive models. *Int. J. Control Autom. Syst.* **2012**, *10*, 639–643.
21. Chen, J.; Zhang, Y.; Ding, R. Gradient-based parameter estimation for input nonlinear systems with ARMA noises based on the auxiliary model. *Nonlinear Dyn.* **2013**, *72*, 865–871.
22. Ma, J.; Xiong, W.; Ding, F.; Alsaedi, A.; Hayat, T. Data filtering based forgetting factor stochastic gradient algorithm for Hammerstein systems with saturation and preload nonlinearities. *J. Frankl. Inst.* **2016**, *353*, 4280–4299.
23. Li, W.; Shah, S.L.; Xiao, D. Kalman filters in non-uniformly sampled multirate systems: For FDI and beyond. *Automatica* **2008**, *44*, 199–208.
24. Albertos, P.; Salt, J. Non-uniform sampled-data control of MIMO systems. *Ann. Rev. Control* **2011**, *35*, 65–76.
25. Ding, F.; Wang, F.F. Recursive least squares identification algorithms for linear-in-parameter systems with missing data. *Control Decis.* **2016**, *31*, 2261–2266.
26. Yang, H.; Xia, Y.; Shi, P. Stabilization of networked control systems with nonuniform random sampling periods. *Int. J. Robust Nonlinear Control* **2011**, *21*, 501–526.

27. Aibing, Q.; Bin, J.; Chenglin, W.; Zehui, M. Fault estimation and accommodation for networked control systems with nonuniform sampling periods. *Int. J. Adapt. Control Signal Process.* **2015**, *29*, 427–442.
28. Sheng, J.; Chen, T.; Shah, S.L. Generalized predictive control for non-uniformly sampled systems. *J. Process Control* **2002**, *12*, 875–885.
29. Khan, S.; Goodall, R.; Dixon, R. Non-uniform sampling strategies for digital control. *Int. J. Syst. Sci.* **2013**, *44*, 2234–2254.
30. Jing, S.; Pan, T.; Li, Z. Recursive bayesian algorithm with covariance resetting for identification of Box-Jenkins systems with non-uniformly sampled input data. *Circuits Syst. Signal Process.* **2016**, *35*, 919–932.
31. Xie, L.; Yang, H.; Ding, F. Identification of non-uniformly sampled-data systems with asynchronous input and output data. *J. Frankl. Inst.* **2017**, *354*, 1974–1991.
32. Li, X.; Ding, R.; Zhou, L. Least-squares-based iterative identification algorithm for Hammerstein nonlinear systems with non-uniform sampling. *Int. J. Comput. Math.* **2013**, *90*, 1524–1534.
33. Salt, J.; Albertos, P. Model-based multirate controllers design. *IEEE Trans. Control Syst. Technol.* **2005**, *13*, 988–997.
34. Huang, J.; Shi, Y.; Huang, H.; Li, Z. l_2-l_∞ filtering for multirate nonlinear sampled-data systems using T-S fuzzy models. *Digit. Signal Process.* **2013**, *23*, 418–426.
35. Xie, L.; Yang, H.; Ding, F.; Huang, B. Novel model of non-uniformly sampled-data systems based on a time-varying backward shift operator. *J. Process Control* **2016**, *43*, 38–52.
36. Chang, F.; Luus, R. A noniterative method for identification using Hammerstein model. *IEEE Trans. Autom. Control* **1971**, *16*, 464–468.
37. Shen, Q.; Ding, F. Least squares identification for Hammerstein multi-input multi-output systems based on the key-term separation technique. *Circuits Syst. Signal Process.* **2016**, *35*, 1–14.
38. Zhang, Y.; Zhao, Z.; Cui, G. Auxiliary model method for transfer function estimation from noisy input and output data. *Appl. Math. Model.* **2014**, *39*, 4257–4265.
39. Jin, Q.; Wang, Z.; Liu, X. Auxiliary model-based interval-varying multi-innovation least squares identification for multivariable OE-like systems with scarce measurements. *J. Process Control* **2015**, *35*, 154–168.
40. Ding, J. Data filtering based recursive and iterative least squares algorithms for parameter estimation of multi-input output systems. *Algorithms* **2016**, *9*, 49, doi: 10.3390/a9030049.
41. Bai, E.W. An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems. *Automatica* **1998**, *34*, 333–338.
42. Wang, X.; Ding, F. Convergence of the auxiliary model-based multi-innovation generalized extended stochastic gradient algorithm for Box-Jenkins systems. *Nonlinear Dyn.* **2015**, *82*, 269–280.