Research Article

An Improved Method for Stochastic Nonlinear System’s Identification Using Fuzzy-Type Output-Error Autoregressive Hammerstein–Wiener Model Based on Gradient Algorithm, Multi-Innovation, and Data Filtering Techniques

Donia Ben Halima Abid 1, 2, Saif Eddine Abouda 1, Hanane Medhaffar 1, and Mohamed Chtourou 1

1University of Sfax, National Engineering School of Sfax (ENIS), Control & Energy Management Laboratory (CemLab), Sfax 3038, Tunisia
2University of Sfax, National Engineering School of Sfax (ENIS), Sfax 3038, Tunisia

Correspondence should be addressed to Donia Ben Halima Abid; donia.benhalima-abid@enis.tn

Received 17 May 2021; Revised 9 July 2021; Accepted 28 July 2021; Published 20 August 2021

Academic Editor: Carlos Aguilar-Ibáñez

Copyright © 2021 Donia Ben Halima Abid et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes an innovative identification approach of nonlinear stochastic systems using Hammerstein–Wiener (HW) model with output-error autoregressive (OEA) noise. Two fuzzy systems are suggested for the identification of the input and output nonlinear blocks of a proposed model from given input-output data measurements. In this work, the need for the commonly used assumptions including well-known structure of input and/or output nonlinearities and/or reversible nonlinear output is eliminated by replacing the intermediate variables and noise with their estimates. Four parametric estimation algorithms to identify the proposed fuzzy-type stochastic output-error autoregressive HW (FSOEAHW) model are derived based on backpropagation algorithm and multi-innovation and data filtering identification techniques. The proposed algorithms are improved backpropagation gradient (IBPG) algorithm, multi-innovation IBPG (MIIBPG) algorithm, a data filtering IBPG (FIBPG) algorithm, and a multi-innovation-based FIBPG (MIFIBPG) algorithm. The convergence of the parameter estimation algorithms is studied. The effectiveness of the proposed algorithms is shown by a given simulation example.

1. Introduction

The system identification is an approach of modeling a dynamic system based on data observations [1–9]. It consists of finding a suitable mathematical model which approximates, as well as possible, the system’s behavior by giving an appropriate structure and parameters model, establishing a relation between inputs and outputs system measurements [10–13].

All physics systems are nonlinear, and it is natural to use nonlinear model to describe such systems. Therefore, nonlinear identification and control techniques have received, recently, more attention, since linear identification and control-based approaches are widely developed and are becoming mature. One class of such nonlinear modeling system is so-called block-oriented models that can be represented in various configurations where linear dynamic blocks and nonlinear static or dynamic subsystems are cascaded. The Hammerstein (H) model (static nonlinear block followed by a dynamic linear one) and the Wiener (W) system (a linear dynamic subsystem followed by a static nonlinear block) are the basic class of the cascaded systems which are widely used in many industrial practice engineering applications [14–22] and, therefore, the modeling approaches of such class of block-oriented models have received great attention for many years [23–36]. Hammerstein and Wiener systems are combined together to produce more complex subcategories, namely, Hammerstein–Wiener (HW) model (a linear block is cascaded between two nonlinear subsystems)
and a Wiener–Hammerstein (WH) system (a nonlinear block is embedded between two linear blocks). Particularly, in many cases, the real system contains both actuator and sensor nonlinearities. Then, it is appropriate to consider the input nonlinear block as actuator nonlinearity and the output nonlinear subsystem as sensor or/and process nonlinearity. The HW model has several advantages: (i) It has a physical view of the nonlinear characteristics of the real system, which is important in the analysis, monitoring, diagnosis, and control of the system. (ii) The HW dynamics are mainly produced by a linear subsystem. Then, developed linear approaches could be used. (iii) When the output nonlinearity has an inverse function, the linear control techniques can be easily applied with desired performances. This block-oriented model form is perfectly used in different fields, such as electrical, mechanical, hydraulic, and chemical fields [37–46]. Consequently, the identification of HW models has been an active research topic nowadays. In literature, different methods are proposed and can be roughly divided into different categories: the recursive, the iterative, the blind, the subspace, the frequency domain, the overparameterization, decomposition methods, and so on.

The basic idea of the abovementioned methods is that the model parameters are approximated either by constructing hybrid model of nonlinear and linear parts (e.g., overparameterization, subspace, decomposition, and blind methods [36, 43–46]) or by separated steps where the estimation input and output nonlinear subsystems parameters and the dynamic linear part ones are established based on the unmeasured intermediate variables estimation step (such as frequency domain, iterative, recursive, stochastic, correlation methods, and a special input-based one [35, 47–51]).

However, the common representation of the abovementioned works is that the process noise (i.e., noise given between the linear dynamic part and the nonlinear block) or output disturbances (i.e., measurement noise given after nonlinear output block) have not been considered, which is not the case in practical physical process. Then, it is more evident to consider some stochastic disturbances. Along these lines, a very few papers that deal with stochastic H-W model have been proposed in the literature. For example, in [43–49], the authors treated a particular structure of H-W presenting some forms of measurement noise. For H-W with process noise, some papers were proposed in [49–53]. Esmaeilani et al. and Wills et al. [54–56] have proposed identification methods for an H-W system’s class presenting measurement and process noises.

It should be noted that the aforementioned approaches’ applications are severely limited due to some problems. In fact, correspondent-cited ideas are restrictive to polynomial forms of the input and/or output nonlinear blocks or well-known input and/or output nonlinear characteristics (like dead zone or backlash) but with unknown parameters. Therefore, if the nonlinearity is not continuous or not in the polynomial form or with unknown characteristics, the algorithms do not give satisfactory performances. Moreover, the redundancy problem producing an oversizing dimension of HW’s matrix parameters is another loophole of previously mentioned methods. For compensation of these shortcomings, artificial intelligent systems such as neural networks and fuzzy systems can be explored to model the H-W model owing to their universal approximation property and their ability to model a given nonlinear function to any arbitrary accuracy. Considering that the H-W structure should be ensured, the identification problem of neural network or fuzzy-type H-W is different (where the modeling of the two nonlinear blocks and the dynamic linear part are altogether needed) from that of traditional neural network and fuzzy system which focuses only on the global data process nonlinear transformation. Nowadays, only scattered works were reported in the identification of H-W model based on neural networks and fuzzy systems. In [57, 58], a multistage approach is proposed to establish a recurrent neural network type H-W model based on an active region boundary initialization, a frequency domain eigensystem realization, and least squares and recursive recurrent learning algorithms. A neuro-fuzzy-type H-W system was presented in [59] using two-stage input signal. For the same form of system, Jia et al. and Li et al. [60, 61] combined special input signal and correlation technique to identify separately and nonrecursively system parameters. However, most of the above-cited algorithms are based on some restrictive conditions, especially the prior knowledge of some parameters and the output nonlinearity's invertibility assumption to obtain intermediate variable's estimate, which is not always the case. Even in [62], a backpropagation gradient algorithm is used to estimate jointly unknown parameters and intermediate variables of neural network type H-W system with polynomial form of input and output nonlinear blocks.

It is clear that there are some problems that need to be addressed in nonlinear process identification using H-W model to achieve satisfactory results. The first one is how to consider stochastic nonlinear process disturbances in the H-W model description. The second is how to identify H-W model without any prior system knowledge and using, only, input and output measurement. The third is how to acquire the powerful identification aptitude of the H-W model with minimum attractive theoretical assumptions. The last is to identify the learning algorithm that can be used to achieve good performance.

To deal with the above-described issues, this paper presents a novel modeling and parameter identification method for the identification of nonlinear stochastic process described by fuzzy-type output-error autoregressive H-W (FSOEAHW) model. Thus, compared with the existent studies, the main contributions of this paper are as follows.

The first originality lies in the proposed fuzzy-type H-W scheme. In fact, the linear part is considered as a discrete transfer function and an output-error autoregressive mathematical model describes the process disturbance. Moreover, two fuzzy models are designed to describe the input and output nonlinearities of H-W model with only input and output measurement knowledge. As a result, the proposed model can not only describe the dynamics of such nonlinear system operating in stochastic environment and avoid the encountered input and output nonlinear blocks.
restrictions based on particular or basis functions forms but
also eliminate the reversibility assumption of the static
output part based on internal variable estimates without
parameter model oversizing problem.

The second involvement is related to the proposed pa-
rameter estimation algorithms. The major idea is to combine
a feedback propagation gradient algorithm with a multi-
innovation and data filtering technique for the fuzzy-type
H-W model identification. The proposed identification al-
gorithms are based on input/output measurement and the
approximation of all internal variables resulting from the
preceding corresponding parameter estimates. In this in-
stance, four algorithms are suggested: an improved back-
propagation gradient (IBPG) algorithm, a multi-innovation
improved backpropagation gradient (MIIBPG) algorithm
for improving the convergence rate through the multi-in-
novation identification technique, a data filtering IBPG
(FIBPG) algorithm, and a multi-innovation FIBPG
(MIFIBPG).

The remainder of this paper is organized as follows. The
considered system is described in Section 2. In Section 3, the
FSoEAHW parameter estimation algorithms, IBPG,
MIIBPG, FIBPG, and MIFIBPG, are developed and their
convergence analysis is studied. Section 4 presents simu-
lation results. Some conclusions are offered in Section 5.

2. System Description

2.1. Stochastic Output-Error Autoregressive Hammer-
stein–Wiener (SOEAHW) System. Consider a stochastic
output-error autoregressive Hammerstein–Wiener
(SOEAHW) system given by Figure 1.

\[
\begin{align*}
    w_k &= \frac{1}{C(z)} v_k = \frac{1}{1 + c_1 z^{-1} + \cdots + c_n z^{-n}} v_k = -\sum_{p=1}^{n} c_p w_{k-p} + v_k = c^T \psi_k + v_k.
\end{align*}
\]

We have \( c = [c_1, c_2, \ldots, c_n]^T \) and \( \psi_k = [-w_{k-1}, -w_{k-2}, \ldots, -w_{k-n}]^T \). Given equations (4) and (5), the system’s output is expressed as

\[
y_k = f_2(x_k) + w_k = y'_{k} + w_k = y'_{k} + c^T \psi_k + v_k.
\]

The objective of this paper is to give the following:

(i) A good description of an unknown nonlinear system
operating in a stochastic environment using HW
model despite the presence of disturbance and the
great model complexity, with minimum attractive
theoretical assumptions that are not always valid,
especially the rigorous restrictions related to the
nonlinear output block’s reversibility and well-
known input and/or output nonlinear blocks
characteristics

(ii) A suitable approximation of the above-described
system such as the sum of square residual term errors

In that system, \( u_k \) and \( y_k \) are system input and output, respectively. \( y'_{k} \), \( x_k \), and \( h_k \) are internal variables. The discrete linear transfer function \( G(.) \) is surrounded by an
input static nonlinear block \( f_1(.) \) and an output static
nonlinear block \( f_2(.) \). It is assumed that the measured
output \( y_k \) contains an unknown additive noise component
\( w_k \) described by an autoregressive mathematical model. \( v_k \) is
a white noise with zero mean and unknown variance \( \sigma^2 \).

The first unmeasurable intermediate variable \( h_k \) is the
output of input nonlinear block and it is expressed by the
following equation:

\[
h_k = f_1(u_k).
\]

The linear dynamic part is given by the following
expression:

\[
G(z) = \frac{B(z)}{A(z)} = \frac{b_1 z^{-1} + b_2 z^{-2} + \cdots + b_n z^{-n}}{1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}}
\]

The second unmeasurable intermediate variable can be
written as follows:

\[
x_k = G(z)h_k = -\sum_{i=1}^{na} a_i x_{k-i} + \sum_{j=1}^{nb} b_j h_{k-j}.
\]

The output of the nonlinear output block is expressed as
follows:

\[
y_{1*} = f_2(x_k).
\]

Noise \( w_k \) is given by

\[
E \left( f_1, \theta, f_2 \right) = \sum_{k=1}^{M} \left( y_k - \hat{y}_k \right)^2,
\]

where \( \hat{y}_k \) is a priori estimated output.

To achieve the abovementioned objectives and based on
the universal approximation properties of fuzzy systems, we
propose using two fuzzy models to describe the input and
output nonlinearities.

2.2. Fuzzy-Type Stochastic Output-Error Autoregressive
Hammerstein–Wiener (FSoEAHW) Model. The crucial
challenge for the identification of Hammerstein–Wiener
model is to result in effective modeling of two static
nonlinear functions. Better modeling requires not only approximating the nonlinear function accurately but also simplifying the identification process. In the literature, several researches are restrictive to a polynomial form of the input and/or output nonlinear blocks or well-known input and/or output nonlinear characteristics (such as dead zone or backlash) but with unknown parameters [41–60, 63–74]. Therefore, if the nonlinearity is not continuous or not in the polynomial form or with unknown characteristic, the algorithms do not give a satisfactory performance. For compensation of the aforementioned shortcomings encountered in the existing structure of Hammerstein–Wiener compensation of the aforementioned algorithms do not give a satisfactory performance. For kV herefore, if the nonlinearity is not continuous or not in the input and output nonlinearities of H-W model with only input and output measurement knowledge.

In this section, we propose developing a modeling approach of a nonlinear dynamic process operating in the stochastic environment identification based on fuzzy model. According to the previous section, we propose a new fuzzy-type stochastic output-error autoregressive H-W (FSOEAHW) model given in Figure 2. It consists of two static nonlinear blocks described by two independent fuzzy systems: a dynamic linear block and autoregressive noise block. Then the first fuzzy system’s output \( h_k \) can be formulated as follows:

\[
h_k = \sum_{n=1}^{N_1} W_n g_n(u_k) = W_1^T \phi_1(u_k). \tag{8}
\]

Giving that \( W_n \) (\( n = 1, 2, \ldots, N_1 \)) is the \( n \)th fuzzy rule’s consequence and \( N_1 \) is the total rule number, \( g_n(u_k) = (\mu_n(u_k) / \sum_{n=1}^{N_1} \mu_n(u_k)) \) with \( \mu_n(u_k) = \exp(-1/2) \left[ \left( (u_k - \mu_n)/\sigma_n \right)^2 \right] \) is \( n \)th Gaussian membership function where \( \mu_n \) and \( \sigma_n \) present, respectively, the correspondent center and width. \( W_1 = [W_1, W_2, \ldots, W_{N_1}]^T \) is a vector parameter, and \( \phi_1(u_k) = [g_{11}(u_k), g_{12}(u_k), \ldots, g_{1N_1}(u_k)]^T \) is an input information vector.

It should be noted that the membership functions can be of several shapes such as triangular, trapezoidal, and Gaussian. The only condition that must be fulfilled is that it must be in the interval [0, 1]. In the literature, Gaussian shape is commonly used because of its simplicity, its smoothness, and nonzero at all points. It is defined by only two parameters (the center and the width) and it is a continuously differentiable function [75–77].

The second unmeasurable intermediate variable of the FSOEAHW can be written as follows:

\[
x_k = -\sum_{i=1}^{ma} a_i x_{k-i} + \sum_{j=1}^{mb} b_j h_{k-j} = \theta_j^T \varphi_j, \tag{9}
\]

where \( \theta_j = [a_1, a_2, \ldots, b_1, \ldots, b_m] \) and \( \varphi_j = [-x_{k-1}, \ldots, -x_{k-n_j}, h_{k-1}, \ldots, h_{k-n_j}] \) are, respectively, the vector parameters and the observation vector of the linear dynamic part.

The output of the second fuzzy system \( y_k \) is expressed by

\[
y_k = \sum_{m=1}^{N_2} W_m g_m(x_k) = W_2^T \varphi_2(x_k), \tag{10}
\]

where \( g_m(x_k) = (\mu_m(x_k) / \sum_{m=1}^{N_2} \mu_m(x_k)) \) with \( \mu_m(x_k) = \exp(-1/2) \left[ \left( (x_k - \mu_m)/\sigma_m \right)^2 \right] \) in which \( \mu_m \) and \( \sigma_m \) are the center and width of the \( m \)th membership function \( \mu_m(x_k) \). \( W_2 = [W_1, W_2, \ldots, W_{N_2}]^T \) and \( \varphi_2(x_k) = [g_{11}(x_k), g_{12}(x_k), \ldots, g_{N_2}(x_k)]^T \) are, respectively, the vector parameter and the observation vector of the output nonlinear block. \( W_m \) is the \( m \)th consequence fuzzy rule and \( N_2 \) is the correspondent total rule number.

Noise \( w_k \) is expressed by (5). Thus, the system output can be rewritten as

\[
y_k = \sum_{m=1}^{N_1} W_m g_m(x_k) + u_k = \sum_{m=1}^{N_3} W_m g_m(x_k) - \sum_{p=1}^{n} c_p w_{k-p} + v_k, \tag{11}
\]

or, equivalently, it can be written in the following matrix form:

\[
y_k = W_2^T \psi_2(x_k) + c^T \phi_k + v_k = W_2^T \phi_k + v_k. \tag{12}
\]

Hence, the parameter vector is \( W = [W_1, W_2, \ldots, W_{N_2}, c_1, c_2, \ldots, c_{N_3}]^T \) and the observation vector is equal to \( \phi_k = [\psi_2(x_k), \psi_k]^T = [g_{11}(x_k), g_{12}(x_k), \ldots, g_{N_2}(x_k), \varphi_k]^T = [-x_{k-1}, \ldots, -x_{k-n}, h_{k-1}, \ldots, h_{k-n}]^T \).

![Figure 1: SOEAHW model.](attachment:image.png)
Note that all vector parameters and observation vectors \( \phi_k(u_k), \phi_k(x_k), \) and \( \phi_k(y_k) \) are unknown. The problem then consists in developing a novel identification algorithm estimating the unknown parameters and variables using the input measurement and output measurement \( \{u_k\} \) and \( \{y_k\} \) that respect the objectives cited in the previous paragraph. Founded on the estimated model parameters, the mathematical model could be built identifying a given practical stochastic nonlinear dynamical system.

### 3. FSOEAHW Parameter Estimation

In this section, we suggest using a backpropagation gradient algorithm to estimate unknown parameters and unmeasured variable.

The gradient algorithm is an important tool in the linear and nonlinear problems in which a modification of parameter estimates is reached using the negative gradient direction of the criterion function. Today and in the subject of identification and control, different recursive and iterative gradient algorithms are proposed [16, 17, 78–84]. Stochastic gradient algorithm is a basic recursive identification algorithm which is used to study different types of systems such as multivariable systems [85, 86] and nonlinear block-oriented systems [16, 17, 26, 87]. In [88], an iterative gradient parameter estimation for output-error autoregressive systems using hierarchical principle has been presented. Yu et al. [31] used a gradient-based backpropagation algorithm to identify Hammerstein neural network type system. An equal algorithm has been extended by [62] for neural network type H-W system identification.

The weakness of the gradient-based methods lies in the fact that estimation accuracy is not good enough for precision control purpose and the convergence rate is very slow. To improve these drawbacks, different approaches have been proposed in the literature. Among them, the multi-innovation technique is one of the most popular techniques which can improve the parameter estimation quality [89]. It consists of parameter estimation based not only on the current data but also on previous finite data at each iteration. Different papers are presented, in the literature, for the identification of different classes of systems like bilinear-in-parameter systems [90], multivariable linear systems [91, 92], block-oriented systems [93–95], and so on.

Another meaningful way for improving parameter estimation is manipulating the powerful data filtering technique. The focal idea based on identification algorithm is to generate system parameter estimates using a special filter to filter the measurement data and then identify the filtered system model and the filtered noise model. In these regards, the structure of the system to identify will be changed without eliminating noise from data or changing the relationship between variables. This technique has shown the effectiveness in the identification of different types of disturbed systems such as those considered in [70, 92, 96–104].

In pursuit, we present four estimation algorithms to give the unknown parameters of the proposed FSOEAHW model. The first algorithm is the improved backpropagation gradient (IBPG) that utilized the backpropagation-based gradient algorithm. The second algorithm, namely, MIBPG, employs a multi-innovation technique and IBPG algorithm for improving the convergence rate through the multi-innovation identification technique. Lastly, and for the same purpose, a data filtering-based IBPG (FIBPG) algorithm and a multi-innovation-based FIBPG (MIFIBPG) algorithm are proposed.

#### 3.1. Improved Backpropagation Gradient (IBPG) Algorithm.

Let \( h_{k,t}, x_{k,t} \), and \( y_{k,t} \) be the estimated variables, respectively, of intermediate variables \( h_k, x_k, \) and \( y_k \) at iteration step \( t \) as follows:

\[
\hat{h}_{k,t} = \hat{W}_{t,1}^T \tilde{\phi}_{t,1-1}(u_k),
\]

\[
\hat{x}_{k,t} = -\sum_{i=1}^{n_a} \tilde{a}_{i,t} \tilde{x}_{i,k,t-1} + \sum_{j=1}^{n_b} \tilde{b}_{j,t} \hat{h}_{k,t},
\]

\[
= \hat{a}_{i,t}^T \tilde{\phi}_{t,1-1},
\]

\[
\hat{y}_{j,k,t} = \hat{W}_{2,t}^T \tilde{\phi}_{2,t-1}(\tilde{x}_{k,t}),
\]

where the estimated vector parameters at recursive step \( t \) are \( \hat{W}_{1,t} = [\hat{W}_{1,1}, \hat{W}_{1,2}, \ldots, \hat{W}_{1,N_1}, \ldots]^T, \) \( \hat{\theta}_{1,t} = [\tilde{a}_{1,t}, \ldots, \tilde{a}_{n_a,t}, \tilde{b}_{1,t}, \ldots, \tilde{b}_{n_b,t}], \) and \( \hat{W}_{2,t} = [\hat{W}_{2,1,t}, \hat{W}_{2,2,t}, \ldots, \hat{W}_{2,N_2,t}]^T. \) The predicted observation vectors are \( \tilde{\phi}_{1,t} = \tilde{\phi}_{1,t} \) and \( \tilde{\phi}_{2,t} = \tilde{\phi}_{2,t} \) as mentioned in the previous section.

In this regard, the estimated variables \( h_k, x_k, \) and \( y_k \) are the center and width estimates \( \mu_{n,t} \) and \( \sigma_{n,t} \) at iteration step \( t \) as follows:

\[
\mu_{n,t} = \text{exp} \left( -1/2 \left( (u_{k,t} - \tilde{\phi}_{n,t}) / \tilde{\sigma}_{n,t} \right)^2 \right),
\]

\[
\sigma_{n,t} = \text{exp} \left( -1/2 \left( (x_{k,t} - \tilde{\phi}_{m,t}) / \tilde{\sigma}_{m,t} \right)^2 \right),
\]

\[
\hat{a}_{i,t}^T \tilde{\phi}_{t,1-1},
\]

\[
\hat{b}_{i,t}^T \tilde{\phi}_{t,1-1},
\]

where \( n \) and \( m \) are parameters of membership functions \( \mu_{n,t} \) and \( \mu_{m,t} \), \( n = 1: N_1 \) and \( m = 1: N_2 \).
As a result, the a priori estimated output \( \hat{y}_k \) can be expressed by the following equivalent adjustable model:

\[
\hat{y}_{kj} = \hat{y}_{kj} - \sum_{p=1}^{n} \hat{c}_{p,j-1} \hat{w}_{k-p} = \hat{W}_{2j-1}^T \hat{\theta}_{2j-1} \left( \hat{x}_{kj} \right) + \hat{c}_{j-1} \hat{\theta}_{kj} = \hat{W}_{1j-1}^T \hat{\theta}_{kj},
\]

(16)

where \( \hat{c}_t = \left[ \hat{c}_{1,t}, \hat{c}_{2,t}, \ldots, \hat{c}_{n,t} \right]^T \), \( \hat{\theta}_{kj} = \left[ \hat{\theta}_{kj} \right]^T \), \( \hat{W}_t = \left[ \hat{W}_{1,t}, \hat{W}_{2,t}, \ldots, \hat{W}_{2n,t} \right] \), \( \hat{c}_{kj}, \hat{c}_{2,t}, \ldots, \hat{c}_{n,t} \), and \( \hat{\theta}_{kj} = \left[ \hat{\theta}_{kj} \right]^T \). \( \hat{W}_t \) is the estimate of output noise \( w \).

To give \( \hat{w}_k, y_k \) is replaced by its estimate \( \hat{y}_{kj} \), which leads to the following expression:

\[
\hat{w}_{kj} = y_k - \hat{W}_{2j}^T \hat{\theta}_{2j}^T \hat{\phi}_{kj}. \tag{17}
\]

Using equation (16), we can define an estimation error term as

\[
e_{kj} = y_k - \hat{y}_{kj}. \tag{18}
\]

Using backpropagation gradient algorithm, each unknown parameter in the FSUENHW model is updated according to adjustment formula given by the minimization of the following quadratic error function with respect to this parameter:

\[
J_{kj} = \frac{1}{2} e_{kj}^2 = \frac{1}{2} \left( y_k - \hat{y}_{kj} \right)^2. \tag{19}
\]

By recursive application of the chain rule, all unmeasurable variables should be calculated first based on the FSOEHW parameter values before adjustment (i.e., the parameters resulting from the previous adjustment step t-1).

Then each unknown parameter is adjusted according to the following expression:

\[
\alpha_t = \alpha_{t-1} + \Delta \alpha_t = \alpha_{t-1} - \eta_{\alpha} \left( \frac{\partial J_{kj}}{\partial \alpha_{t-1}} \right), \tag{20}
\]

where \( \eta_{\alpha} \) is the learning rate of parameter updates. Parameters \( \alpha_t \) and \( \alpha_{t-1} \) are, respectively, values of parameter \( \alpha \) after and before each adjustment step. \( \alpha \) represents one of a tuning FSUENHW’s set of parameters \( \{ \hat{W}_{1n}, \hat{\alpha}_{1n}, \hat{\alpha}_{2n}, \hat{b}_j, \hat{W}_{1n}, \hat{\alpha}_{1n}, \hat{\alpha}_{2n}, \hat{\psi} \} \) with \( n = 1, \ldots, N_1 \), \( i = 1, \ldots, n_1, j = 1, \ldots, n_1, m = 1, \ldots, N_2, p = 1, \ldots, n_1 \).

Applying IBPG algorithm, the adjustment equations of the first fuzzy model parameters are given as follows:

\[
\hat{W}_{1n,t} = \hat{W}_{1n,t-1} + \eta_{\hat{W}_{1n}} e_{kj} \frac{\partial \hat{y}_{kj}}{\partial \hat{\theta}_{kj}} \frac{\partial \hat{\theta}_{kj}}{\partial \hat{W}_{1n,t-1}}, \quad n = 1, 2, \ldots, N_1, \tag{21}
\]

\[
\hat{\alpha}_{1n,t} = \hat{\alpha}_{1n,t-1} + \eta_{\hat{\alpha}_{1n}} e_{kj} \frac{\partial \hat{y}_{kj}}{\partial \hat{\theta}_{kj}} \frac{\partial \hat{\theta}_{kj}}{\partial \hat{\alpha}_{1n,t-1}}, \tag{22}
\]

\[
\hat{\alpha}_{2n,t} = \hat{\alpha}_{2n,t-1} + \eta_{\hat{\alpha}_{2n}} e_{kj} \frac{\partial \hat{y}_{kj}}{\partial \hat{\theta}_{kj}} \frac{\partial \hat{\theta}_{kj}}{\partial \hat{\alpha}_{2n,t-1}}, \tag{23}
\]

where

\[
\frac{\partial \hat{y}_{kj}}{\partial \hat{x}_{kj}} = \sum_{s=1}^{N_2} \left( \frac{\hat{x}_{s,t} - \hat{\alpha}_{s,t-1}}{\hat{\alpha}_{s,t-1}} \right)^t \hat{\alpha}_{s,t-1} (\hat{y}_{kj} - \hat{W}_{s,t-1} \hat{x}_{kj}), \tag{24}
\]

\[
\frac{\partial \hat{\theta}_{kj}}{\partial \hat{W}_{1n,t-1}} = \sum_{j=1}^{n} \hat{b}_{j,t-1} \hat{\theta}_{n,t-1} (\hat{u}_{k-j}) - \hat{a}_{n,t-1} \frac{\partial \hat{\theta}_{kj}}{\partial \hat{W}_{1n,t-1}}, \tag{25}
\]

\[
\frac{\partial \hat{x}_{kj}}{\partial \hat{\alpha}_{1n,t-1}} = \sum_{j=1}^{n} \hat{b}_{j,t-1} \left( \frac{\hat{u}_{k-j} - \hat{\alpha}_{1n,t-1}}{\hat{\alpha}_{1n,t-1}} \right) \hat{\alpha}_{n,t-1} (\hat{w}_{1n,t-1} - \hat{h}_{k-j}) - \sum_{s=1}^{N_2} \hat{a}_{s,t-1} \frac{\partial \hat{x}_{kj}}{\partial \hat{\alpha}_{1n,t-1}}, \tag{26}
\]

\[
\frac{\partial \hat{\theta}_{kj}}{\partial \hat{\alpha}_{2n,t-1}} = \sum_{j=1}^{n} \hat{b}_{j,t-1} \left( \frac{\hat{u}_{k-j} - \hat{\alpha}_{2n,t-1}}{\hat{\alpha}_{2n,t-1}} \right) \hat{\alpha}_{n,t-1} (\hat{w}_{1n,t-1} - \hat{h}_{k-j}) - \sum_{s=1}^{N_2} \hat{a}_{s,t-1} \frac{\partial \hat{x}_{kj}}{\partial \hat{\alpha}_{2n,t-1}}, \tag{27}
\]
The linear dynamic part parameters are adjusted according to equations (28) and (29):

$$\tilde{a}_{i,t} = \tilde{a}_{i,t-1} + \eta_{a_i} e_{i,t} \frac{\partial \tilde{y}_{k,t}}{\partial a_{i,t-1}} \frac{\partial \tilde{x}_{k,t}}{\partial \tilde{a}_{i,t-1}}$$  \hspace{1cm} i = (1, 2, \ldots, n_a), \hspace{1cm} (28)

$$\tilde{b}_{j,t} = \tilde{b}_{j,t-1} + \eta_{b_j} e_{j,t} \frac{\partial \tilde{y}_{k,t}}{\partial b_{j,t-1}} \frac{\partial \tilde{x}_{k,t}}{\partial \tilde{b}_{j,t-1}}$$  \hspace{1cm} j = (1, 2, \ldots, n_b), \hspace{1cm} (29)

Finally, the parameter adjustment equations of the second fuzzy system and the additive noise can be obtained by the following formulas:

$$\tilde{W}_{j,m,t} = \tilde{W}_{j,m,t-1} + \eta_{\tilde{w}_j} e_{j,t} \tilde{g}_{j,m,t-1} \tilde{x}_{k,t}, \hspace{1cm} m = (1, 2, \ldots, N_a), \hspace{1cm} (30)$$

$$\tilde{a}_{j,m,t} = \tilde{a}_{j,m,t-1} + \eta_{a_j} e_{j,t} \left( \tilde{x}_{k,t} - \tilde{a}_{j,m,t-1} \tilde{x}_{k,t} \right), \hspace{1cm} (31)$$

$$\tilde{\sigma}_{j,m,t} = \tilde{\sigma}_{j,m,t-1} + \eta_{\sigma_j} e_{j,t} \left( \tilde{x}_{k,t} - \tilde{\sigma}_{j,m,t-1} \tilde{x}_{k,t} \right)^2, \hspace{1cm} (32)$$

$$\tilde{c}_{p,t} = \tilde{c}_{p,t-1} + \eta_{c_p} e_{p,t} \left( -\tilde{w}_{k,t+1} \right), \hspace{1cm} p = (1, 2, \ldots, n_c). \hspace{1cm} (33)$$

Step 6: if the stopping criteria $E$ given by (7) and the parameter variations $|a_i - a_{i-1}|$ are less than fixed small values, then stop; else, go to step 2; $t = t + 1$.

3.1.1. IBPG Algorithm Convergence Analysis. The adjustment procedures of FSOEAHW parameters are based on the learning rate choice ($\eta_a$). Too small $\eta_a$ guarantees convergence but with slow training speed, whereas too big $\eta_a$ guides to parametric divergence. In this section, Theorem 1 gives a selecting approach of a convenient $\eta_a$.

**Theorem 1.** The asymptotic convergence of the IBPG algorithm is guaranteed if each learning rate $\eta_a$ of each corresponding adjustable parameter $a$ is chosen to satisfy

$$0 < \eta_a < \frac{2}{\left( P_{a,\max} \right)^2}, \hspace{1cm} (34)$$

where $P_{a,\max} = \max_{s,t} |\partial \tilde{y}_s / \partial a_{s,t-1}|$.

**Proof.** See Appendix A.

3.2. Multi-Innovation Improved Backpropagation Gradient (MIIBPG) Algorithm. The multi-innovation technique is an effective tool to enhance the convergence rate of estimation methods. That is why it is joined with different estimation
The elementary idea is to expand the error term $e_{k,t}$ given by (8) and denoted scalar innovation to an innovation vector (called multi-innovation) [109, 111, 113].

Define an innovation vector $E_{i,k,t}$ as

$$E_{i,k,t} = \begin{bmatrix} e_{k,t} e_{k-1,t} \cdots e_{k-L+1,t} \end{bmatrix} \in \mathbb{R}^L,$$

where $L$ is a positive integer denoted innovation length and $e_{k-l,t} = y_{k-l} - \tilde{y}_{k-l,t}$ is the $l$th error scalar term at time $k-l$ ($l = 0, 1, \ldots, L$) expressed by

$$e_{k-l,t} = y_{k-l} - \left( \hat{y}_{k-l,t} - \sum_{p=1}^{n_c} \varepsilon_{p,t-1} \hat{w}_{k-l-p,t} \right) = y_{k-l} - \left( \hat{W}_{k-l-1}^T \hat{\psi}_{k-l,t} + \varepsilon_{l,t} \hat{\psi}_{k-l,t} \right) = y_{k-l} - \hat{W}_{k-l-1}^T \hat{\psi}_{k-l,t}$$

According to the recursive gradient algorithm minimizing the cost function given by equation (39), the MIIBPG technique provides a set of parameters updating equations listed in the following:

$$f_{j,k,t} = \frac{1}{2} E_{i,k,t}^T E_{j,k,t},$$

$$\tilde{c}_{p,t} = \tilde{c}_{p,t-1} - \eta_c \sum_{l=0}^{L-1} \frac{\partial \tilde{f}_{i,k,t}}{\partial \tilde{c}_{p,t-l}} \tilde{e}_{k-l,t} = \tilde{c}_{p,t-1} - \eta_c \sum_{l=0}^{L-1} e_{k-l,t} (-\tilde{w}_{k-l-1}, t), \quad (p: n_c),$$

$$\tilde{W}_{m,t} = \tilde{W}_{m,t-1} + \eta_{\tilde{W}} \sum_{l=0}^{L-1} e_{k-l,t} \tilde{\varphi}_{m,t-1}(\tilde{x}_{k-l,t}), \quad (m: N_z),$$

$$\tilde{\sigma}_{m,t} = \tilde{\sigma}_{m,t-1} + \eta_{\tilde{\sigma}} \sum_{l=0}^{L-1} e_{k-l,t} \left( \frac{\tilde{x}_{k-l,t} - \tilde{\sigma}_{m,t-1}}{\tilde{\sigma}_{m,t-1}^2} \right) \tilde{\varphi}_{m,t-1}(\tilde{x}_{k-l,t}) \left( \tilde{W}_{m,t-1} - \tilde{y}_{k-l,t} \right),$$

$$\tilde{\sigma}_{m,t} = \tilde{\sigma}_{m,t-1} + \eta_{\tilde{\sigma}} \sum_{l=0}^{L-1} e_{k-l,t} \left( \frac{\tilde{x}_{k-l,t} - \tilde{\sigma}_{m,t-1}}{\tilde{\sigma}_{m,t-1}^2} \right)^2 \tilde{\varphi}_{m,t-1}(\tilde{x}_{k-l,t}) \left( \tilde{W}_{m,t-1} - \tilde{y}_{k-l,t} \right),$$

$$\tilde{a}_{i,t} = \tilde{a}_{i,t-1} + \eta_{\tilde{a}} \sum_{l=0}^{L-1} e_{k-l,t} \frac{\partial \tilde{y}_{k-l,t}}{\partial \tilde{x}_{k-l,t}} \frac{\partial \tilde{x}_{k-l,t}}{\partial \tilde{a}_{i,t-1}}, \quad (i: n_a),$$

where

$$\frac{\partial \tilde{y}_{k-l,t}}{\partial \tilde{x}_{k-l,t}} = \sum_{m=1}^{N_z} \left( \frac{\tilde{x}_{k-l,t} - \tilde{\sigma}_{m,t-1}}{\tilde{\sigma}_{m,t-1}^2} \right) \tilde{\varphi}_{m,t-1}(\tilde{x}_{k-l,t}) \left( \tilde{W}_{m,t-1} - \tilde{y}_{k-l,t} \right),$$

algorithms such as gradient algorithm and least-square method in its recursive and iterative form [33, 105–113]. Particularly, it is well known that recursive gradient algorithm presents a slow convergence rate compared to other estimation approaches [113]. Many factors contribute to this disadvantage, principally its dependence on current data $\{u_k, y_k\}$ only. As a matter of fact, at each recursive step, gradient algorithm does not use previous data $\{u_{k-1}, y_{k-1}, \ldots, z_k\}$, $l = 1, 2, \ldots$. It does not have the capacity to use available data in the same step.

To overcome this problem, we propose in this section using multi-innovation IBPG (MIIBPG) algorithm based on previous finite data; that is, the MIIBPG algorithm uses the current and the previous data at each iteration $t$, which can improve the parameter estimation accuracy.
\[
\frac{\partial \tilde{x}_{k,t}}{\partial \tilde{a}_{j,t-1}} = -\tilde{x}_{k,t-1} - \sum_{s=1}^{n_s} \tilde{a}_{s,t-1} \frac{\partial \tilde{x}_{k,t-s}}{\partial \tilde{a}_{j,t-1}},
\]

\[
\tilde{b}_{j,t} = \tilde{b}_{j,t-1} + \eta_{\tilde{b}} \sum_{l=0}^{L-1} e_{k,t-l} \frac{\partial \tilde{y}_{k,t-l}}{\partial \tilde{b}_{j,t-1}}, \quad (j = 1, \ldots, n_b),
\]

\[
\frac{\partial \tilde{x}_{k,t}}{\partial \tilde{b}_{j,t-1}} = \tilde{h}_{k,t-1} - \sum_{s=1}^{n_s} \tilde{a}_{s,t-1} \frac{\partial \tilde{x}_{k,t-s}}{\partial \tilde{b}_{j,t-1}},
\]

\[
\tilde{W}_{i,n,t} = \tilde{W}_{i,n,t-1} + \eta_{\tilde{W}} \sum_{l=0}^{L-1} e_{k,t-l} \frac{\partial \tilde{y}_{k,t-l}}{\partial \tilde{x}_{k,t-l}} \frac{\partial \tilde{x}_{k,t}}{\partial \tilde{W}_{i,n,t-1}}
\]

\[
= \tilde{W}_{i,n,t-1} + \eta_{\tilde{W}} \sum_{l=0}^{L-1} e_{k,t-l} \frac{\partial \tilde{y}_{k,t-l}}{\partial \tilde{x}_{k,t-l}} \left( \sum_{j=1}^{n_b} \tilde{b}_{j,t-1} \tilde{g}_{j,n,t-1}(u_{k,t-j}) \right),
\]

\[
\tilde{\alpha}_{i,n,t} = \tilde{\alpha}_{i,n,t-1} + \eta_{\tilde{\alpha}} \sum_{l=0}^{L-1} e_{k,t-l} \frac{\partial \tilde{y}_{k,t-l}}{\partial \tilde{x}_{k,t-l}} \frac{\partial \tilde{x}_{k,t}}{\partial \tilde{\alpha}_{i,n,t-1}},
\]

\[
\frac{\partial \tilde{x}_{k,t}}{\partial \tilde{\alpha}_{i,n,t-1}} = \tilde{W}_{i,n,t-1} + \sum_{l=0}^{L-1} e_{k,t-l} \left( \frac{u_{k,t-l} - \tilde{\alpha}_{i,n,t-1}}{\tilde{\alpha}_{i,n,t-1}} \right)^2 \tilde{g}_{i,n,t-1}(u_{k,t-l}) \left( \tilde{W}_{i,n,t-1} - \tilde{h}_{k,t-l,j} \right),
\]

\[
\frac{\partial \tilde{x}_{k,t}}{\partial \tilde{\alpha}_{i,n,t-1}} = \sum_{l=0}^{L-1} \tilde{b}_{j,t-1} \left( \frac{u_{k,t-l} - \tilde{\alpha}_{i,n,t-1}}{\tilde{\alpha}_{i,n,t-1}} \right)^2 \tilde{g}_{i,n,t-1}(u_{k,t-l}) \left( \tilde{W}_{i,n,t-1} - \tilde{h}_{k,t-l,j} \right).
\]

The MIIBPG identification algorithm can be implemented based on the following steps:

Step 1: initialize the parameters of the linear dynamic parts of the two fuzzy systems and of the output noise randomly. Fix the learning rates and the innovation length $L_1, t = 1$.

At each iteration $t$ and for each sample $k$, repeat the following steps:

Step 2: for $l = 0, \ldots, L - 1$, and for given input $u(k - j - l)$, calculate $\tilde{h}_{k,t-l, j}$, $\tilde{x}_{k-t-l, j}$, $\tilde{y}_{k-t-l, j}$, and $e_{k,t-l}$ $(j = 1, \ldots, n_b)$ using equations (13)–(16) and (18) replacing $k$ with $k-l$.

Step 3: adjust the FSOEAWH parameters using equations (40)–(53).

Step 4: calculate noise’s estimate by equation (17) replacing $k$ with $k-l$.

Step 5: if the stopping criteria $E$ given by (7) and the parameter variations $|a_t - a_{t-1}|$ are less than fixed small values, then stop; else, go to step 2; $t = t + 1$.  

\[ e_{k,t-l} \]
3.2.1. MIIBPG Algorithm Convergence Analysis. This section surmises the convergence analysis of MIIBPG algorithm which can be applied to estimate a nonlinear stochastic system that can be illustrated by the proposed FSOEAHW model. The convergence properties of the MIIBPG algorithm are introduced by the following theorem.

**Theorem 2.** The asymptotic convergence of MIIBPG algorithm is guaranteed if the learning rate $\eta_\alpha$ of the corresponding adjustable parameter $\alpha$ is chosen to satisfy

$$0 < \eta_\alpha < \frac{2}{(P_{a, \max})^2},$$

where $P_{a, \max} = \max_k \left| \sum_{i=0}^{L-1} \frac{\partial y_{k-1,i}}{\partial \alpha_{i-1}} \right|$.

**Proof.** See Appendix B. □

3.3. Data Filtering-Based Improved Backpropagation Gradient (FIBPG) Algorithm. The data filtering technique in system identification is used to deal with the parameter estimation issues of systems disturbed by colored noises. Specifically, the elementary idea is to use a linear filter to filter input-output data so that the original systems with colored noises are transformed into new ones with white noises. Then the system’s structure is transformed to be simpler without changing the relationship between the system inputs and outputs. Owing to the advantages of the data filtering technique, it has been widely applied for different system identification and parameter estimation. For example, each of recursive, iterative, and hierarchical least-square algorithms is combined with data filtering technique in the identification problem of output-error autoregressive linear systems [114], two-input single-output controlled autoregressive moving average linear systems [115], Hammerstein finite impulse response systems with moving average output noise [116], and a multivariable box Jenkins-like system [117]. Similarly, a gradient algorithm in its recursive and iterative form is used with data filtering technique for identification of some classes of linear and nonlinear systems such as linear multivariable autoregressive moving average system [101], a state-space linear system in its observability canonical form disturbed by colored noise [118], and a block-oriented system like Hammerstein finite impulse response system with moving average output noise [119].

Stirred by the above description, we propose in this section combining IBPG algorithm (given in Section 3.1) with a data filtering technique to improve the accuracy rate of IBPG algorithm. In our case, we propose introducing a linear filter $C(z)$ to filter a measurement data in the purpose of parameter and intermediate variables estimation based on two criteria.

We define, respectively, the filtered intermediate variable $y_{j,k}$ and the filtered output $y_k$ as

$$y_{j,k} = C(z)y_k = y_k + \sum_{p=1}^{n} c_p y_{k-p},$$

$$y_{j,k} = C(z)y_{j,k} = \sum_{p=0}^{n} c_p y_{j,k-p}.$$  \hspace{1cm} (55)

We have $c_0 = 1$. Multiplying both sides of (11) by $C(z)$ and using (10), we obtain

$$C(z)y_k = C(z)y_{j,k} + v_k = C(z) \sum_{m=1}^{N_1} W_{1,m} g_{1,m}(x_k) + v_k.$$  \hspace{1cm} (56)

Thus, the filtered output can be written in matrix form as

$$y_{j,k} = W_1^T C(z) \psi_2(x_k) + v(k).$$  \hspace{1cm} (57)

In that, we can write the filtered observation vector as follows:

$$\psi_{2j}(x_k) = C(z) \psi_2(x_k).$$  \hspace{1cm} (58)

Consequently, the filtered system output can be defined as

$$y_{j,k} = y_{j,k} + v(k) = W_2^T \psi_{2j}(x_k) + v(k).$$  \hspace{1cm} (59)

The objective, now, is to develop an FSOEAHW parameters estimation approach based on gradient algorithm, using a filtered data measurement and minimizing a filtered estimation error given by

$$e_{j,k,t} = y_{j,k} - \tilde{y}_{j,k,t},$$

where the filtered estimated output $\tilde{y}_{j,k,t}$ can be expressed by the equivalent adjustable model:

$$\tilde{y}_{j,k,t} = \tilde{W}_2^{T}\tilde{\psi}_{2j,t-1}(\tilde{x}_{k,t}),$$

where $\tilde{\psi}_{2j,t-1}(\tilde{x}_{k,t})$ will be expressed later. It represents the estimated filtered observation vector in which the estimated intermediate variable $\tilde{x}_{k,t}$ is given by (14).

Nevertheless, $y_{j,k}$ in equation (59) depends on unknown parameter $c_{p}, p = (1, 2, \ldots, n_k)$. To neutralize this problem, we submit a two-step gradient algorithm. In the first step, the unknown parameter $c_{p}, p = (1, 2, \ldots, n_k)$, will be estimated minimizing the criteria $J_{j,k,t}$ (given by equation (62)). Subsequently, unknown parameters set $\beta = \left\{ \tilde{W}_{jm,\tilde{\alpha}_{m}}, \tilde{\beta}_{m}, \tilde{\alpha}_{m}, \tilde{\beta}_{j,\tilde{\alpha}_{j}}, \tilde{\beta}_{j,\tilde{\alpha}_{j}} \right\}$ with $(n = 1, \ldots, N_1, i = 1, \ldots, n_{\tilde{\alpha}_{i}}, j = 1, \ldots, n_{\tilde{\alpha}_{j}}, m = 1, \ldots, N_2)$ will be estimated minimizing the criteria $J_{j,k,t}$ given by equation (66).

$$J_{j,k,t} = \frac{1}{2} (w_k - \tilde{c}_{T_{k-1}}^T \tilde{\psi}_{k,t})^2.$$  \hspace{1cm} (62)

In reality, the quadratic criterion $J_{j,k,t}$ contains the unmeasurable output noise vector $w_k$, so the estimation of unknown parameter $c_{p}, p = (1, 2, \ldots, n_k)$, cannot be
implemented. Then, we propose replacing $w_k$ in (62) by its estimated $\tilde{w}_k$ given by equation (64). Afterward, the new quadratic criterion functions will be defined as

$$J_{k,t} = \frac{1}{2} e_{k,t}^2 = \frac{1}{2} (\tilde{w}_{k,t} - \tilde{c}_{i,t}^T \hat{p}_{k,t})^2,$$

(63)

where

$$\tilde{w}_{k,t} = y_k - \tilde{W}_{2,t}^T \tilde{q}_{2,t-1} (\tilde{x}_{k,t}).$$

Consequently, the estimation equation of parameter $c_p$, $p = (1, 2, \ldots, n)$, can be inferred using gradient algorithm minimizing the criteria $J_{k,t}$ as follows:

$$\tilde{c}_{p,t} = \tilde{c}_{p,t-1} + \eta_c e_{k,t} (\tilde{w}_{k-p,t}), \quad p = (1, 2, \ldots, n).$$

(65)

The estimation procedure of unknown parameters set $\beta$ is derived using the gradient method minimizing the quadratic criterion function defined as follows:

$$J_{i,k,t} = \frac{1}{2} e_{i,k,t}^2 = \frac{1}{2} (y_{i,k,t} - \tilde{y}_{i,k,t})^2,$$

(66)

where $y_{j,k,t}$ and $\tilde{y}_{j,k,t}$ are expressed, respectively, by

$$y_{j,k,t} = \sum_{p=0}^{n_c} \hat{c}_{p,t} y_{k-p},$$

(67)

$$\tilde{y}_{j,k,t} = \tilde{W}_{j,k,t}^T \tilde{q}_{2,t-1} (\tilde{x}_{k,t}),$$

(68)

where

$$\tilde{q}_{2,t-1} (\tilde{x}_{k,t}) = \tilde{C}_t (z) \tilde{q}_{2,t-1} (\tilde{x}_{k,t}).$$

(69)

We have

$$\tilde{C}_t (z) = 1 + \hat{c}_{n,t} z^{-1} + \cdots + \hat{c}_{1,t} z^{-n_c}.$$  

(70)

Therefore, the rest of parameters are adjusted using the following equations with $\tilde{c}_{n,t} = 1$:

$$\tilde{W}_{i,n,t} = \tilde{W}_{i,n,t-1} + \eta_{\tilde{W}_{i,n}} e_{i,k,t} \frac{\partial \tilde{y}_{k,t}}{\partial \tilde{W}_{i,n,t-1}} = \tilde{W}_{i,n,t-1} + \eta_{\tilde{W}_{i,n}} e_{i,k,t} \sum_{p=0}^{n_c} \hat{c}_{p,t} \frac{\partial \tilde{y}_{k-p,t}}{\partial \tilde{x}_{k-p,t}} \frac{\partial \tilde{x}_{k-p,t}}{\partial \tilde{W}_{i,n,t-1}} = \tilde{W}_{i,n,t},$$

(71)

where

$$\frac{\partial \tilde{x}_{k-p,t}}{\partial \tilde{W}_{i,n,t-1}} = \sum_{j=1}^{n_c} \tilde{\beta}_{j,i,t} g_{j,n,t-1}(u_{k-p-j}) - \sum_{j=1}^{n_c} \tilde{\alpha}_{j,i,t} \tilde{g}_{j,n,t-1}(u_{k-p-i-j}),$$

(72)

$$\frac{\partial \tilde{y}_{k-p,t}}{\partial \tilde{x}_{k-p,t}} = \sum_{j=1}^{N_c} \left( \tilde{x}_{k-p,t} - \frac{\tilde{x}_{j,n,t-1}}{2} \right) \tilde{g}_{j,n,t-1}(u_{k-p-j}) \tilde{g}_{j,n,t-1}(u_{k-p-j}) \tilde{W}_{j,n,t-1},$$

(73)

$$\tilde{\alpha}_{i,n,t} = \tilde{\alpha}_{i,n,t-1} + \eta_{\tilde{\alpha}_{i,n}} e_{i,k,t} \frac{\partial \tilde{y}_{k,t}}{\partial \tilde{\alpha}_{i,n,t-1}} = \tilde{\alpha}_{i,n,t-1} + \eta_{\tilde{\alpha}_{i,n}} e_{i,k,t} \sum_{p=0}^{n_c} \hat{c}_{p,t} \frac{\partial \tilde{y}_{k-p,t}}{\partial \tilde{x}_{k-p,t}} \frac{\partial \tilde{x}_{k-p,t}}{\partial \tilde{\alpha}_{i,n,t-1}},$$

(74)

$$\tilde{\beta}_{i,n,t} = \tilde{\beta}_{i,n,t-1} + \eta_{\tilde{\beta}_{i,n}} e_{i,k,t} \frac{\partial \tilde{y}_{k,t}}{\partial \tilde{\beta}_{i,n,t-1}} = \tilde{\beta}_{i,n,t-1} + \eta_{\tilde{\beta}_{i,n}} e_{i,k,t} \sum_{p=0}^{n_c} \hat{c}_{p,t} \frac{\partial \tilde{y}_{k-p,t}}{\partial \tilde{x}_{k-p,t}} \frac{\partial \tilde{x}_{k-p,t}}{\partial \tilde{\beta}_{i,n,t-1}},$$

(75)

$$\tilde{\sigma}_{i,n,t} = \tilde{\sigma}_{i,n,t-1} + \eta_{\tilde{\sigma}_{i,n}} e_{i,k,t} \frac{\partial \tilde{y}_{k,t}}{\partial \tilde{\sigma}_{i,n,t-1}} = \tilde{\sigma}_{i,n,t-1} + \eta_{\tilde{\sigma}_{i,n}} e_{i,k,t} \sum_{p=0}^{n_c} \hat{c}_{p,t} \frac{\partial \tilde{y}_{k-p,t}}{\partial \tilde{x}_{k-p,t}} \frac{\partial \tilde{x}_{k-p,t}}{\partial \tilde{\sigma}_{i,n,t-1}},$$

(76)
To summarize the FIBPG algorithm, we give the following steps:

Step 1: Initialize the parameters of the linear dynamic parts of the two fuzzy systems and the output noise parameters randomly. Fix the learning rates; \( t = 1 \).

At each iteration \( t \) and for each sample \( k \), repeat the following steps:

Step 2: For input \( u(k - P - j) \), calculate \( \hat{h}_{k-p-j, t} \) (\( j = 1, \ldots, n_\theta, p = 0, \ldots, n_c \)), \( \bar{x}_{k-p-j, t}, \bar{y}_{k-p-j, t} \), and \( \check{y}_{k-p-j, t} \) using, respectively, equations (13)–(16) replacing \( k \) with \( k-p \).

Step 3: Compute \( \hat{\bar{w}}_{k, t} \) using equation (64).

Step 4: Adjust parameter \( \check{c}_p \), \( p = (1, 2, \ldots, n_c) \), by equation (65).

Step 5: Compute \( \check{y}_{k,t}, \check{y}_{k,t} \) and \( e_{i,j,t} \) based, respectively, on equations (66)–(68).

Step 6: Update the rest of the parameters using equations (71)–(84).

Step 7: If the stopping criteria \( E \) given by (7) and the parameter variations \( (|a_i - a_i|) \) are less than fixed small values, then stop; else, go to step 2; \( t = t + 1 \).

### 3.3.1. FIBPG Convergence Analysis.

This subsection aims at the convergence analysis of the FIBPG which can be applied to estimate the parameters of a nonlinear system that can be described by the FSEOFHW model. The convergence properties of the FIBPG algorithm are determined in the following theorem.

**Theorem 3.** The asymptotic convergence of FIBPG is ensured if the learning rate \( \eta_i \) \( (p = 1, \ldots, n_c) \) and those of other FSOEAHW parameters satisfy the following conditions:
0 < \eta_c < \frac{2}{(P_c, \text{max})^2}, \quad \eta_p < \frac{2}{(P_p, \text{max})^2}, \quad \beta \text{ is an element of a parameter set } \\
\{\tilde{W}, \tilde{d}_i, \tilde{d}_j, \tilde{a}_i, \tilde{b}_j, \tilde{W}_m, \tilde{d}_m, \tilde{a}_m\} \text{ with } (n = 1, \ldots, N_1, i = 1, \ldots, n_i, j = 1, \ldots, n_j, m = 1, \ldots, N_2). \quad \square \\

Proof. See Appendix C.

3.4. Multi-Innovation and Data Filtering-Based Improved Backpropagation Gradient (MIFIBPG) Algorithm. Similar to Section 3.2 and to improve the convergence accuracy, we can combine multi-innovation approach with data filtering technique and then we obtain a multi-innovation FIBPG (MIFIBPG) algorithm.

Define \( E_{1l/k,t} \) and \( E_{1f/k,t} \) as follows:

\[
E_{2l/k,t} = \left[ e_{j/k,t} e_{j-1/k,t} : e_{j-1/k,t} - L + 1, t \right] = \begin{bmatrix}
\tilde{w}_{k,l} - \tilde{c}_{l-1} \tilde{y}_{k,l} \\
\tilde{w}_{k-l-1,t} - \tilde{c}_{l-1} \tilde{y}_{k-l-1,t} \\
\vdots \\
\tilde{w}_{k-l-1,t} - \tilde{c}_{l-1} \tilde{y}_{k-L,t}
\end{bmatrix},
\]

\[
E_{1l/k,t} = \left[ e_{j/k,t} e_{j-1/k,t} - L + 1, t \right] = \begin{bmatrix}
\tilde{y}_{j,k,t} - \tilde{y}_{j,k-1,t} - \tilde{y}_{j,k-1,t} - y_{j,k-1,t} \\
\tilde{y}_{j,k-1,t} - \tilde{y}_{j,k-1,t} - y_{j,k-1,t} - y_{j-1,k,t} - L + 1, t
\end{bmatrix},
\]

where

\[
e_{j,k,t} = \tilde{w}_{k,l,t} - \tilde{c}_{l} \tilde{y}_{k,l} = y_{k} - \tilde{W}^{T}_{2,t-1} \tilde{y}_{2,t-1} (\tilde{x}_{k,l,t}) - \tilde{c}_{l-1} \tilde{y}_{k,l-1}, \quad e_{j,k,t} = y_{j,k,t} - \tilde{y}_{j,k,t} - \tilde{y}_{j,k-1,t} = \sum_{l=0}^{n_l} \tilde{c}_{p,t} y_{j,k-1-p} - \tilde{W}_{2,t-1} \tilde{y}_{2,t-1} (\tilde{x}_{k,l,t}), \quad l = (0: L - 1) \text{ and } \tilde{c}_{0} = 1.
\]

Like the previous section and referring to MIIIBPG algorithm, initially the vector parameter \( \tilde{c} \) is adjusted in order to calculate the innovation vector error \( E_{1l/k,t} \). It is updated according to formula (91) by minimizing the quadratic criterion given by the following equation:

\[
J_{2l/k,t} = \frac{1}{2} E_{2l/k,t}^{T} E_{2l/k,t}, \quad \tilde{c}_{p,t} = \tilde{c}_{p,t-1} + \eta_{\tilde{c}} \sum_{t=0}^{L-1} e_{j,k-t} (\tilde{w}_{k-t-1,t}), \quad (p = 1: n_l), \quad (90)
\]

\[
\tilde{W}_{m,t} = \tilde{W}_{m,t-1} + \eta_{\tilde{W}} \sum_{t=0}^{L-1} e_{j,k-t} \left( \sum_{p=0}^{n_l} \tilde{c}_{p,t} \tilde{y}_{j,m-1,t} (\tilde{x}_{k-1-p}) \right), \quad m = (1, 2, \ldots, N_2), \quad (91)
\]

The adjustment equations of all other parameters are gotten using the MIFIBPG algorithm minimizing the cost function \( J_{1f/k,t} = (1/2) E_{1f/k,t}^{T} E_{1f/k,t}. \) They are listed as follows:
\[ \sigma_{j,m,t} = \sigma_{j,m,t-1} + \eta_{\sigma m} \sum_{l=0}^{L-1} e_{j,k-l} \sum_{p=0}^{n_p} \tilde{c}_{p,t} \left( \frac{\left( \tilde{x}_{k-l-p,t} - \sigma_{j,m,t-1} \right)}{\sigma_{j,m,t-1}^3} \right) \tilde{g}_{j,m,t-1}(\tilde{x}_{k-l-p,t})(\tilde{W}_{j,m,t-1} - \tilde{y}_{j,t} - l - \tilde{p}, t), \]  

(92)

\[ \tilde{a}_{i,t} = \tilde{a}_{i,t-1} + \eta_{\tilde{a} i} \sum_{l=0}^{L-1} e_{i,j-k-l} \sum_{p=0}^{n_p} \tilde{c}_{p,t} \left( \frac{\partial \tilde{y}_{j,k-l}}{\partial a_{i,t-1}} \right) \tilde{g}_{j,m,t-1}(\tilde{y}_{j,k} - l - \tilde{p}, t - \tilde{W}_{j,m,t-1}), \]  

(93)

where

\[ \frac{\partial \tilde{y}_{j,k-l-p,t}}{\partial x_{k-l-p,t}} = \sum_{m=1}^{N_i} \left( \frac{\left( \tilde{x}_{k-l-p,t} - \sigma_{j,m,t-1} \right)}{\sigma_{j,m,t-1}^2} \right) \tilde{g}_{j,m,t-1}(\tilde{x}_{k-l-p,t})(\tilde{y}_{j,k} - l - \tilde{p}, t - \tilde{W}_{j,m,t-1}) \right), \]  

(94)

\[ \frac{\partial \tilde{x}_{k-l-p,t}}{\partial a_{i,t-1}} = -\tilde{x}_{k-l-p,i} - \sum_{s=1}^{n_s} \tilde{a}_{s,t-1} \frac{\partial \tilde{x}_{k-l-p,i}}{\partial a_{i,t-1}} \]  

(95)

\[ \tilde{b}_{j,t} = \tilde{b}_{j,t-1} + \eta_{\tilde{b} j} \sum_{l=0}^{L-1} e_{i,j,k-l} \sum_{p=0}^{n_p} \tilde{c}_{p,t} \frac{\partial \tilde{y}_{j,k-l}}{\partial b_{j,t-1}} \tilde{b}_{j,t-1} + \eta_{\tilde{b} j} \sum_{l=0}^{L-1} e_{i,j,k-l} \sum_{p=0}^{n_p} \tilde{c}_{p,t} \frac{\partial \tilde{y}_{j,k-l}}{\partial b_{j,t-1}} \frac{\tau}{\partial b_{j,t-1}} \]  

(96)

\[ \tilde{W}_{i,n,t} = \tilde{W}_{i,n,t-1} + \eta_{\tilde{W} i} \sum_{l=0}^{L-1} e_{i,j,k-l} \frac{\partial \tilde{y}_{j,k-l}}{\partial W_{i,n,t-1}} = \tilde{W}_{i,n,t-1} + \eta_{\tilde{W} i} \sum_{l=0}^{L-1} e_{i,j,k-l} \sum_{p=0}^{n_p} \tilde{c}_{p,t} \frac{\partial \tilde{y}_{j,k-l}}{\partial W_{i,n,t-1}} \frac{\partial \tilde{x}_{k-l-p,t}}{\partial W_{i,n,t-1}}, \]  

(97)

where

\[ \frac{\partial \tilde{x}_{k-l-p,t}}{\partial W_{i,n,t-1}} = \sum_{j=1}^{n} \tilde{b}_{j,t-1} \tilde{g}_{j,n,t-1}(u_{k-l-p,i}) - \sum_{j=1}^{n} \tilde{a}_{i,j-1} \sum_{j=1}^{n} \tilde{b}_{j,t-1} \tilde{g}_{j,n,t-1}(u_{k-l-p,i}), \]  

(98)

\[ \tilde{a}_{i,n,t} = \tilde{a}_{i,n,t-1} + \eta_{\tilde{a} i} \sum_{l=0}^{L-1} e_{i,j,k-l} \sum_{p=0}^{n_p} \tilde{c}_{p,t} \frac{\partial \tilde{y}_{j,k-l}}{\partial a_{i,t-1}} \tilde{g}_{j,n,t-1}(u_{k-l-p,i}) \]  

(99)

\[ \frac{\partial \tilde{x}_{k-l-p,i}}{\partial a_{i,n,t-1}} = \sum_{j=1}^{n} \tilde{b}_{j,t-1} \left( \frac{u_{k-l-p,i} - \tilde{a}_{i,n,t-1}}{\tilde{a}_{i,n,t-1}} \right) \tilde{g}_{j,n,t-1}(u_{k-l-p,i}) \left( \tilde{W}_{i,n,t-1} - \tilde{h}_{k-l-p,i} \right) \right) - \left( \sum_{j=1}^{n} \tilde{a}_{i,j-1} \tilde{b}_{j,t-1} \left( \frac{u_{k-l-p,i} - \tilde{a}_{i,n,t-1}}{\tilde{a}_{i,n,t-1}} \right) \tilde{g}_{j,n,t-1}(u_{k-l-p,i}) \left( \tilde{W}_{i,n,t-1} - \tilde{h}_{k-l-p,i} \right) \right), \]  

(100)

\[ \tilde{a}_{i,n,t} = \tilde{a}_{i,n,t-1} + \eta_{\tilde{a} i} \sum_{l=0}^{L-1} e_{i,j,k-l} \frac{\partial \tilde{y}_{j,k-l}}{\partial a_{i,n,t-1}} \tilde{g}_{j,n,t-1}(u_{k-l-p,i}) \]  

(101)
Theorem 4. The convergence of the MIFIBPG algorithm is ensured if the learning rates \( \eta_r \) and \( \eta_\beta \), where \( \beta \) represents each element of the parameter set 
\[
\{ \bar{W}_n, \bar{n}, \bar{m}, \bar{b}_j, \bar{W}_m, \bar{n}, \bar{m}, \bar{b}_j \}, \quad \text{with} \quad (n = 1, \ldots, N_1, i = 1, \ldots, n, j = 1, \ldots, n, m = 1, \ldots, N_2),
\]

satisfy the following inequalities:
\[
0 < \eta_r \leq \frac{2}{(P_{c, \max})^2}, \quad 0 < \eta_\beta \leq \frac{2}{(P_{\beta, \max})^2},
\]

(103)

where \( P_{c, \max} = \max_k |\sum_{l=0}^{L-1} \partial \bar{w}_{k-l} / \partial \bar{c}_{l}^{\beta_{l-1}}| \), \( P_{\beta, \max} = \max_k |\sum_{l=0}^{L-1} \partial \bar{y}_{k-l} / \partial \bar{c}_{l}^{\beta_{l-1}}| \).

The theorem’s proof is similar to those given in Appendices B and C for the MIIIBPG and FIBPG algorithms convergence study.

4. Example

In this section, we achieve an illustration simulation in order to test the performance and the efficiency of the developed algorithms. Therefore, we consider a nonlinear stochastic system given by stochastic output-error Hammerstein–Wiener mathematical model in which its blocks are represented using the following expressions:

\[
h_k = f_1(u_k) = \begin{cases} 
0, & \text{if } u_k \leq 0.15, \\
1 - \exp(-u_k) - 0.2 \exp(-2u_k) \cos(3u_k), & \text{if } 0.15 < u_k < 3.5, \\
0.96, & \text{if } u_k \geq 3.5,
\end{cases}
\]

\[
G(z) = \frac{B(z)}{A(z)} = \frac{0.15z^{-2}}{1 + 0.15z^{-2}},
\]

(104)

\[
y'_{ik} = f_2(x_k) = \begin{cases} 
0, & \text{if } x_k \leq -0.3, \\
1 + \tanh(x_k - 1), & \text{if } -0.3 < x_k < 1.3, \\
0.96, & \text{if } x_k \geq 1.3,
\end{cases}
\]

\[
\frac{1}{C(z)} = \frac{1}{1 + c_1z^{-1}} = \frac{1}{1 + 0.15z^{-1}}.
\]

A set of 9000 random values between 0 and 4 are used as input \( u_k \). \( \{v_k\} \) is a white noise sequence with zero mean and noise-to-signal ratio as \( \gamma_{ns} = \sqrt{\text{var}(v_k) / \text{var}(y_{ik})} \times 100 = 13\% \). The input \( u_k \) and the corresponding output \( y_k \) are presented in Figure 3.
function parameters $\tilde{\alpha}_i$ and $\tilde{\sigma}_m$ ($n = 1, \ldots, 8$) are initialized, respectively, in intervals $[0, 0.4]$ and $[0.1, 0.35]$.

The initial values of parameters $\tilde{\alpha}_i, \tilde{\beta}_j, \tilde{c}_m, i = 1, 2, j = 1, 2,$ and $p = 1$ are considered, respectively, as $0.25, -0.25, 0.5, 0.1$ and $0.26$.

The output nonlinear part is depicted by 6 fuzzy rules where the consequent parts parameters are initialized randomly between 0 and 1.5 and the membership function parameters $\tilde{\alpha}_{m,0}$ and $\tilde{\sigma}_{m,0}$ are taken, respectively, in intervals $[0, 0.5]$ and $[0.15, 0.25]$. Numerical simulations are realized to estimate the parameters intervening in the discrete time FSOEAHW.

4.1. Parametric Estimation Using IBPG Algorithm. We propose estimating the FSOEAHW parameters using the IBPG algorithm described in Section 3.1. Thus, we demonstrate the estimation results by giving the evolution curves of the intermediate parameters and output estimates $\hat{h}_{k,t}$, $\hat{x}_{k,t}$, $\hat{y}_{k,t}$, and $\hat{y}_{k,t}$ in Figure 4.

In addition, Figure 5 represents the evolution curves of the estimation error $e_{k,t}$ and the parametric distance $\delta_t = \|\theta - \hat{\theta}_t\|/\|\theta\|$ in which $\theta = [\tilde{\alpha}_f^T, \tilde{c}_f^T]$ and $\hat{\theta}_t = [\hat{\alpha}_f^T, \hat{c}_f^T]$. In Table 1, we present some obtained numerical values of vector parameter estimates $\hat{\theta}_t$ and the corresponding parametric error for $t = 0, 2000, 4000, 6000,$ and $9000$. In addition, to evaluate the estimation quality, we depict in Table 2 the estimation errors $e_h = \|h - \hat{h}\|/\|h\|$, $e_x = \|x - \hat{x}\|/\|x\|$, $e_y = \|y - \tilde{y}\|/\|y\|$, $e_\omega = \|\omega - \tilde{\omega}\|/\|\omega\|$, and $e_c = \|y - \tilde{y}\|/\|y\|$.

As is noticed in the above-presented simulation results, a satisfactory estimation is achieved with the proposed IBPG algorithm despite the existence of the noise acting on the process output without restrictive assumptions. Furthermore, the parameter estimation errors gradually become, iteratively, smaller. Thus, the parameter estimates are closer to the true values and the fuzzy models generate a satisfactory estimation accuracy of the nonlinear input and output blocks.

For comparison purpose, we propose giving simulation results in which the input and output nonlinear parts are estimated by two independent polynomial functions expressed by the following equations:

$$h_k = \sum_{n=1}^{N_1} \chi_n h_{k,n}^n,$$

$$y_{k,t} = \sum_{m=1}^{N_2} \beta_m y_{k,m}^m,$$

where $\chi_n$ ($n = 1: N_1$) and $\beta_m$ ($m = 1: N_2$) are constant parameters. Then the corresponding output-error autoregressive HW (OEAHW) system parameters are $\chi_n$, $a_i$, $b_j$, $c_p$, and $\beta_m$ ($n = 1: 3, i = 1: 2, j = 1: 2, p = 1, m = 1: 3$). The proposed IBPG algorithm is adopted for parameter adjustment based on updating equations given in Appendix D. It should be noticed that OEAHW is inspired from [24], where the IBPG algorithm is used to estimate a polynomial Wiener model.

In this sense, Figure 6 illustrates the evolution curves of the estimation and parametric errors for the OEAHW and FSOEAHW models. Table 3 presents the OEAHW quality estimation.

The obtained results confirm that FSOEAHW model generates a more satisfactory estimation accuracy compared to the OEAHW model with polynomial form. Due to the ability of the fuzzy systems for providing a good approximation of the two nonlinear parts, the proposed FSOEAHW reaches a good performance with a comparatively small amount of calculation and the resulting estimated model can capture systems dynamics correctly.

4.2. Parametric Estimation Using MIIBPG Algorithm. Using the MIIBPG algorithm for FSOEAHW parameter identification, the vector parameter estimates $\hat{\theta}_t$ and their corresponding parametric errors $\delta_t$ ($t = 0, 2000, 4000, 6000, 9000$) are given in Table 4 for $L = 2$. 

---

**Figure 3:** The system's input and output versus $k$. 

**Figure 4:** The system's input and output versus $k$. 

**Figure 5:** The system's input and output versus $k$. 

**Figure 6:** The system's input and output versus $k$. 

**Table 1:** The obtained numerical values of vector parameter estimates $\hat{\theta}_t$ and the corresponding parametric error for $t = 0, 2000, 4000, 6000,$ and $9000$. 

**Table 2:** The estimation errors $e_h = \|h - \hat{h}\|/\|h\|$, $e_x = \|x - \hat{x}\|/\|x\|$, $e_y = \|y - \tilde{y}\|/\|y\|$, $e_\omega = \|\omega - \tilde{\omega}\|/\|\omega\|$, and $e_c = \|y - \tilde{y}\|/\|y\|$. 

**Table 3:** The OEAHW quality estimation. 

**Table 4:** The vector parameter estimates $\hat{\theta}_t$ and their corresponding parametric errors $\delta_t$ ($t = 0, 2000, 4000, 6000, 9000$) are given in Table 4 for $L = 2$. 

---
Furthermore, for same values of $L$, the evolution curves of $\delta$ and estimation error variances $\sigma^2_k,t$ are presented, respectively, in Figures 7 and 8 with $\sigma^2_k,t = \sum_{i=1}^{K} (e_{i,k,t} - \bar{e}_{i,k,t})^2/\bar{k}$ in which $\bar{e}_{i,k,t}$ is a statistical mean of the estimation error. The estimation errors of system variables are given in Table 5.

By examining Tables 4 and 5 and Figures 7 and 8, we can affirm that the parameter and estimation errors gradually
become smaller with the iterative variable and innovation length increasing. In addition, the developed MIIBPG algorithm drives to have higher performance than the standard IBPG algorithm. In fact, the system parameter estimates converge to their true values and the estimation errors decrease to smaller values. In this direction, the

| Table 1: The vector parameter estimates \( \hat{\theta}_t \) and their parametric errors \( \delta_t \) for the FSOEAHW model using IBPG algorithm. |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Algorithm | \( t, k \) | \( \hat{a}_1 \) | \( \hat{a}_2 \) | \( \hat{b}_1 \) | \( \hat{b}_2 \) | \( \hat{e}_1 \) | \( \hat{\delta}_1 \) |
| IBPG | 0.000 | 0.2500 | -0.2500 | 0.5000 | 0.1000 | 0.2600 | 0.8252 |
| | 2000 | -0.0338 | 0.2503 | 0.9796 | 0.1365 | 0.1112 | 0.1808 |
| | 4000 | -0.1761 | 0.3322 | 0.9758 | 0.1491 | 0.1507 | 0.0350 |
| | 6000 | -0.1930 | 0.3757 | 1.0022 | 0.1518 | 0.1577 | 0.0254 |

| Table 2: Estimation errors for the FSOEAHW model using IBPG algorithm. |
|----------------|----------------|----------------|----------------|----------------|----------------|
| Algorithm | \( \epsilon_h \) | \( \epsilon_x \) | \( \epsilon_y \) | \( \epsilon_w \) | \( \epsilon_y \) |
| IBPG | 0.0021 | 0.0061 | 0.0007 | 0.5102 | 0.0017 |

| Figure 6: Evolution curves of estimation error (a) and parametric error (b) for the FSOEAHW and OEAHW models. |

| Table 3: Estimation errors for the OEAHW model using IBPG algorithm. |
|----------------|----------------|----------------|----------------|----------------|----------------|
| Algorithm | \( \epsilon_h \) | \( \epsilon_x \) | \( \epsilon_y \) | \( \epsilon_w \) | \( \epsilon_y \) |
| IBPG | 0.0131 | 0.0212 | 0.0243 | 18.6439 | 0.0315 |
Table 4: The vector parameter estimates \( \hat{\theta}_t \) and their parametric errors \( \delta_t \) for the FSOEAW model using MIIBPG algorithm with \( L = 2, 5, \) and 7.

| \( L \) | \( t, k \) | \( \hat{a}_1 \) | \( \hat{a}_2 \) | \( \hat{b}_1 \) | \( \hat{b}_2 \) | \( \hat{c}_1 \) | \( \delta_t \) |
|-----|-----|-----|-----|-----|-----|-----|-----|
| 2   | 0.000 | 0.2500 | -0.2500 | 0.5000 | 0.1000 | 0.2600 | 0.8252 |
|     | 2000  | -0.1036 | 0.3009 | 0.9897 | 0.1380 | 0.1514 | 0.0975 |
| 5   | 4000  | -0.1901 | 0.3409 | 0.9800 | 0.1487 | 0.1514 | 0.0213 |
|     | 6000  | -0.1900 | 0.3542 | 1.0125 | 0.1495 | 0.1514 | 0.0172 |
|     | 9000  | -0.2001 | 0.3512 | 1.0131 | 0.1513 | 0.1514 | 0.0121 |

| 7   | 0.000 | 0.2500 | -0.2500 | 0.5000 | 0.1000 | 0.2600 | 0.8252 |
|     | 2000  | -0.1705 | 0.3480 | 0.9420 | 0.1444 | 0.1542 | 0.0428 |
|     | 4000  | -0.2019 | 0.3493 | 0.9608 | 0.1511 | 0.1542 | 0.0333 |
|     | 6000  | -0.1938 | 0.3572 | 0.9913 | 0.1515 | 0.1542 | 0.0089 |
|     | 9000  | -0.1999 | 0.3529 | 1.0081 | 0.1522 | 0.1542 | 0.0090 |

**Figure 7:** Evolution curves of parametric error for the FSOEAW model using MIIBPG algorithm.

**Figure 8:** Evolution curves of estimation error variances for the FSOEAW model using MIIBPG algorithm.
Table 5: Estimation errors for the FSOEAHW model using MIIBPG algorithm with $L = 2$, 5, and 7.

| $L$ | $e_h$   | $e_y$   | $e_{jy}$ | $e_w$ | $e_y$ |
|-----|---------|---------|----------|-------|-------|
| 2   | 0.0008  | 0.0046  | 0.0007   | 0.4740| 0.0018|
| 5   | 0.0006  | 0.0033  | 0.0006   | 0.3830| 0.0016|
| 7   | 0.0006  | 0.0013  | 0.0005   | 0.3026| 0.0015|

Figure 9: Evolution curves of estimation error variance for the FSOEAHW model using FIBPG algorithm.

Figure 10: Evolution curves of parametric error for the FSOEAHW model using FIBPG algorithm.

Table 6: The vector parameter estimates $\hat{\theta}_t$ and their parametric errors $\delta_t$ for the FSOEAHW model using FIBPG algorithm.

| FIBPG algorithm | $t$, $k$ | $\hat{\alpha}_1$ | $\hat{\alpha}_2$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\gamma}_1$ | $\hat{\delta}_t$ |
|----------------|----------|-------------------|-------------------|------------------|------------------|------------------|------------------|
| 0000           | 0.2500   | -0.2500           | 0.5000            | 0.1000           | 0.2600           | 0.8252           |
| 2000           | -0.0292  | 0.2634            | 1.0185            | 0.1305           | 0.1445           | 0.1748           |
| 4000           | -0.1675  | 0.3268            | 1.0057            | 0.1422           | 0.1445           | 0.0384           |
| 6000           | -0.1801  | 0.3558            | 1.0046            | 0.1438           | 0.1445           | 0.0203           |
| 9000           | -0.1995  | 0.3548            | 1.0046            | 0.1457           | 0.1445           | 0.0088           |
The proposed multi-innovation approach can estimate the parameters effectively and increasing the innovation length can improve parameter and variable estimation accuracy and accelerate the convergence rate because the algorithm uses more information in each iteration. Hence, the proposed algorithm is effective for the FSOEAHW model identification.

### 4.3. Parametric Estimation Using FIBPG Algorithm

In this section, the efficiency of the FIBPG algorithm compared to the IBPG algorithm is well noticed by visualizing the evolutions curves of the estimation error variance $\sigma_{e_{1,k,t}}^2$ and a parametric error $\delta_t$ given, respectively, in Figures 9 and 10. In fact, a smooth and a more precise parametric convergence is ensured. This is more approved by examining Tables 6 and 7 that show the estimation accuracy of FSOEAHW model.

### 4.4. Parametric Estimation Using MIFIBPG Algorithm

In this part, we exploit the MIFIBPG algorithm to estimate the parameters of the FSOEAHW. Thus, Tables 8 and 9 present the parameter estimates and their estimation errors for $L = 2$, 5, and 7. For the same innovation length values, Figures 11 and 12 give the estimation error variance $\sigma_{e_{1,k,t}}^2$ and the parametric distance $\delta_t$.

The proposed multi-innovation approach combined with data filtering technique can estimate the parameters effectively and smoothly and increasing the innovation length gives a good parameter estimation accuracy and accelerates the convergence rate and then improves the gradient algorithm drawbacks.

### Tables

#### Table 7: Estimation errors for the FSOEAHW model using FIBPG algorithm.

| FIBPG algorithm | $e_h$ | $e_x$ | $e_y$ | $e_\omega$ | $e_y$ |
|-----------------|------|------|------|----------|------|
|                 | 0.0007 | 0.0047 | 0.0006 | 0.4608 | 0.0015 |

#### Table 8: The vector parameter estimates $\hat{\theta}_t$ and their parametric errors $\delta_t$ for the FSOEAHW model using MIFIBPG algorithm with $L = 2$, 5, and 7.

| $L$ | $t$, $k$ | $\hat{a}_1$ | $\hat{a}_2$ | $\hat{b}_1$ | $\hat{b}_2$ | $\hat{c}_1$ | $\delta_t$ |
|-----|---------|-------------|-------------|-------------|-------------|-------------|----------|
| 2   | 0.0000 | 0.2500     | -0.2500     | 0.5000      | 0.1000      | 0.2600      | 0.8252   |
|     | 2000   | -0.1118    | 0.3071      | 1.0059      | 0.1346      | 0.1543      | 0.0885   |
|     | 4000   | -0.1921    | 0.3354      | 1.0058      | 0.1438      | 0.1543      | 0.0174   |
|     | 6000   | -0.1887    | 0.3501      | 1.0058      | 0.1438      | 0.1543      | 0.0124   |
|     | 9000   | -0.2006    | 0.3489      | 1.0058      | 0.1462      | 0.1543      | 0.0075   |
| 5   | 0.0000 | 0.2500     | -0.2500     | 0.5000      | 0.1000      | 0.2600      | 0.8252   |
|     | 2000   | -0.1932    | 0.3481      | 1.0039      | 0.1472      | 0.1478      | 0.0068   |
|     | 4000   | -0.2121    | 0.3427      | 1.0029      | 0.1507      | 0.1478      | 0.0128   |
|     | 6000   | -0.1909    | 0.3528      | 1.0029      | 0.1495      | 0.1478      | 0.0066   |
|     | 9000   | -0.1984    | 0.3505      | 1.0029      | 0.1503      | 0.1478      | 0.0037   |
| 7   | 0.0000 | 0.2500     | -0.2500     | 0.5000      | 0.1000      | 0.2600      | 0.8252   |
|     | 2000   | -0.2239    | 0.3774      | 0.9999      | 0.1489      | 0.1501      | 0.0332   |
|     | 4000   | -0.2165    | 0.3429      | 1.0000      | 0.1505      | 0.1501      | 0.0144   |
|     | 6000   | -0.1944    | 0.3531      | 1.0000      | 0.1495      | 0.1501      | 0.0034   |
|     | 9000   | -0.1999    | 0.3506      | 1.0000      | 0.1499      | 0.1501      | 0.0006   |

#### Table 9: Estimation errors for the FSOEAHW model using MIFIBPG algorithm with $L = 2$, 5, and 7.

| MIFIBPG algorithm | $e_h$ | $e_x$ | $e_y$ | $e_\omega$ | $e_y$ |
|-------------------|------|------|------|----------|------|
| 2                 | 0.0005 | 0.0038 | 0.0005 | 0.4076    | 0.0016 |
| 5                 | 0.0005 | 0.0016 | 0.0005 | 0.3325    | 0.0015 |
| 7                 | 0.0006 | 0.0022 | 0.0005 | 0.3046    | 0.0015 |
5. Conclusion

This paper deals with modeling and identification of SOEAHW systems based on an artificial intelligence technique. Two fuzzy models with adjustable parameters are used to identify the input and output nonlinear blocks. The output nonlinearity may be noninvertible. A fuzzy model having as input the linear dynamic output estimate approximates it. The four given parameter estimation algorithms (IBPG, MIBPG, FIBPG, and MIFIBPG) are based on recursive gradient algorithm, multi-innovation technique, and data filtering approach. They have convergence properties and validity for online implementation. Simulation results validate the effectiveness of the proposed algorithms. However, the drawbacks of the proposed algorithms lie, particularly, in the updated parameter initialization and the learning rate choice. In fact, it is well known that adjustment equations based on gradient algorithm depend directly on these parameter choices and these have a crucial effect on the convergence and accuracy. The work presented in this paper can be extended to other classes of stochastic nonlinear multivariable systems with colored noise. Furthermore, we can suggest proposing an adaptive fuzzy control of nonlinear industrial process (such as practical hydraulic process) based on the presented SOEAHW model. Finally, a more suitable approach for an appropriate learning rate’s research and other optimization techniques can be proposed to give better performance.

Appendix

A: IBPG Algorithm’s Convergence Proof

Consider a discrete-type Lyapunov function as [15, 51]

\[ V_k = \frac{1}{2}e_k^2. \]  

(A.1)

For simplicity of writing, the index \( t \) will be omitted. The alternation of the Lyapunov function due to the training process is hence

\[ \Delta V_k = V_{k+1} - V_k = \frac{1}{2} \left[ e_{k+1}^2 - e_k^2 \right] = \Delta \psi_k \left[ e_k^2 + \frac{1}{2} \Delta e_k^2 \right]. \]  

(A.2)

where the error difference \( \Delta e_k \) can be computed by

\[ \Delta e_k = e_{k+1} - e_k. \]  

(A.3)

Denote by \( \rho = [I \dot{W}_{1n}, \dot{a}_{1n}, \dot{\bar{b}}_1, \dot{W}_{2m}, \dot{a}_{2m}, \dot{\bar{c}}_p]^T \) a parameter vector to be adjusted based on IBPG algorithm and its corresponding learning rate \( \eta_i = \text{diag} \left\{ \eta_{W_{1n}}, \eta_{a_{1n}}, \eta_{\bar{b}_1}, \eta_{W_{2m}}, \eta_{a_{2m}}, \eta_{\bar{c}_p} \right\} \) where \( n = 1, \ldots, N_1, i = 1, \ldots, n_i, j = 1, \ldots, n_j, m = 1, \ldots, N_2, \) \( p = 1, \ldots, n_p \). Then, \( \Delta \psi_k \) can be computed as

\[ \Delta \psi_k = \left[ \frac{\partial \psi_k}{\partial \rho} \right]^T \Delta \rho, \]  

(A.4)

where

\[ \frac{\partial \psi_k}{\partial \rho} = \left[ \frac{\partial \psi_k}{\partial W_{1n}}, \frac{\partial \psi_k}{\partial a_{1n}}, \frac{\partial \psi_k}{\partial \bar{b}_1}, \frac{\partial \psi_k}{\partial W_{2m}}, \frac{\partial \psi_k}{\partial a_{2m}}, \frac{\partial \psi_k}{\partial \bar{c}_p} \right]^T. \]  

(A.5)

\[ \Delta \rho = \left[ \Delta \dot{W}_{1n}, \Delta \dot{a}_{1n}, \Delta \bar{b}_1, \Delta \dot{W}_{2m}, \Delta \dot{a}_{2m}, \Delta \bar{c}_p \right]^T. \]  

(A.6)

\( \Delta \rho \) denotes the change in parameter vector \( \rho \) formulated using equations (20) and (A.3) as

\[ \Delta \rho = -\eta_i e_k \frac{\partial \psi_k}{\partial \rho} = \eta_i e_k \frac{\partial \psi_k}{\partial \rho}, \]  

(A.7)

where

\[ \frac{\partial \psi_k}{\partial \rho} = \left[ \frac{\partial \psi_k}{\partial W_{1n}}, \frac{\partial \psi_k}{\partial a_{1n}}, \frac{\partial \psi_k}{\partial \bar{b}_1}, \frac{\partial \psi_k}{\partial W_{2m}}, \frac{\partial \psi_k}{\partial a_{2m}}, \frac{\partial \psi_k}{\partial \bar{c}_p} \right]^T. \]  

(A.8)

The Lyapunov function (A.1) is a time-invariant positive definite function. Thus, from (A.3)–(A.8), equation (A.2) will be expressed as

\[ \Delta V_k = \chi \Delta e_k \]  

where

\[ \chi = \sum_{t} \left[ \frac{1}{2} \eta_i \left( \frac{\partial \psi_k}{\partial \rho} \right)^2 \frac{1}{2} \left( 2 - \eta_i \left( \frac{\partial \psi_k}{\partial \rho} \right) \right) \right]. \]  

(A.10)

Now, it is obvious that the convergence of the FOEAHW is guaranteed if \( 0 < \chi \). Then we get conditions of Theorem 1.

This completes the proof.
**B: MIIBPG Algorithm’s Convergence Proof**

In this case, we define a Lyapunov function as

\[ V_k = \frac{1}{2} E_k^T E_k. \]  

(B.1)

The alternation of \( V_k \) due to the training process is hence

\[
\Delta V_k = V_{k+1} - V_k = \frac{1}{2} \left[ E_{k+1}^T E_{k+1} - E_k^T E_k \right] = \frac{1}{2} \sum_{i=0}^{L-1} \left( e_{k+1} - e_{k} \right)^2.
\]  

(B.2)

where the error difference \( \Delta e_{k-1} \) can be computed by

\[
\Delta e_{k-1} = e_{k+1} - e_k.
\]  

(B.3)

where

\[
\frac{\partial e_{k-1}}{\partial \rho} = \left[ \frac{\partial e_{k-1}}{\partial W_n^m}, \frac{\partial e_{k-1}}{\partial \sigma_n^m}, \frac{\partial e_{k-1}}{\partial \sigma_m^m}, \frac{\partial e_{k-1}}{\partial \sigma_n^k}, \frac{\partial e_{k-1}}{\partial \sigma_m^k}, \frac{\partial e_{k-1}}{\partial \sigma_n^L}, \frac{\partial e_{k-1}}{\partial \sigma_m^L} \right]^T.
\]  

(B.4)

\[ \Delta \rho \] is a variation of parameter vector which can be expressed by

\[
\Delta \rho = -\eta e_{k-1} \frac{\partial \tilde{\rho}_{k-1}}{\partial \rho} = -\eta e_{k-1} \frac{\partial \tilde{\rho}_{k-1}}{\partial \rho}.
\]  

(B.5)

Using equations (B.3)–(B.6), equation (B.2) is reformulated as

\[
\Delta V_k = \sum_{i=0}^{L-1} \left( e_{k-1} + \frac{1}{2} \left( e_{k-1} \frac{\partial \tilde{\rho}_{k-1}}{\partial \rho} \right) \right) = \sum_{i=0}^{L-1} \left[ \frac{\partial \tilde{\rho}_{k-1}}{\partial \rho} \right] \eta e_{k-1} \frac{\partial \tilde{\rho}_{k-1}}{\partial \rho}.
\]  

(B.7)
where
\[ x_i = \sum z_i \eta_i x_i, \]  
\[ x_i = \left( \frac{\partial \bar{y}_{k-1}}{\partial p_s} \right)^2 \left( 2 - \eta_i \left( \frac{\partial \bar{y}_{k-1}}{\partial p_s} \right)^2 \right). \]  

(B.8)

To ensure that \( \Delta V_k < 0 \), it is sufficient to have conditions described in Theorem 2.

C: FIBPG Algorithm’s Convergence Proof

To study the FIBPG algorithm’s convergence, we define a Lyapunov candidate \( V \) which is positive definite, expressed according to the following equation:
\[ V_{\text{FIBPG}} = V_{1k} + V_{2k}, \]  
where the first discrete type Lyapunov function can be defined as
\[ V_{1k} = \frac{1}{2} \epsilon_{k}^2. \]  

(C.2)

The alternation of the Lyapunov function due to the training process is hence
\[ \Delta V_{1k} = V_{1k+1} - V_{1k} = \frac{1}{2} \left[ \epsilon_{k+1}^2 - \epsilon_{k}^2 \right] = \Delta \epsilon_{k} \left[ \epsilon_{k} + \frac{1}{2} \Delta \epsilon_{k} \right], \]  
where
\[ \Delta \epsilon_{k} = \epsilon_{k+1} - \epsilon_{k}. \]  

(C.4)

Like Appendix A, we have the following expression:
\[ \Delta V_{1k} = \frac{1}{2} \epsilon_{k}^2, \]  
where
\[ \chi_2 = \frac{1}{2} \eta_i \left( \bar{w}_{k-p} \right)^2 \left( 2 - \eta_i \left( \bar{w}_{k-p} \right)^2 \right). \]  

(C.6)

In order that \( \Delta V_{1k} < 0 \), equation (B.5) must be satisfied:
\[ 0 < \eta_i < \frac{2}{\left( \bar{w}_{k-p} \right)^2}. \]  

(C.7)

The second part of the difference is
\[ \Delta V_{2k} = V_{2k+1} - V_{2k} = \frac{1}{2} \left[ \epsilon_{k+1}^2 - \epsilon_{k}^2 \right]. \]  

(C.9)

By the same reasoning, the alternation of the Lyapunov function \( \Delta V_{2k} \) can be expressed by
\[ \Delta V_{2k} = \frac{1}{2} \left[ \epsilon_{k+1}^2 - \epsilon_{k}^2 \right] = -\chi_3 \epsilon_{k}^2, \]  
where
\[ \chi_3 = \sum z_i \left[ \frac{1}{2} \eta_i \left( \frac{\partial \bar{y}_{k}}{\partial p_s} \right)^2 \left( 2 - \eta_i \left( \frac{\partial \bar{y}_{k}}{\partial p_s} \right)^2 \right) \right] \]  
\[ \rho = \left[ \bar{W}_{m}, \bar{a}_{m}, \bar{a}_{i}, \bar{b}_j, \bar{W}_{m}, \bar{a}_{i}, \bar{a}_{m} \right]^T. \]  

(C.11)

D: Adjustment Equations of OEAHW Model

The adjustment equation of polynomial function based OEAHW model using IBPG algorithm is as follows:
\[ \tilde{x}_{nt} = \tilde{x}_{nt-1} + \eta_{nt} \epsilon_{nt} \frac{\partial \tilde{x}_{nt}}{\partial \tilde{x}_{nt-1}} \left( \frac{\partial \tilde{x}_{k_l}}{\partial \tilde{x}_{nt-1}} \right) \]  
\[ n = (1, 2, \ldots, N_1), \]  
where
\[ \Delta \epsilon_{k} = \epsilon_{k+1} - \epsilon_{k}. \]  

(C.4)
\[ \frac{\partial \tilde{Y}_{k,t}}{\partial \tilde{X}_{k,t}} = \sum_{i=1}^{N_1} (-s \tilde{p}_{j,t-1} \tilde{X}_{k,t}), \]

\[ \frac{\partial \tilde{X}_{k,t}}{\partial \tilde{X}_{k,t-1}} = \sum_{j=1}^{n} \tilde{b}_{j,t-1} \mu_{k-1} - \sum_{i=1}^{n} \tilde{a}_{i,t-1} \frac{\partial \tilde{X}_{k,t}}{\partial \tilde{X}_{k,t-1}}, \]

\[ \tilde{a}_{i,t} = \tilde{a}_{i,t-1} + \eta \tilde{e}_{i,k}, i = (1, 2, \ldots, n_a), \]

\[ \tilde{b}_{j,t} = \tilde{b}_{j,t-1} + \eta \tilde{e}_{j,k}, j = (1, 2, \ldots, n_b), \]

\[ \frac{\partial \tilde{Y}_{k,t}}{\partial \tilde{a}_{i,t-1}} = \frac{\partial \tilde{X}_{k,t}}{\partial \tilde{X}_{k,t-1}}, \]

\[ \frac{\partial \tilde{X}_{k,t}}{\partial \tilde{b}_{j,t-1}} = \tilde{h}_{k,t} - \sum_{i=1}^{n} \tilde{a}_{i,t-1} \frac{\partial \tilde{X}_{k,t}}{\partial \tilde{X}_{k,t-1}}, \]

\[ \tilde{p}_{m,t} = \tilde{p}_{m,t-1} + \eta \tilde{e}_{m,k}, m = (1, 2, \ldots, N_2), \]

\[ \tilde{e}_{p,t} = \tilde{e}_{p,t-1} + \eta \tilde{e}_{p,k}, p = (1, 2, \ldots, n_c). \]

\[ (D.2) \]

**Data Availability**

No data were used to support this study.

**Conflicts of Interest**

The authors declare that they have no conflicts of interest.

**References**

[1] Z. Hafezi and M. M. Arefi, "Recursive generalized extended least squares and RML algorithms for identification of bilinear systems with ARMA noise," *ISA Transactions*, vol. 88, pp. 50–61, 2019.

[2] G. Fan, "Systematic parameter identification of a control-oriented electrochemical battery model and its application for state of charge estimation at various operating conditions," *Journal of Power Sources*, vol. 470, 2020.

[3] L. Li and X. Ren, "Identification of nonlinear Wiener-Hammerstein systems by a novel adaptive algorithm based on cost function framework," *ISA Transactions*, vol. 80, pp. 146–159, 2018.

[4] L. Wan, F. Ding, X. Liu, and C. Chen, "A new iterative least squares parameter estimation approach for equation-error autoregressive systems," *International Journal of Control, Automation and Systems*, vol. 18, no. 3, pp. 780–790, 2020.

[5] C.-R. Rad and O. Hancu, "An improved nonlinear modelling and identification methodology of a servo-pneumatic actuating system with complex internal design for high-accuracy motion control applications," *Simulation Modelling Practice and Theory*, vol. 75, pp. 29–47, 2017.

[6] C. Wang and J. Xun, "Novel recursive least squares identification for a class of nonlinear multiple-input single-output systems using the filtering technique," *Advances in Mechanical Engineering*, vol. 8, no. 11, Article ID 168781401668033, 2016.

[7] A. Martinčević, M. Vašak, and V. Lesić, "Identification of a control-oriented energy model for a system of fan coil units," *Control Engineering Practice*, vol. 91, Article ID 104100, 2019.

[8] J. Ding, "The hierarchical iterative identification algorithm for multi-input-output-error systems with autoregressive noise," *Complexity*, vol. 2017, Article ID 5929894, 11 pages, 2017.

[9] Q. Wu, X. Wang, B. Chen, and H. Wu, "Modeling, online identification, and compensation control of single tendon sheath system with time-varying configuration," *Mechanical Systems and Signal Processing*, vol. 130, pp. 56–73, 2019.

[10] M. A. Islas, J. d. J. Rubio, S. Muñiz et al., "A fuzzy logic model for hourly electrical power demand modeling," *Electronics*, vol. 10, p. 448, 2021.

[11] H.-S. Chang, M.-Y. Chen, and Y.-J. Huang, "Wavelet-based EEG processing for epilepsy detection using fuzzy entropy and associative petri net," *IEEE Access*, vol. 7, pp. 103255–103262, 2019.

[12] F. Furlán, E. Rubio, H. Sossa, and V. Ponce, "CNN based detectors on planetary environments: a performance evaluation," *Frontiers in Neurorobotics*, vol. 14, 2020.

[13] J. de Jesus Rubio, "SOFMLS: online self-organizing fuzzy modified least-squares network," *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 6, pp. 1296–1309, 2009.

[14] M. Liu, Y. Xiao, and R. Ding, "Iterative identification algorithm for Wiener nonlinear systems using the Newton method," *Applied Mathematical Modelling*, vol. 37, no. 9, pp. 6584–6591, 2013.

[15] J. Peng, R. Dubay, J. M. Hernandez, and M. M. Abu-Ayyad, "A Wiener neural network-based identification and adaptive generalized predictive control for nonlinear SISO systems," *Industrial & Engineering Chemistry Research*, vol. 50, no. 12, pp. 7388–7397, 2011.

[16] L. Zhou, X. Li, and F. Pan, "Gradient-based iterative identification for MISO Wiener nonlinear systems: application to a glutamate fermentation process," *Applied Mathematics Letters*, vol. 26, no. 8, pp. 886–892, 2013.

[17] Y. Xia, J. Zou, W. Yan, and H. Li, "Adaptive tracking constrained controller design for solid oxide fuel cells based on a Wiener-type neural network," *Applied Sciences*, vol. 8, no. 10, 2018.

[18] C. Qi, H.-X. Li, X. Zhao, S. Li, and F. Gao, "Hammerstein modeling with structure identification for multi-input multi-output nonlinear industrial processes," *Industrial & Engineering Chemistry Research*, vol. 50, no. 19, pp. 11153–11169, 2011.

[19] D. Wang, S. Zhang, M. Gan, and J. Qiu, "A novel EM identification method for hammerstein systems with missing output data," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2500–2508, 2020.

[20] B. Zhang, Z. Mao, and T. Zhang, "Intelligent control for Hammerstein nonlinear systems with arbitrary deadzone input," *Transactions of the Institute of Measurement and Control*, vol. 39, no. 4, pp. 567–578, 2017.

[21] J. Wang, A. Sano, T. Chen, and B. Huang, "Identification of Hammerstein systems without explicit parameterisation of non-linearity," *International Journal of Control*, vol. 82, no. 5, pp. 937–952, 2009.

[22] L. Jia, M.-S. Chiu, and S. S. Ge, "A noniterative neuro-fuzzy based identification method for Hammerstein processes," *Journal of Process Control*, vol. 15, no. 7, pp. 749–761, 2005.
[23] J. Li and X. Li, “Particle swarm optimization iterative identification algorithm and gradient iterative identification algorithm for wiener systems with colored noise,” Complexity, vol. 2018, Article ID 7335171, 8 pages, 2018.

[24] X. Shao, “Identification for polynomial wiener model using improved back propagation method,” Physics Procedia, vol. 33, pp. 718–725, 2012.

[25] L. Zhou, X. Li, and F. Pan, “Least-squares-based iterative identification algorithm for wiener nonlinear systems,” Journal of Applied Mathematics, vol. 2013, Article ID 565641, 6 pages, 2013.

[26] L. Zhou, X. Li, and F. Pan, “Gradient based iterative parameter identification for Wiener nonlinear systems,” Applied Mathematical Modelling, vol. 37, no. 16-17, pp. 8203–8209, 2013.

[27] W. Xiong, W. Fan, and R. Ding, “Least-squares parameter estimation algorithm for a class of input nonlinear systems,” Journal of Applied Mathematics, vol. 2012, Article ID 684074, 14 pages, 2012.

[28] J. Ma and F. Ding, “Filtering-based multistage recursive identification algorithm for an input nonlinear output-error autoregressive system by using the key term separation technique,” Circuits, Systems, and Signal Processing, vol. 36, no. 2, pp. 577–599, 2017.

[29] F. Li and L. Jia, “Correlation analysis-based error compensation recursive least-square identification method for the Hammerstein model,” Journal of Statistical Computation and Simulation, vol. 88, no. 1, pp. 56–74, 2018.

[30] W. Jian-hong, “Recursive identification of Hammerstein systems with polynomial function approximation,” International Journal of Management and Fuzzy Systems, vol. 3, p. 87, 2017.

[31] H. Yu, J. Peng, and Y. Tang, “Identification of nonlinear dynamic systems using Hammerstein-type neural network,” Mathematical Problems in Engineering, vol. 2014, Article ID 959507, 9 pages, 2014.

[32] N. I. Chaudhary, M. A. Raja, J. A. Khan, and M. S. Aslam, “Identification of input nonlinear control autoregressive systems using fractional signal processing approach,” The Scientific World Journal, vol. 2013, Article ID 467276, 13 pages, 2013.

[33] H. Hu, X. Yongsong, and R. Ding, “Multi-innovation stochastic gradient identification algorithm for Hammerstein controlled autoregressive autoregressive systems based on the key term separation principle and on the model decomposition,” Journal of Applied Mathematics, vol. 2013, Article ID 596141, 7 pages, 2013.

[34] S. Dong, L. Yu, W.-A. Zhang, and X. Yang, “Recursive identification of Hammerstein systems with dead zone input nonlinearity,” in Proceedings of the 2019 Chinese Control And Decision Conference (CCDC), pp. 4101–4106, Nanchang, China, June 2019.

[35] L. Fang, J. Wang, and Q. Zhang, “Identification of extended Hammerstein systems with hysteresis-type input nonlinearities described by Preisach model,” Nonlinear Dynamics, vol. 79, no. 2, pp. 1257–1273, 2015.

[36] J. Ma, W. Xiong, and F. Ding, “Iterative identification algorithms for input nonlinear output error autoregressive systems,” International Journal of Control, Automation and Systems, vol. 14, no. 1, pp. 140–147, 2016.

[37] Y. J. Wang, J. L. Ho, and Y. Bin Jiang, “A self-positioning linear actuator based on a piezoelectric slab with multiple pads,” Mechanical Systems and Signal Processing, vol. 150, 2021.

[38] W. Chen, R. Zhang, H. Liu, X. Xie, and L. Yan, “A novel method for solar panel temperature determination based on a wavelet neural network and Hammerstein-Wiener model,” Advances in Space Research, vol. 66, no. 8, pp. 2035–2046, 2020.

[39] H. C. Park, S. W. Sung, and J. Lee, “Modeling of Hammerstein–Wiener processes with special input test signals,” Industrial & Engineering Chemistry Research, vol. 45, no. 3, pp. 1029–1038, 2006.

[40] Y. Yan, B. Li, H. F. Ling, H. S. Chen, and M. J. Zhang, “Nonlinear state space modeling and system identification for electrohydraulic control,” Mathematical Problems in Engineering, vol. 2013, Article ID 973903, 9 pages, 2013.

[41] A. Nemati and M. Faeghi, “The performance comparison of ANFIS and Hammerstein-Wiener models for BLDC motors,” in Lecture Notes in Electrical Engineering, pp. 29–37, Springer, Berlin, Germany, 2011.

[42] J. Du, L. Zhang, J. Chen, J. Li, and C. Zhu, “Multi-model predictive control of Hammerstein-Wiener systems based on balanced multi-model partition,” Mathematical and Computer Modelling of Dynamical Systems, vol. 25, no. 4, pp. 333–353, 2019.

[43] R. Abbasi-Asl, R. Khorsand, and B. Vosoughi-Vahdat, “Hammerstein-Wiener model: a new approach to the estimation of formal neural information,” Basic and Clinical Neuroscience, vol. 3, pp. 45–51, 2012.

[44] N. Patcharaprakiti, K. Kirtikara, V. Monyakul et al., “Modeling of single phase inverter of photovoltaic system using Hammerstein-Wiener nonlinear system identification,” Current Applied Physics, vol. 10, no. 3, pp. S532–S536, 2010.

[45] C. Napole, O. Barambones, I. Calvo, M. Derbeli, M. Y. Silaa, and J. Velasco, “Advances in tracking control for piezoelectric actuators using fuzzy logic and Hammerstein-Wiener compensation,” Mathematics, vol. 8, no. 11, 2020.

[46] K.-K. Xu, H.-D. Yang, and C.-J. Zhu, “A novel extreme learning Machine-based Hammerstein-Wiener model for complex nonlinear industrial processes,” Neurocomputing, vol. 358, pp. 246–254, 2019.

[47] E. W. Bai, “An optimal two stage identification algorithm for Hammerstein-Wiener nonlinear systems,” in Proceedings of the 1998 American Control Conference, pp. 2756–2760, Philadelphia, PA, USA, June 1998.

[48] E.-W. Bai, “A blind approach to the Hammerstein-Wiener model identification,” Automatica, vol. 38, no. 6, pp. 967–979, 2002.

[49] J. Wang, T. Chen, and L. Wang, “A blind approach to identification of Hammerstein-Wiener systems corrupted by nonlinear-process noise,” in Proceedings of the 2009 7th Asian Control Conference, pp. 1340–1345, Hong Kong, China, August 2009.

[50] J. Vörös, “Identification of nonlinear block-oriented systems with backlash and saturation,” Journal of Electrical Engineering, vol. 70, no. 2, pp. 138–144, 2019.

[51] P. Crama and J. Schoukens, “Hammerstein-Wiener system estimator initialization,” Automatica, vol. 40, no. 9, pp. 1543–1550, 2004.

[52] A. Brouiri, F. Giri, F. Ikhouane, and F. Z. Chaoui, “Identification of Hammerstein-Wiener systems with backlash input nonlinearity bordered by straight lines,” IFAC Proceedings Volumes, vol. 47, no. 3, pp. 475–480, 2014.

[53] H. Hong and Z. Mao, “An identification algorithm for Hammerstein-Wiener system with dead zone input nonlinearity using gradient method,” in Proceedings of the 26th
F. Yu, Z. Mao, M. Jia, and P. Yuan, "Recursive parameter estimation for Hammerstein-Wiener nonlinear blocks in presence of noise and disturbance," *International Journal of Information and Electronics Engineering*, vol. 11, pp. 9–15, 2017.

A. Wills, T. B. Schön, L. Ljung, and B. Ninness, "Identification of hammerstein-wiener models," *Automatica*, vol. 49, no. 1, pp. 70–81, 2013.

L. Esmailian, J. Ghaisari, and M. A. Bagherzadeh, "Bayesian approach to identify Hammerstein-Wiener non-linear model in presence of noise and disturbance," *IET Control Theory & Applications*, vol. 13, no. 3, pp. 367–376, 2019.

J. S. Wang and Y. C. Chen, "A Hammerstein-Wiener recurrent neural network with universal approximation capability," in *Proceedings of the 2008 IEEE International Conference on Systems, Man and Cybernetics*, pp. 1832–1837, Singapore, October 2008.

Y. C. Chen and J. S. Wang, "A Hammerstein-Wiener recurrent neural network with frequency-domain eigensystem realization algorithm for unknown system identification," *Journal of Universal Computer Science*, vol. 15, pp. 2565–2583, 2009.

L. Jia, A. Yang, and M. Chiu, "Special two-stage input signal based neuro-fuzzy model for Hammerstein-Wiener processes," *International Journal of System Control and Information Processing*, vol. 1, p. 199, 2012.

F. Li, L. Chen, S. Wo, S. Li, and Q. Cao, "Modeling and parameter learning method for the Hammerstein-Wiener model with disturbance," *Measurement and Control*, vol. 53, no. 5-6, pp. 971–982, 2020.

J. Li and F. Qiliang, "Combined separable signals based neuro-fuzzy Hammerstein-Wiener model," *Memetic Computing*, vol. 9, no. 3, pp. 245–259, 2017.

S. E. Abouda, D. B. H. Abid, M. Elloumi, Y. Koubaa, and A. Chaari, "Identification of non-linear stochastic systems using a new Hammerstein-Wiener neural network: a simulation study through a non-linear hydraulic process," *International Journal of Computer Applications in Technology*, vol. 63, no. 3, pp. 241–256, 2020.

F. Yu, Z. Mao, and M. Jia, "Recursive identification for Hammerstein-Wiener systems with dead-zone input non-linearity," *Journal of Process Control*, vol. 23, no. 8, pp. 1108–1115, 2013.

M. Biegański, "Multi-level identification of Hammerstein-Wiener (N-L-N) system in active experiment," in *Proceedings of the 15th International Conference on Informatics in Control, Automation and Robotics*, pp. 355–361, Porto, Portugal, 2018.

B. Zhang and Z. Mao, "Adaptive control of stochastic Hammerstein-Wiener nonlinear systems with measurement noise," *International Journal of Systems Science*, vol. 47, no. 1, pp. 162–178, 2016.

F. Yu, Z. Mao, and D. He, "Identification of time-varying Hammerstein-Wiener systems," *IEEE Access*, vol. 8, pp. 136906–136916, 2020.

F. Yu, Z. Mao, P. Yuan, D. He, and M. Jia, "Recursive parameter estimation for Hammerstein-Wiener systems using modified EKF algorithm," *ISA Transactions*, vol. 70, pp. 104–115, 2017.

F. Yu, Z. Mao, M. Jia, and P. Yuan, "Recursive parameter identification of Hammerstein-Wiener systems with measurement noise," *Signal Processing*, vol. 105, pp. 137–147, 2014.

D. Wang and F. Ding, "Extended stochastic gradient identification algorithms for Hammerstein-Wiener ARMAX systems," *Computers & Mathematics with Applications*, vol. 56, no. 12, pp. 3157–3164, 2008.

Y. Wang and F. Ding, "Parameter estimation algorithms for hammerstein-wiener systems with autoregressive moving average noise," *Journal of Computational and Nonlinear Dynamics*, vol. 11, pp. 1–8, 2016.

Y. Li, Z. Z. Mao, and Y. Wang, "BCRLS identification method for Hammerstein-Wiener model," in *Proceedings of the 2010 International Conference on Measuring Technology and Mechatronics Automation*, no. 1, pp. 745–748, Changsha, China, March 2010.

J. Bai, Z. Mao, and T. Pu, "Recursive identification for multi-input-multi-output Hammerstein-Wiener system," *International Journal of Control*, vol. 92, no. 6, pp. 1457–1469, 2019.

S. E. Abouda, M. Elloumi, Y. Koubaa, and A. Chaari, "Over parameterisation and optimisation approaches for identification of nonlinear stochastic systems described by Hammerstein-Wiener models," *International Journal of Modelling, Identification and Control*, vol. 33, no. 1, pp. 61–75, 2019.

S. E. Abouda, M. Elloumi, Y. Koubaa, and A. Chaari, "Design of parametric estimation algorithm for Hammerstein-Wiener mathematical models," in *Proceedings of the 2019 19th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, pp. 371–375, Sousse, Tunisia, March 2019.

A. Sadollah, "Introductory chapter: which membership function is appropriate in fuzzy system?" in *Fuzzy Logic Based in Optimization Methods and Control Systems and Its Applications*, p. 13, IntechOpen, London, UK, 2018.

J. A. Hameed, "Using Gaussian membership functions for improving the reliability and robustness of students’ evaluation systems," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7135–7142, 2011.

V. O. S. Olunloyo, A. M. Ajofoyinbo, and O. Ibibado-Obe, "On development of fuzzy controller: the case of Gaussian and triangular membership functions," *Journal of Signal and Information Processing*, vol. 2, no. 4, pp. 257–265, 2011.

J. Vörös, "An iterative method for Wiener-Hammerstein systems parameter identification," *Journal of Electrical Engineering*, vol. 55, pp. 328–331, 2004.

K. D. Yamada, "Hyperparameter-free optimizer of stochastic gradient descent that incorporates unit correction and momentum estimation," *BioRxiv*, pp. 1–4, 2018.

D. B. H. Abid and M. Chiourou, "Indirect adaptive fuzzy control of non-linear systems using fuzzy supervisory term," *International Journal of Computer Applications in Technology*, vol. 59, no. 2, p. 130, 2019.

S. E. Abouda, D. B. H. Abid, M. Elloumi, Y. Koubaa, and A. Chaari, "Identification of nonlinear dynamic systems using fuzzy Hammerstein-Wiener systems," in *Proceedings of the 2019 19th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, pp. 365–370, Sousse, Tunisia, March 2019.

J. Park, D. Yi, and S. Ji, "A novel learning rate schedule in optimization for neural networks and it’s convergence,", *Symmetry (Basel)*, vol. 12, pp. 1–16, 2020.

R. Q. Dong, Y. Zhang, and A. G. Wu, "Weighted hierarchical stochastic gradient identification algorithms for ARX..."
models,” *International Journal of Systems Science*, vol. 52, no. 2, 2020.

[84] Q. Tu, Y. Rong, and J. Chen, “Parameter identification of ARX models based on modified momentum gradient descent algorithm,” *Complexity*, vol. 2020, Article ID 9537075, 11 pages, 2020.

[85] F. Ding, L. Wan, Y. Guo, and F. Chen, “The filtering based auxiliary model generalized extended stochastic gradient identification for a multivariable output-error system with autoregressive moving average noise using the multi-innovation theory,” *Journal of the Franklin Institute*, vol. 357, no. 9, pp. 5591–5609, 2020.

[86] X. Wang and F. Ding, “Partially coupled extended stochastic gradient algorithm for nonlinear multivariable output error moving average systems,” *Engineering Computations*, vol. 34, no. 2, pp. 629–647, 2017.

[87] J. Ding, Z. Cao, J. Chen, and G. Jiang, “Weighted parameter estimation for Hammerstein nonlinear ARX systems,” *Circuits, Systems, and Signal Processing*, vol. 39, no. 4, pp. 2178–2192, 2020.

[88] L. Sha and Y. Ji, “Gradient-based iterative parameter estimation for output error autoregressive systems using hierarchical principle,” in *Proceedings of the 2020 Chinese Control and Decision Conference (CCDC)*, pp. 1684–1689, Hefei, China, August 2020.

[89] F. Ding and T. Chen, “Performance analysis of multi-innovation gradient type identification methods,” *Automatica*, vol. 43, no. 1, pp. 1–14, 2007.

[90] M. Chen, F. Ding, and E. Yang, “Gradient-based iterative parameter estimation for bilinear-in-parameter systems using the model decomposition technique,” *IET Control Theory & Applications*, vol. 12, no. 17, pp. 2380–2389, 2018.

[91] Q. Liu, F. Ding, Q. Zhu, and T. Hayat, “Two-stage multi-innovation stochastic gradient algorithm for multivariable output-error ARMA systems based on the auxiliary model,” *International Journal of Systems Science*, vol. 50, no. 15, pp. 2870–2884, 2019.

[92] P. Ma and F. Ding, “New gradient based identification methods for multi-variable pseudo-linear systems using the multi-innovation and the data filtering,” *Journal of the Franklin Institute*, vol. 354, no. 3, pp. 1568–1583, 2017.

[93] S. Cheng, Y. Wei, D. Sheng, Y. Chen, and Y. Wang, “Identification for Hammerstein nonlinear ARMAX systems based on multi-innovation fractional order stochastic gradient,” *Signal Processing*, vol. 142, pp. 1–10, 2018.

[94] F. Li, J. Li, and D. Peng, “Identification method of neuro-fuzzy-based Hammerstein model with coloured noise,” *IET Control Theory & Applications*, vol. 11, no. 17, pp. 3026–3037, 2017.

[95] L. Xu, F. Ding, and E. Yang, “Auxiliary model multi-innovation stochastic gradient parameter estimation methods for nonlinear sandwich systems,” *International Journal of Robust and Nonlinear Control*, vol. 31, no. 1, pp. 148–165, 2021.

[96] H. Xu, F. Ding, and E. Yang, “Three-stage multi-innovation parameter estimation for an exponential autoregressive time-series model with moving average noise by using the data filtering technique,” *International Journal of Robust and Nonlinear Control*, vol. 31, no. 1, pp. 166–184, 2021.

[97] Y. Wang and F. Ding, “A filtering based multi-innovation gradient estimation algorithm and performance analysis for nonlinear dynamical systems,” *IMA Journal of Applied Mathematics*, vol. 82, no. 6, pp. 1171–1191, 2017.

[98] B. Shen, F. Ding, L. Xu, and T. Hayat, “Data filtering based multi-innovation gradient identification methods for feedback nonlinear systems,” *International Journal of Control, Automation and Systems*, vol. 16, no. 5, pp. 2225–2234, 2018.

[99] Y. Mao and F. Ding, “Data filtering-based multi-innovation stochastic gradient algorithm for nonlinear output error autoregressive systems,” *Circuits, Systems, and Signal Processing*, vol. 35, no. 2, pp. 651–667, 2016.

[100] Q. Liu and F. Ding, “The data filtering based generalized stochastic parameter estimation algorithms for multivariate output-error autoregressive systems using the auxiliary model,” *Multidimensional Systems and Signal Processing*, vol. 29, no. 4, pp. 1781–1800, 2018.

[101] J. Pan, H. Ma, X. Jiang, W. Ding, and F. Ding, “Adaptive gradient-based iterative algorithm for multivariable controlled autoregressive moving average systems using the data filtering technique,” *Complexity*, vol. 2018, Article ID 9589307, 11 pages, 2018.

[102] Z. Wang, Y. Wang, and Z. Ji, “Filtering based multi-innovation extended stochastic gradient algorithm for Hammerstein nonlinear system modeling,” *Applied Mathematical Modelling*, vol. 39, no. 18, pp. 5724–5732, 2015.

[103] Y. Wang and F. Ding, “Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model,” *Automatica*, vol. 71, pp. 308–313, 2016.

[104] Y. Mao and F. Ding, “A novel data filtering based multi-innovation stochastic gradient algorithm for Hammerstein nonlinear systems,” *Digital Signal Processing*, vol. 46, pp. 215–225, 2015.

[105] Z. Yue and X. Yang, “Multi-innovation identification algorithm based on auxiliary model for Hammerstein output error systems,” in *Proceedings of the 2020 Chinese Automation Congress (CAC)*, pp. 5635–5640, IEEE, Shanghai, China, November 2020.

[106] L. Li, H. Zhang, and X. Ren, “A modified multi-innovation algorithm to turntable servo system identification,” *Circuits, Systems, and Signal Processing*, vol. 39, no. 9, pp. 4339–4353, 2020.

[107] X. Wang and F. Ding, “Modelling and multi-innovation parameter identification for Hammerstein nonlinear state space systems using the filtering technique,” *Mathematical and Computer Modelling of Dynamical Systems*, vol. 22, no. 2, pp. 113–140, 2016.

[108] Z. Hua, J. Xin, C. Hang, and X. Shou, “The parameter identification of the autonomous underwater vehicle based on multi-innovation least squares identification algorithm,” *International Journal of Advanced Robotic Systems*, vol. 17, pp. 1–11, 2020.

[109] F. Ding, P. X. Liu, and G. Liu, “Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises,” *Signal Processing*, vol. 89, no. 10, pp. 1883–1890, 2009.

[110] P. Ma, F. Ding, and Q. Zhu, “Decomposition-based recursive least squares identification methods for multivariate pseudo-linear systems using the multi-innovation,” *International Journal of Systems Science*, vol. 49, no. 5, pp. 920–928, 2018.

[111] F. Ding, “Several multi-innovation identification methods,” *Digital Signal Processing*, vol. 20, no. 4, pp. 1027–1039, 2010.

[112] L. Han and F. Ding, “Identification for multivariate multi-input systems using the multi-innovation identification theory,” *Computers & Mathematics with Applications*, vol. 57, no. 9, pp. 1438–1449, 2009.
[113] F. Ding and T. Chen, “Multi-innovation stochastic gradient
identification methods,” in Proceedings of the 2006 6th World
Congress on Intelligent Control and Automation, pp. 1501–
1505, IEEE, Dalian, China, June 2006.

[114] H. Chen, W. Zhang, and F. Ding, “Data filtering based least
squares iterative algorithm for parameter identification of
output error autoregressive systems,” Information Processing
Letters, vol. 114, no. 10, pp. 573–578, 2014.

[115] X. Lu, W. Zhou, and W. Shi, “Data filtering based recursive
least squares algorithm for two-input single-output systems
with moving average noises,” Journal of Applied Mathe-
matics, vol. 2014, Article ID 694053, 8 pages, 2014.

[116] Z. Wang, Y. Shen, Z. Ji, and R. Ding, “Filtering based re-
cursive least squares algorithm for Hammerstein FIR-MA
systems,” Nonlinear Dynamics, vol. 73, no. 1-2, pp. 1045–1054, 2013.

[117] Y. Wang, L. Xu, and F. Ding, “Data filtering based parameter
estimation algorithms for multivariable box-Jenkins-like
systems,” IFAC-Papersonline, vol. 48, no. 8, pp. 849–852,
2015.

[118] T. Cui, F. Ding, A. Alsaedi, and T. Hayat, “Data filtering-
based parameter and state estimation algorithms for state-
space systems disturbed by coloured noises,” International
Journal of Systems Science, vol. 51, no. 9, pp. 1669–1684, 2020.

[119] Z. Wang and Z. Ji, “Data-filtering-based iterative identifi-
cation methods for nonlinear FIR-MA systems,” Journal of
Vibration and Control, vol. 20, no. 14, pp. 2193–2201, 2014.