Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Separation identification of a neural fuzzy Wiener–Hammerstein system using hybrid signals*

Feng LI[†‡1], Hao YANG[1], Qingfeng CAO[2]

[1]College of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou 213001, China

[2]College of Electrical, Energy and Power Engineering, Yangzhou University, Yangzhou 225127, China

[†]E-mail: lifeng@jsut.edu.cn

Received Jan. 31, 2023; Revision accepted July 21, 2023; Crosschecked May 25, 2024

Abstract: A novel separation identification strategy for the neural fuzzy Wiener–Hammerstein system using hybrid signals is developed in this study. The Wiener–Hammerstein system is described by a model consisting of two linear dynamic elements with a nonlinear static element in between. The static nonlinear element is modeled by a neural fuzzy network (NFN) and the two linear dynamic elements are modeled by an autoregressive exogenous (ARX) model and an autoregressive (AR) model, separately. When the system input is Gaussian signals, the correlation technique is used to decouple the identification of the two linear dynamic elements from the nonlinear element. First, based on the input and output of Gaussian signals, the correlation analysis technique is used to identify the input linear element and output linear element, which addresses the problem that the intermediate variable information cannot be measured in the identified Wiener–Hammerstein system. Then, a zero-pole match method is adopted to separate the parameters of the two linear elements. Furthermore, the recursive least-squares technique is used to identify the nonlinear element based on the input and output of random signals, which avoids the impact of output noise. The feasibility of the presented identification technique is demonstrated by an illustrative simulation example and a practical nonlinear process. Simulation results show that the proposed strategy can obtain higher identification precision than existing identification algorithms.

Key words: Wiener–Hammerstein system; Neural fuzzy network; Correlation analysis technique; Hybrid signals; Separation identification

https://doi.org/10.1631/FITEE.2300058 **CLC number:** TP273; TP18

1 Introduction

Block-oriented systems have been recognized as particularly useful for identification modeling of practical nonlinear processes owing to their simple structure and excellent approximation capability (Mu and Chen, 2014). The systems most representative of this particular type are the Hammerstein and the Wiener systems, which are composed of a static nonlinear element and a linear dynamic element (Li F et al., 2022, 2023a, 2023b; Wang et al., 2022).

The Wiener–Hammerstein system (Ghanmi et al., 2020) is a type of block structure system consisting of a static nonlinear element and two linear dynamic elements, inheriting properties from both Wiener and Hammerstein systems (Fig. 1). Thus, it can more effectively approximate a nonlinear system than either of the two separate systems. Such a block structure system may be used to represent many practical nonlinear applications, such as an electronic circuit (Piroddi et al., 2012), heat exchanger

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (No. 62003151), the Changzhou Science and Technology Bureau (Nos. CJ20220065 and CM20223015), the Qinglan Project of Jiangsu Province, China (No. 2022[29]), and the Zhongwu Youth Innovative Talents Support Program of Jiangsu University of Technology, China (No. 202102003)

ORCID: Feng LI, https://orcid.org/0000-0001-9445-1627

© Zhejiang University Press 2024

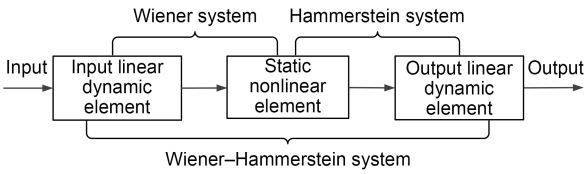

Fig. 1 A diagram of the Wiener–Hammerstein system

(Ławryńczuk, 2016), wing flutter clearance of an F-18 system (Janjanam et al., 2022), and a continuous stirred tank reactor (Škrjanc, 2021). Much effort has been made in identifying the nonlinear Wiener–Hammerstein systems. In general, there are two kinds of identification schemes: synchronous identification and decoupled identification. The synchronous identification approaches identify hybrid parameters composed of nonlinear and linear elements, and include subspace methods (Ase and Katayama, 2015; Katayama and Ase, 2016), multi-innovation-based methods (Li LW et al., 2020), and over-parameter methods (Falck et al., 2012). These methods contain a parameter product term and need to be further separated by adopting a parameter separation technique, such as a singular value decomposition method or an average method. Therefore, the identification accuracy of the synchronous identification approaches is reduced and the computational complexity is increased. In contrast, a decoupled identification scheme can separate parameter identification of the static nonlinear and linear dynamic elements by estimating the immeasurable intermediate variable. Thus, nonlinear element and linear element parameters are carried out independently, such as in iterative methods (dos Santos et al., 2012; Han and de Callafon, 2012) and frequency-domain methods (Rijlaarsdam et al., 2012; Shaikh and Barbé, 2019). In addition, some intelligent optimization techniques are used to identify the Wiener–Hammerstein system, such as the genetic method (Naitali and Giri, 2016) and particle swarm optimization method (Zong et al., 2021).

In recent years, identification modeling methods of Wiener–Hammerstein systems have been widely studied. Substantial research work has focused on the identification modeling issue of the Wiener–Hammerstein nonlinear systems with modeling of static nonlinear functions by polynomial expression, piecewise linear function, and basis function expansion (Paduart et al., 2012; Sjöberg and Schoukens, 2012; Tiels et al., 2014). Li LW and Ren (2018) investigated parameter identification of Wiener–Hammerstein systems by using a filter gain and novel cost function, in which smooth nonlinearity modeled by polynomials and non-smooth nonlinearity approximated by dead-zone were considered. Hafsi et al. (2012) applied the multi-segment piecewise-linear characteristic to represent the nonlinearity of the Wiener–Hammerstein system and the recursive least-squares method to fulfill simultaneous estimation of both linear blocks and nonlinear block parameters. Based on the best linear approximation theory, Sjöberg et al. (2012) used a maximum likelihood estimate and iterative search method to describe the identification of a Wiener–Hammerstein system with basis function expansion and obtained good initial estimates. Note that the nonlinearities used in the above scheme can be considered as a linear combination of basis functions, which are not suitable for approximating complex discontinuous functions because the order of the nonlinear model may need to be large to obtain accurate modeling.

Neural networks and fuzzy systems have been applied widely to nonlinear dynamic system modeling owing to their ability to model nonlinearity to high precision. However, very few algorithms for identification of the Wiener–Hammerstein system based on neural networks or fuzzy systems have been proposed and formally analyzed. Weber and Gühmann (2021) used a non-autoregressive neural network architecture to represent the Wiener–Hammerstein system, and considered an optimized caching algorithm for identification. Škrjanc (2021) combined a filtered recursive least-squares method with incrementally evolving Gaussian clustering to identify an interval fuzzy model in an online manner for the Wiener–Hammerstein structure. Note that neural networks and fuzzy systems have their own advantages, namely their self-studying capability and fuzzy reasoning capability. In this context, neural fuzzy systems consider both the self-studying ability and the reasoning ability, which can more effectively model non-smooth or discontinuous nonlinear functions than the two models. Therefore, the main goal is to use a neural fuzzy system to represent the nonlinearity of the Wiener–Hammerstein system. In addition, the redundancy

of the parameters and the existence of the internal variables, which are inaccessible to measurement, are the major difficulties for the identification of Wiener–Hammerstein systems.

In this research, we focus on Wiener–Hammerstein system identification. The neural fuzzy network (NFN) is used to model only the static nonlinear element of the Wiener–Hammerstein system. Note that the identification of an NFN-based Wiener–Hammerstein system is fundamentally different from that of a single NFN. The main difference lies in the system structure constraint imposed by the Wiener–Hammerstein system, in which a static nonlinear element is embedded between the linear elements. It is apparent that the resulting identification problem needs to consider both nonlinear mapping and linear dynamics. In contrast, single NFN research results focus on the identification of one global nonlinear mapping.

In this study, we develop a new two-step separation identification approach for identifying a Wiener–Hammerstein system with output noise using hybrid signals. Hybrid signals, consisting of Gaussian signals and random signals, are applied to separate the identification of the two linear dynamic elements from the nonlinear element, which avoids decomposing parameter product terms in synchronous identification and improves identification accuracy. In the first step, in the case of input and output of Gaussian signals, the correlation analysis technique is used to identify the input linear element and output linear element. Thus, the issue that the intermediate variables of the Wiener–Hammerstein system cannot be measured is effectively handled. In the second step, the recursive least-squares technique is used to identify the nonlinear element based on the input and output of a random signal, which avoids the impact of output noise. The significant innovations of this work lie in the following aspects:

1. The parameter product term of the nonlinear element and the linear element is included for synchronous identification methods (Ase and Katayama, 2015; Katayama and Ase, 2016; Li LW et al., 2020). Unlike synchronous identification, the presented method separates parameter identification of the nonlinear element from that of the linear elements using the designed hybrid signals. This avoids decomposing parameter product terms in synchronous identification and improves identification accuracy.

2. To address the issue that the intermediate variables of the Wiener–Hammerstein system cannot be measured, unlike the auxiliary model technique used in Zong et al. (2021), a correlation analysis technique is used. The technique involves the autocorrelation function of input and cross-correlation function of input and output, which ensures that the identification procedure is realized.

3. The zero-pole match technique is adopted to acquire the parameters of the two linear elements without parameter redundancy.

2 Problem statement and preliminaries

The nonlinear Wiener–Hammerstein system considered consists of a model with an input linear dynamic element $G(q)$ in cascade with a static nonlinear element $f(\cdot)$ with an output linear dynamic element $H(q)$, as described in Mzyk and Wachel (2017) and Li LW et al. (2018). In this proposal, the autoregressive exogenous (ARX) model and autoregressive (AR) model are used to model the input linear dynamic element and output linear dynamic element, and NFN is represented to approximate the static nonlinear element. Fig. 2 shows the Wiener–Hammerstein system with ARX-NFN-AR interfered by output noise.

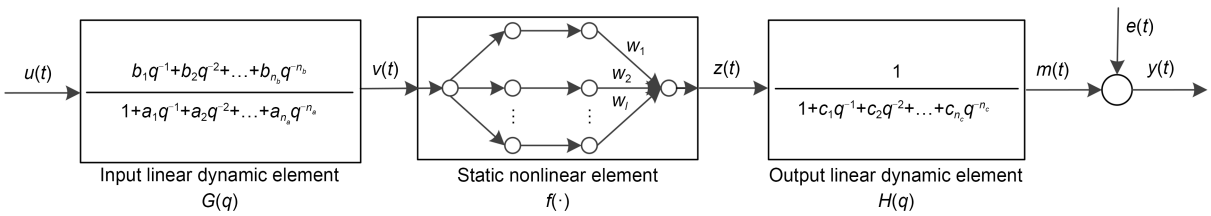

Fig. 2 Wiener–Hammerstein system with ARX-NFN-AR interfered by the output noise

The mathematical representations of the Wiener–Hammerstein system interfered by output noise are described by

$$\begin{cases} v(t)=u(t)G(q)=u(t)\dfrac{B(q)}{A(q)}, \\ z(t)=f(v(t)), \\ m(t)=H(q)z(t)=\dfrac{1}{C(q)}z(t), \\ y(t)=m(t)+e(t), \end{cases} \tag{1}$$

where $v(t)$, $z(t)$, and $m(t)$ represent the intermediate unmeasurable variables, $A(q)=1+\sum_{i=1}^{n_a}a_i q^{-i}$, $B(q)=\sum_{j=1}^{n_b}b_j q^{-j}$, $C(q)=1+\sum_{m=1}^{n_c}c_m q^{-m}$, q^{-1} is the backshift operator, namely $y(t)q^{-1}=y(t-1)$, n_a, n_b, and n_c are given orders of the ARX model and the AR model, and $e(t)$ is white noise.

It is assumed that the static nonlinear element can be represented by a four-layer NFN (Li F et al., 2017, 2022, 2023a, 2023b), and its output can be expressed as

$$f\left(v(t)\right)=\sum_{l=1}^{L}\frac{\mu_l\left(v(t)\right)}{\sum_{l=1}^{L}\mu_l\left(v(t)\right)}w_l, \tag{2}$$

where $\mu_l\left(v(t)\right)=\exp\left(-\dfrac{\left(v(t)-c_l\right)^2}{\sigma_l^2}\right)$, c_l is center, σ_l is width, w_l is weight, and L indicates the number of fuzzy rules.

Our goal is to estimate the unknown parameters of the Wiener–Hammerstein system with ARX-NFN-AR, i.e., the ARX model, NFN, and the AR model. This requires that the following objective functions be optimized:

$$\begin{aligned} &J\left(\hat{G}(q),\hat{f}(\hat{v}(t)),\hat{H}(q)\right)=\frac{1}{2N}\sum_{t=1}^{N}\left[\hat{y}(t)-y(t)\right]^2\le\varepsilon \\ &\text{s.t.}\quad \hat{v}(t)=\hat{G}(q)u(t),\ \hat{z}(t)=\hat{f}(\hat{v}(t)), \\ &\qquad\ \hat{m}(t)=\hat{H}(q)\hat{z}(t),\ \hat{y}(t)=\hat{m}(t)+e(t), \end{aligned} \tag{3}$$

in which symbol "^" denotes estimate, N represents the length of the measured data, and ε is a given threshold.

Remark 1 ARX and AR models have been used widely to handle linear system problems. In this work, ARX and AR are used only to establish the two linear dynamic elements of the Wiener–Hammerstein system. However, we pay more attention to the problem of Wiener–Hammerstein system identification. Note that the identification of an ARX- or AR-based Wiener–Hammerstein system is fundamentally different from that of a single ARX or AR. The main difference is that the system structure constraint imposed by the Wiener–Hammerstein system is a static nonlinear element embedded between linear elements. For the resulting identification problem we need to consider both nonlinear mapping and linear dynamics. In contrast, single ARX or AR research results focus on linear system identification.

3 Outline of the Wiener–Hammerstein system with ARX-NFN-AR identification

The synchronous method for identification of the Wiener–Hammerstein system (Li LW et al., 2020) includes a product term of element parameters, and the parameters of two linear elements are separated under the condition that the first parameter of both linear elements is set to one. In contrast, in our study, a feasible separation estimation approach is developed for estimating independently the Wiener–Hammerstein system with ARX-NFN-AR based on hybrid signals. In our previous work (Jia et al., 2017), the input–output function relationship of the Wiener system in the presence of a zero-mean white Gaussian process was derived, by which separation identification was realized. Considering the advantages of this approach, here we apply it to the Wiener–Hammerstein system with ARX-NFN-AR identification.

Assume that the system input is a zero-mean white Gaussian process, i.e., $u(t)\sim N(0,\sigma_u^2)$, and consider that the system input is a zero initial state. Then $v(t)$ can be expressed as

$$\begin{cases} v(1)=0, \\ v(2)=b_1u(1), \\ v(3)=-a_1v(2)-a_2v(1)+b_1u(2)+b_2u(1) \\ \qquad\ =(-a_1b_1+b_2)u(1)+b_1u(2), \\ \qquad\ \vdots \\ v(t)=\sum_{k=1}^{t-1}h_t(k)u(t-k), \end{cases} \tag{4}$$

in which h_t is a coefficient.

Based on Eq. (4), the conditional expectation of $v(t)$ can be obtained:

$$\begin{aligned}
&E\left(v(t-\tau)\,|\,v(t)\right)\\
&=\sum_{k=1}^{t-\tau-1}h_{t-\tau}(k)E\left(u(t-\tau-k)\,\middle|\,\sum_{\xi=1}^{t-1}h_t(\xi)u(t-\xi)\right)\\
&=\sum_{k=1}^{t-\tau-1}h_{t-\tau}(k)E\left[u(t-\tau-k)\,\middle|\,\left(h_t(\tau+k)u(t-\tau-k)\right.\right.\\
&\quad\left.\left.+\sum_{\xi=1}^{t-1}h_t(\xi)u(t-\xi),\xi\neq\tau+k\right)\right]\\
&=\sum_{k=1}^{t-\tau-1}h_{t-\tau}(k)E\left(U_{t-\tau-k}|(X_{t-\tau-k}+Y_t)\right),
\end{aligned}\tag{5}$$

where

$$X_{t-\tau-k}=h_t(\tau+k)U_{t-\tau-k},$$

$$Y_t=\sum_{\xi=1}^{t-1}h_t(\xi)u(t-\xi),\ \xi\neq\tau+k,\ U_{t-\tau-k}=u(t-\tau-k).$$

From Eq. (5), sequences $U_{t-\tau-k}$ and $X_{t-\tau-k}$ are independent of Y_t, and $U_{t-\tau-k}=X_{t-\tau-k}/h_t(\tau+k)$. Define $V_t=v(t)=X_{t-\tau-k}+Y_t$. Since $U_{t-\tau-k}$ and V_t are related, their joint probability distribution of functions (Ross, 2014; Jia et al, 2017) can be calculated by

$$\begin{aligned}
&f_{(U_{t-\tau-k},V_t)}(u,v)\\
&=f_{(X_{t-\tau-k},Y_t)}\left(h_t(\tau+k)u,\,v-h_t(\tau+k)u\right)\\
&\quad\cdot\left|\mathrm{Jac}(X_{t-\tau-k},Y_t)\right|^{-1}\\
&=f_{X_{t-\tau-k}}\left(h_t(\tau+k)u\right)f_{Y_t}\left(v-h_t(\tau+k)u\right)\\
&\quad\cdot\left|\mathrm{Jac}(X_{t-\tau-k},Y_t)\right|^{-1}\\
&=h_t(\tau+k)\frac{1}{2\pi\sigma_u\sigma_Y h_t(\tau+k)}\\
&\quad\cdot\exp\left(-\frac{z^2}{2\sigma_u^2}-\frac{(v-h_t(\tau+k)u)^2}{2\sigma_Y^2}\right)\\
&=\frac{1}{2\pi\sigma_u\sigma_Y}\exp\left(-\frac{u^2\sigma_Y^2+(v-h_t(\tau+k)u)^2\sigma_u^2}{2\sigma_u^2\sigma_Y^2}\right),
\end{aligned}\tag{6}$$

where

$$\mathrm{Jac}(X_{t-\tau-k},Y_t)=\frac{\partial(U_{t-\tau-k},V_t)}{\partial(X_{t-\tau-k},Y_t)}=\begin{vmatrix}\dfrac{\partial U_{t-\tau-k}}{\partial X_{t-\tau-k}} & \dfrac{\partial U_{t-\tau-k}}{\partial Y_t}\\ \dfrac{\partial V_t}{\partial X_{t-\tau-k}} & \dfrac{\partial V_t}{\partial Y_t}\end{vmatrix}=\frac{1}{h_t(\tau+k)}.$$

Assume that constant c is the value of $v(t)$ at time t. The conditional probability density function of $U_{t-\tau-k}$ in the condition of $v(t)$ can be computed using Eq. (7):

$$\begin{aligned}
f_{U_{t-\tau-k}|V_t=C}(u|v=c)&=\frac{f_{(U_{t-\tau-k},V_t)}(u,c)}{f_{V_t}(c)}\\
&=\frac{\sigma_v}{\sqrt{2\pi}\,\sigma_u\sigma_Y}\exp\left(\frac{c^2}{2\sigma_v^2}-\frac{u^2\sigma_Y^2+(c-h_t(\tau+k)u)^2\sigma_u^2}{2\sigma_u^2\sigma_Y^2}\right)\\
&=\frac{\sigma_v}{\sqrt{2\pi}\,\sigma_u\sigma_Y}\exp\left(\frac{\sigma_u^2\sigma_Y^2c^2}{2\sigma_v^2\sigma_u^2\sigma_Y^2}\right)\exp\left(-\frac{(u^2\sigma_Y^2+(c-h_t(\tau+k)u)^2\sigma_u^2)\sigma_v^2}{2\sigma_v^2\sigma_u^2\sigma_Y^2}\right)\\
&=\frac{\sigma_v}{\sqrt{2\pi}\,\sigma_u\sigma_Y}\exp\left(\frac{\sigma_u^2(\sigma_Y^2-\sigma_v^2)c^2}{2\sigma_v^2\sigma_u^2\sigma_Y^2}\right)\\
&\quad\cdot\exp\left(-\frac{\sigma_v^2(\sigma_Y^2+\sigma_u^2h_t(\tau+k)^2)u^2-2\sigma_v^2\sigma_u^2h_t(\tau+k)cu}{2\sigma_v^2\sigma_u^2\sigma_Y^2}\right)\\
&=k_c\exp\left(-\frac{k_au^2-2k_bcu}{2\sigma_v^2\sigma_u^2\sigma_Y^2}\right)=k_c\exp\left(-\frac{k_a(u-k_bc/k_a)^2}{2\sigma_v^2\sigma_u^2\sigma_Y^2}\right)\exp\left(\frac{k_b^2c^2}{2k_a\sigma_v^2\sigma_u^2\sigma_Y^2}\right)\\
&=k_d\exp\left(-\frac{k_a(u-k_bc/k_a)^2}{2\sigma_v^2\sigma_u^2\sigma_Y^2h_t(\tau+k)^2}\right),
\end{aligned}\tag{7}$$

where

$$k_a=\sigma_v^2(\sigma_Y^2+\sigma_u^2 h_t(\tau+k)^2),\ k_b=\sigma_v^2\sigma_u^2 h_t(\tau+k),$$

$$k_c=\frac{\sigma_v}{\sqrt{2\pi}\,\sigma_u\sigma_Y}\exp\left(\frac{\sigma_u^2(\sigma_Y^2-\sigma_v^2)c^2}{2\sigma_v^2\sigma_u^2\sigma_Y^2}\right),$$

$$k_d=k_c\cdot\exp\left(\frac{k_b^2}{2k_a\sigma_v^2\sigma_u^2\sigma_Y^2}\right),$$

and σ_u^2, σ_v^2, and σ_Y^2 are variances.

Eq. (7) shows that the conditional probability density function $f_{U_{t-\tau-k}|V_t=C}(\cdot)$ follows a normal distribution with mean $k_b c/k_a$ and variance $\sigma_v^2\sigma_u^2\sigma_Y^2 h_t(\tau+k)^2/k_a$.

Therefore, we can obtain the conditional mathematical expectation $u(t-\tau-k)$ in the condition of $v(t)$:

$$E\left(u(t-\tau-k)\,\middle|\,v(t)=c\right)$$
$$=E\left(u(t-\tau-k)\,\middle|\,\sum_{k=1}^{t-1}h_t(k)u(t-k)=c\right)=\frac{k_b}{k_a}c. \quad (8)$$

Using Eq. (8) concludes the following conditional mathematical expectation $v(t-\tau)$:

$$E\left(v(t-\tau)\,\middle|\,v(t)\right)$$
$$=E\left(\sum_{k=1}^{t-\tau-1}h_{t-\tau}(k)u(t-\tau-k)\,\middle|\,\sum_{k=1}^{t-1}h_t(k)u(t-k)\right)$$
$$=\sum_{k=1}^{t-\tau-1}h_{t-\tau}(k)\frac{k_b}{k_a}c=b(\tau)c, \quad (9)$$

where $b(\tau)=\frac{k_b}{k_a}\sum_{k=1}^{t-\tau-1}h_{t-\tau}(k)$ is a constant independent of c.

As mentioned in Martin and Lennart (2005), the process $v(t)$ holds separability in Eq. (9).

Theorem 1 Assume that the input of the Wiener–Hammerstein system described by Eq. (1) is a zero-mean white Gaussian process. Then, there is a constant b_0 that has the following relationship:

$$R_{yu}(\tau)=b_0G(q)H(q)R_u(\tau)=b_0G_H(q)R_u(\tau), \quad (10)$$

where $R_{yu}(\tau)$ is the cross-correlation function of the input and output, $R_u(\tau)$ is the autocorrelation function of the input, $G_H(q)=G(q)H(q)$ indicates a product of two linear elements, and $b_0=\frac{E\left(z(t)v(t)\right)}{E\left(v(t)v(t)\right)}$.

Proof Since $v(t)$ satisfies the separability, the correlation function $R_{zv}(\tau)$ can be computed by

$$R_{zv}(\tau)=E\left(z(t)v(t-\tau)\right)$$
$$=E\left(z(t)E\left(v(t-\tau)|v(t)\right)\right)$$
$$=b(\tau)E\left(z(t)v(t)\right). \quad (11)$$

Similarly, the correlation function is given by

$$R_v(\tau)=E\left(v(t)v(t-\tau)\right)=b(\tau)E\left(v(t)v(t)\right). \quad (12)$$

According to Eqs. (11) and (12), we have $\frac{R_{zv}(\tau)}{R_v(\tau)}=\frac{E\left(z(t)v(t)\right)}{E\left(v(t)v(t)\right)}$. Let $b_0=\frac{E\left(z(t)v(t)\right)}{E\left(v(t)v(t)\right)}$. Thereby, we have

$$R_{zv}(\tau)=b_0R_v(\tau). \quad (13)$$

The input linear dynamic element $G(q)$ is expressed using infinite impulse response, namely

$$v(t)=G(q)u(t)=\sum_{k=1}^{\infty}g(k)q^k u(t)=\sum_{k=1}^{\infty}g(k)u(t-k). \quad (14)$$

Based on Eq. (14), Eq. (11) can be written as

$$R_{zv}(\tau)=E\left(z(t)v(t-\tau)\right)$$
$$=E\left(z(t)\sum_{k=1}^{\infty}g(k)u(t-\tau-k)\right)$$
$$=\sum_{k=1}^{\infty}g(k)R_{zu}(\tau+k)$$
$$=\sum_{k=1}^{\infty}g(k)q^k R_{zu}(\tau). \quad (15)$$

Analogously, the following expression is obtained using Eq. (12):

$$R_v(\tau)=E\left(v(t)v(t-\tau)\right)$$
$$=E\left(\sum_{k=1}^{\infty}g(k)u(t-k)\sum_{\varsigma=1}^{\infty}g(\varsigma)u(t-\tau-\varsigma)\right)$$
$$=\sum_{k=1}^{\infty}\sum_{\varsigma=1}^{\infty}g(k)R_u(\tau+\varsigma-k)g(\varsigma)$$
$$=\sum_{\varsigma=1}^{\infty}g(\varsigma)q^{\xi}R_u(\tau)\sum_{k=1}^{\infty}g(k)q^{-k}. \quad (16)$$

Note that, for the input linear element, $\sum_{k=1}^{\infty} g(k)q^k$ in Eq. (15) and $\sum_{\varsigma=1}^{\infty} g(\varsigma)q^{\xi}$ in Eq. (16) are equal. Therefore, we have

$$R_{zu}(\tau)=b_0\sum_{k=1}^{\infty} g(k)q^{-k}R_u(\tau)=b_0G(q)R_u(\tau). \quad (17)$$

The output linear element $H(q)$ can also be expressed by an infinite impulse response using Eq. (14), and as shown in Eqs. (15) and (16), similar derivation processes are carried out. We can obtain

$$\begin{aligned}R_{mu}(\tau)&=E\big(m(t)u(t-\tau)\big)=E\big(H(q)z(t)u(t-\tau)\big)\\&=H(q)R_{zu}(\tau).\end{aligned} \quad (18)$$

Based on the equation $y(t)=m(t)+e(t)$, the following equation holds:

$$\begin{aligned}R_{yu}(\tau)&=E\big(m(t)u(t-\tau)\big)+E\big(e(t)u(t-\tau)\big)\\&=E\big(H(q)z(t)u(t-\tau)\big)=H(q)R_{zu}(\tau).\end{aligned} \quad (19)$$

Therefore,

$$R_{yu}(\tau)=b_0G(q)H(q)R_u(\tau)=b_0G_H(q)R_u(\tau). \quad (20)$$

The proof is completed.

3.1 Parameter identification for two linear elements

As discussed above, Theorem 1 is used to estimate the parameters of two linear elements, namely $\theta_G=\left[a_1,a_2,\cdots,a_{n_a},b_1,b_2,\cdots,b_{n_b}\right]^{\mathrm{T}}$ and $\theta_H=\left[c_1,c_2,\cdots,c_{n_c}\right]^{\mathrm{T}}$. Using a group of Gaussian signals $u(t)$ as input, according to the linear relationship between the cross-correlation function and the autocorrelation function described in Theorem 1, for the dynamic linear elements of the Wiener–Hammerstein system, the correlation analysis technique can be used to identify parameters of the two linear elements, that is,

$$\begin{aligned}G_H(q)&=G(q)H(q)\\&=\sum_{j=1}^{n_b} b_jq^{-j}\Big/\left[\left(1+\sum_{i=1}^{n_a} a_iq^{-i}\right)\left(1+\sum_{m=1}^{n_c} c_mq^{-m}\right)\right].\end{aligned} \quad (21)$$

Combining Eqs. (20) and (21) gives

$$R_{yu}(\tau)=\frac{b_0\sum_{j=1}^{n_b} b_jq^{-j}}{\left(1+\sum_{i=1}^{n_a} a_iq^{-i}\right)\left(1+\sum_{m=1}^{n_c} c_mq^{-m}\right)}R_u(\tau). \quad (22)$$

Here, set $\tau=1, 2, \cdots, P$ and $\tilde{b}_j=b_0b_j$. Then

$$R_{yu}(\tau)\left(1+\sum_{i=1}^{n_a} a_iq^{-i}\right)\left(1+\sum_{m=1}^{n_c} c_mq^{-m}\right)=\sum_{j=1}^{n_b}\tilde{b}_jq^{-j}R_u(\tau). \quad (23)$$

Hence,

$$\begin{aligned}R_{yu}(\tau)=&-\left(\left(1+\sum_{i=1}^{n_a} a_iq^{-i}\right)\left(1+\sum_{m=1}^{n_c} c_mq^{-m}\right)-1\right)R_{yu}(\tau)\\&+\sum_{j=1}^{n_b}\tilde{b}_jq^{-j}R_u(\tau).\end{aligned} \quad (24)$$

To simplify Eq. (24), define $(\overline{ac})_r$ ($r\in[1, n_a+n_c]$) as the coefficient of expression $-\left(\left(1+\sum_{i=1}^{n_a} a_iq^{-i}\right)\cdot\left(1+\sum_{m=1}^{n_c} c_mq^{-m}\right)\right)-1$, which is from high to low. Thus, Eq. (24) can be rewritten as

$$R_{yu}(\tau)=-\sum_{r=1}^{n_a+n_c}(\overline{ac})_rR_{yu}(\tau-r)+\sum_{j=1}^{n_b}\tilde{b}_jR_u(\tau-j). \quad (25)$$

The least-squares algorithm described in Eq. (26) is used to estimate the parameters of two elements:

$$\theta=R\Phi^{\mathrm{T}}(\Phi\Phi^{\mathrm{T}})^{-1}, \quad (26)$$

where

$$\theta=\left((\overline{ac})_1,(\overline{ac})_2,\cdots,(\overline{ac})_{n_a+n_c},\tilde{b}_1,\tilde{b}_2,\cdots,\tilde{b}_{n_b}\right),$$

$$R=\left(R_{yu}(1),R_{yu}(2),\cdots,R_{yu}(P)\right),$$

$$\Phi=\begin{bmatrix}-R_{yu}(0) & -R_{yu}(1) & -R_{yu}(2) & \cdots & -R_{yu}(P-1)\\ 0 & -R_{yu}(0) & -R_{yu}(1) & \cdots & -R_{yu}(P-2)\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & -R_{yu}(P-n_a-n_c)\\ R_u(0) & R_u(1) & R_u(2) & \cdots & R_u(P-1)\\ 0 & R_u(0) & R_u(1) & \cdots & R_u(P-2)\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & R_u(P-n_b)\end{bmatrix}.$$

Eqs. (27) and (28) can be used to estimate $R_{yu}(\tau)$ and $R_u(\tau)$, respectively:

$$R_{yu}(\tau)=\frac{1}{n}\sum_{t=1}^{n}y(t)u(t-\tau), \tag{27}$$

$$R_u(\tau)=\frac{1}{n}\sum_{t=1}^{n}u(t)u(t-\tau). \tag{28}$$

In this work, the ARX and AR models are used, so the zeros and gain exist only in the input dynamic linear element. According to Eq. (21), the zeros and poles of the product element are still composed of the zeros and poles of these two linear elements. The number of zeros generated by the product element is n_b-1, and the number of poles is n_a+n_c. Thus, when $n_a=i$, $n_b=j$, $n_c=m$, there are $(j-1)\mathrm{C}_{i+m}^{2}$ possible configuration methods.

To configure the zeros and poles correctly, for each set of zeros and poles, when the error between the actual output and the estimated output is the lowest, the zero-pole configuration is correctly configured (Shaikh and Barbé, 2019). After the zeros and poles are correctly configured, the transfer function in the form of zeros and poles is directly converted into the traditional transfer function form, and the linear parameters $\hat{\theta}_G=[\hat{a}_1,\hat{a}_2,\cdots,\hat{a}_{n_a},\hat{b}_1,\hat{b}_2,\cdots,\hat{b}_{n_b}]$ and $\hat{\theta}_H=[\hat{c}_1,\hat{c}_2,\cdots,\hat{c}_{n_c}]$ are identified.

3.2 Parameter identification for the nonlinear element

Based on previous analysis and discussion, parameters $\hat{a}_i$, $\hat{b}_j$, and $\hat{c}_m$ have been identified. In this subsection, a group of random signals $u_1(t)$ is applied to nonlinear element estimation. The clustering algorithm (Li F et al., 2017, 2022, 2023a, 2023b) is used to estimate center c_l and width σ_l. The key problem is to estimate weight w_l.

According to Eqs. (1) and (2), we have

$$y_1(t)+\sum_{m=1}^{n_c}c_m y_1(t-m)=\sum_{l=1}^{L}\phi_l\big(\hat{v}_1(t)\big)w_l+e(t). \tag{29}$$

Then, we can obtain

$$y_1(t)=-\sum_{m=1}^{n_c}c_m y_1(t-m)+\sum_{l=1}^{L}\phi_l\big(\hat{v}_1(t)\big)w_l+e(t). \tag{30}$$

Based on Eq. (30), the Wiener–Hammerstein system can be written as

$$y_1(t)=(\varphi(t))^{\mathrm{T}}\theta_1+e(t), \tag{31}$$

where

$$\varphi(t)=\big[-y_1(t-1),\ -y_1(t-2),\cdots,\ -y_1(t-n_c),\ \phi_1\big(\hat{v}_1(t)\big),\phi_2\big(\hat{v}_1(t)\big),\cdots,\phi_L\big(\hat{v}_1(t)\big)\big]^{\mathrm{T}},$$

$$\theta_1=\big[c_1,c_2,\cdots,c_{n_c},w_1,w_2\cdots,w_L\big]^{\mathrm{T}}.$$

Define the following criterion functions:

$$J(\theta_1)=\sum_{t=1}^{N}\big[y_1(t)-(\varphi(t))^{\mathrm{T}}\theta_1\big]^2. \tag{32}$$

Then, the recursive least-squares method is used to calculate the weights of NFN:

$$\hat{\theta}_1(t)=\hat{\theta}_1(t-1)+L(t)\big[y_1(t)-(\varphi(t))^{\mathrm{T}}\hat{\theta}_1(t-1)\big], \tag{33}$$

$$L(t)=\frac{P(t-1)\varphi(t)}{1+(\varphi(t))^{\mathrm{T}}P(t-1)\varphi(t)}, \tag{34}$$

$$P(t)=\big[I-L(t)(\varphi(t))^{\mathrm{T}}\big]P(t-1). \tag{35}$$

To describe parameter estimation methods, the procedures for estimation of the Wiener–Hammerstein system with ARX-NFN-AR based on hybrid signals are illustrated in Fig. 3.

4 Simulation results

4.1 Numerical simulation

To verify the effectiveness and superiority of the presented method, the Wiener–Hammerstein system was considered in which the nonlinearity is discontinuous functions.

$$v(t)=u(t)G(q)=u(t)\frac{0.6q^{-1}-0.4q^{-2}}{1-0.3q^{-1}+0.5q^{-2}},$$

$$z(t)=\begin{cases}1.2v(t)+0.4, & v(t)\le-2,\\ 0.4v(t)-1.2, & -2<v(t)\le-1,\\ 1.6v(t), & -1<v(t)\le1,\\ 0.4v(t)+1.2, & 1<v(t)\le2,\\ 1.2v(t)-0.4, & v(t)>2,\end{cases}$$

$$m(t)=H(q)z(t)=\frac{1}{1-0.7q^{-1}+0.9q^{-2}}z(t),$$

$$y(t)=m(t)+e(t).$$

In this research, 10 000 Gaussian signals were used as the input of the Wiener–Hammerstein system. The correlation analysis scheme was used to identify the two linear elements, and its zero and pole points were constructed. In this simulation, n_a=n_c=2, n_b=1, and thus there were six configuration methods. The zero and four poles were defined as zero=2/3, pole1=0.1500+0.6910i, pole2=0.1500−0.6910i, pole3=0.3500+0.8818i, pole4=0.3500−0.8818i. Table 1 shows the output error for each zero-pole configuration. The result of sequence 6 was optimal as it had the lowest MSE (mean squared error, $\mathrm{MSE}=\frac{1}{N}\sum_{t=1}^{N}(\mathrm{observed}_t-\mathrm{predicted}_t)^2$). The identification results for the two linear elements are shown in Fig. 4. The results of our proposed ARX model and AR model are close to those of the real model, and the error is small.

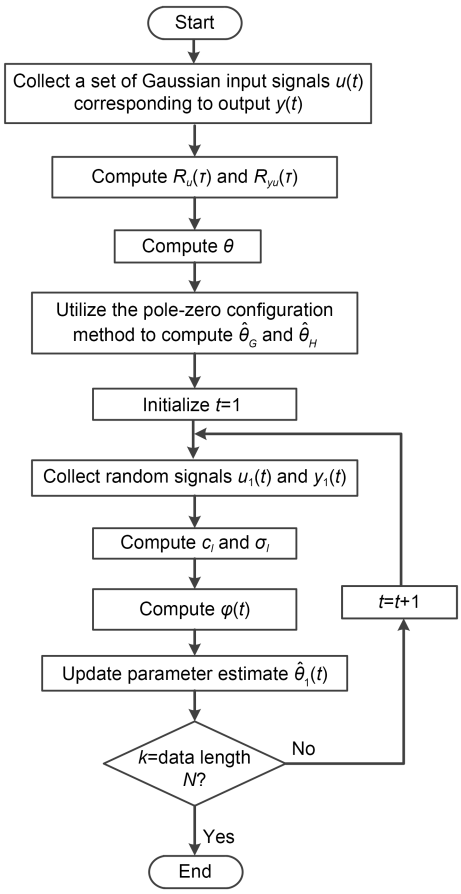

Fig. 3 Flowchart of the developed parameter estimation method

Table 1 Output error comparisons under different zero-pole configurations

Sequence	Zero-pole configuration	MSE
1	G(zero, pole1, pole2), H(pole3, pole4)	1.5136
2	G(zero, pole1, pole3), H(pole2, pole4)	0.6569
3	G(zero, pole1, pole4), H(pole2, pole3)	0.6566
4	G(zero, pole2, pole3), H(pole1, pole4)	1.1539
5	G(zero, pole2, pole4), H(pole1, pole3)	0.1657
6	G(zero, pole3, pole4), H(pole1, pole2)	0.0140

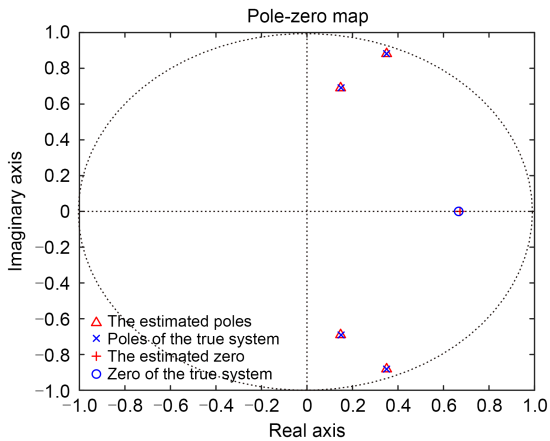

Fig. 4 Identification results for the two linear elements

After the zeros and poles of the two linear elements were obtained, 5000 random signals were used to identify the nonlinear element. The parameters S_0=0.986, ρ=1.1, and λ=0.1 were set, the center and the width of NFN were calculated by a clustering method, and the weights were estimated by the recursive least-squares method. To verify the accuracy of the proposed method, the nonlinear element was modeled by a polynomial model using the same data. Fig. 5 shows the approximation results of the nonlinear element using

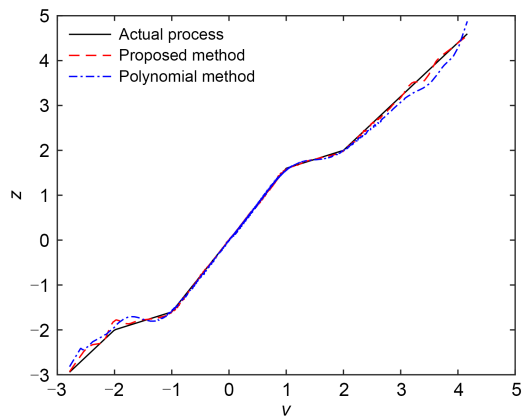

Fig. 5 Nonlinearity approximation using two models

the polynomial model and NFN. The MSE and MAE (maximum absolute error, $\text{MAE}=\frac{1}{N}\sum_{t=1}^{N}\left|\text{observed}_t-\text{predicted}_t\right|$) are listed in Table 2. According to Fig. 5 and Table 2, compared with polynomial modeling, NFN showed a stronger nonlinear approximation ability for piecewise nonlinear functions.

Table 2 Comparisons of MSE and MAE using two modeling methods

Modeling method	Order	MSE	MAE
Polynomial model	22	2.7E-3	0.0170
	23	2.0E-3	0.0166
	24	1.1E-3	0.0192
	25	0.1761	0.0355
NFN		3.2223E-4	0.0116

To demonstrate the effectiveness of the proposed strategy, 200 randomly generated signals of [0, 7] were used for prediction, and the reframed multi-innovation least-squares (RMILS) method (Li LW et al., 2020) and the auxiliary model based multi-innovation improved PSO (AM-MIIPSO) method (Zong et al., 2021) were compared. Comparisons of the predicted output using the different methods are shown in Fig. 6, and comparisons of the respective MSE and MAE in Table 3.

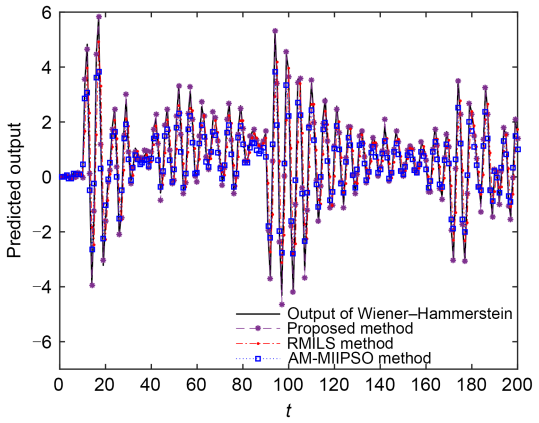

Fig. 6 Comparisons of the predicted output using different methods

Table 3 Comparisons of MSE and MAE using different methods

Method	MSE	MAE
Proposed	0.0014	0.0930
RMILS	0.0828	0.6657
AM-MIIPSO	0.0336	0.4378

Fig. 6 shows that the proposed strategy can predict the output of the Wiener–Hammerstein system better than the RMILS method and the AM-MIIPSO method. From Table 3, taking MSE and MAE as the criteria for evaluating the prediction performance, the MSE and MAE values of the proposed strategy were the smallest among the three methods, so the obtained quality and prediction accuracy of the Wiener–Hammerstein system were the best. Thus, the proposed strategy can effectively identify the neural fuzzy Wiener–Hammerstein system.

Remark 2 In the simulation, λ=0.1 and ρ=1.1. Simulation results showed that the nonlinear element can be approximated exactly by selecting suitable parameters λ and ρ. A smaller parameter λ and larger parameter ρ led to a smaller model error, but slower convergence. Thus, the parameter selection should balance the system error and convergence speed.

4.2 Application to a flexible robot arm

A benchmark data set taken from the identification database DAISY (de Moor et al., 1997) was used. The system consists of an arm mounted on a motor whose input is reaction torque and output is acceleration. The measurement data set contains 1024 samples, divided into two parts: the first 800 samples are used for identification and the remaining 204 samples for prediction. The input and output of the experimental robot arm are shown in Fig. 7.

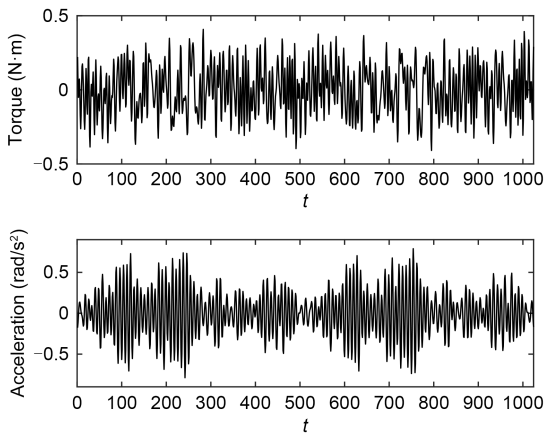

Fig. 7 Input and output of the flexible robot arm

Taking the reaction torque of the ground structure as the input and acceleration as the output, a neural fuzzy nominal model was trained. The nominal model

training results of the flexible robot arm system are shown in Fig. 8, and the output of the nominal model in Fig. 9. Figs. 8 and 9 showed that the convergence speed of the training nominal model was high, and that the trained acceleration can track the actual acceleration well.

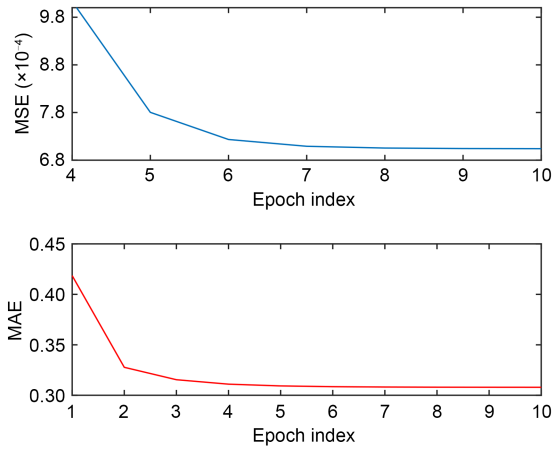

Fig. 8 MSE and MAE of the neural fuzzy nominal model

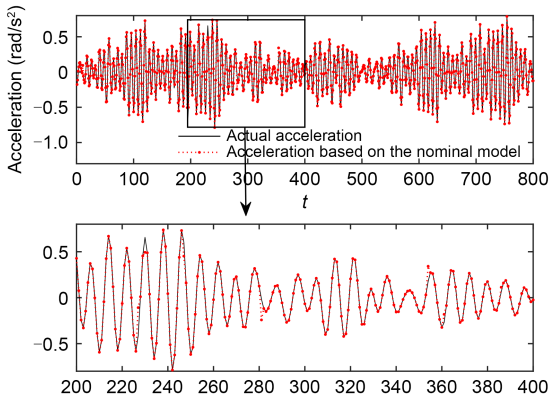

Fig. 9 Output of the nominal model

For the two-stage identification method proposed in this research, 1000 Gaussian signals and 800 reaction torque were used to form hybrid signals. The output of the Gaussian signal through the nominal model and the acceleration data corresponding to the reaction torque were used as the output of the Wiener–Hammerstein system.

First, based on the input and output of the 1000 Gaussian signals, parameters of the two linear elements were identified using the correlation analysis technique and the zero-pole configuration scheme. In this simulation, the pole of the ARX model was set as n_a=2, zero-point n_b=1, and the pole of the AR model as n_c=2. Thus, there were six configuration methods. Table 4 shows the output error for each zero-pole configuration. The result of sequence 1 was optimal. Thus, the ARX model $G(q)=\dfrac{-0.3246q^{-1}-0.6926q^{-2}}{1+0.4472q^{-1}+0.6372q^{-2}}$ and the AR model $H(q)=\dfrac{1}{1-0.1389q^{-1}+0.07178q^{-2}}$ can be obtained.

Table 4 Comparison of MSE under different zero-pole configurations

Sequence	Zero-pole configuration	MSE
1	*G*(zero, pole1, pole2), *H*(pole3, pole4)	0.0017
2	*G*(zero, pole1, pole3), *H*(pole2, pole4)	0.1817
3	*G*(zero, pole1, pole4), *H*(pole2, pole3)	0.0538
4	*G*(zero, pole2, pole3), *H*(pole1, pole4)	0.0533
5	*G*(zero, pole2, pole4), *H*(pole1, pole3)	0.1818
6	*G*(zero, pole3, pole4), *H*(pole1, pole2)	1.2082

Second, based on the reaction torque and acceleration data, parameters of the nonlinear element were estimated based on the clustering algorithm and recursive least-squares method. The number of nonlinear fuzzy rules (L) was varied to optimize the system. Fig. 10 shows the MSE at different L's. The MSE was the smallest when L was 15.

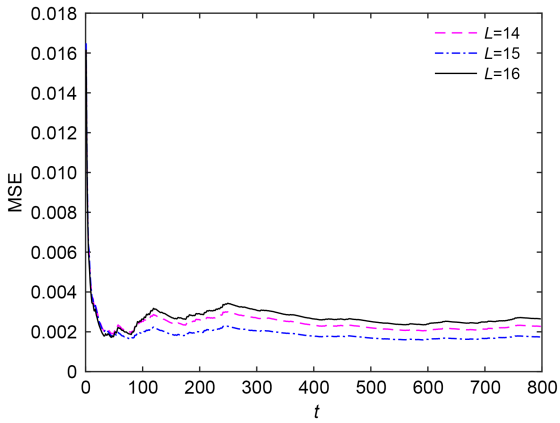

Fig. 10 MSE at different *L*'s

Finally, the remaining data were used as a test set to confirm the effectiveness of the proposed approach in comparison with the polynomial model based Wiener–Hammerstein system (P-W-H) (the polynomial order was eight), RMILS method (Li LW et al., 2020),

and AM-MIIPSO method (Zong et al., 2021). The results of the acceleration prediction comparison of the robot arm are shown in Fig. 11, and the MSE and MAE comparisons of the three methods are given in Table 5. Fig. 12 shows that, compared with the P-W-H, RMILS, and AM-MIIPSO methods, our proposed method can better predict the acceleration of the robot arm using the identified Wiener–Hammerstein system. From Table 5, the MSE and MAE of the predicted acceleration were the lowest using the proposed method. The MSE was reduced by 74.07% and the MAE by 49.93% compared with the minimum value of the other methods. Thus, the proposed strategy can effectively identify the Wiener–Hammerstein system.

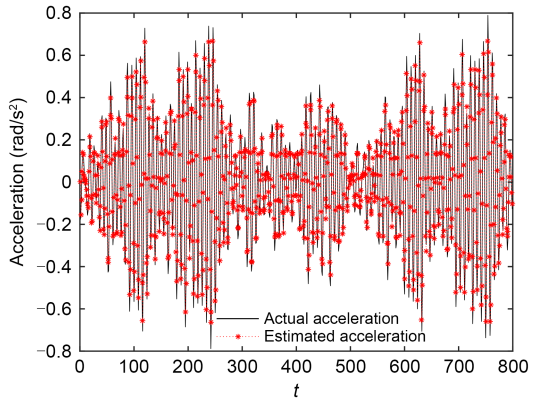

Fig. 11 Actual and estimated acceleration

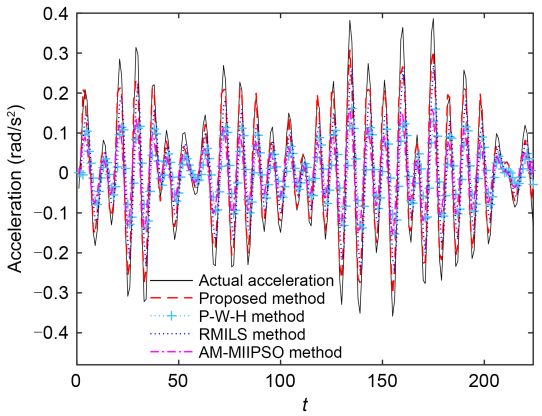

Fig. 12 Comparison of the acceleration prediction of the robot arm using different methods

Table 5 Comparisons of MSE and MAE among different methods

Method	MSE	MAE
Proposed	0.0021	0.0375
P-W-H	0.0229	0.1256
RMILS	0.0081	0.0749
AM-MIIPSO	0.0151	0.1013

5 Conclusions

In this paper, a novel separation identification approach of the Wiener–Hammerstein system with ARX-NFN-AR based on hybrid signals is proposed. The two main advantages of the presented scheme are: (1) Based on the characteristic that Gaussian signals do not excite the linear element, the parameters of the two linear elements and the nonlinear element are identified independently using the designed hybrid signals, which simplifies the identification process and reduces complexity; (2) The correlation analysis technique is used for Wiener–Hammerstein system estimation, thereby solving the problem that the intermediate variable information cannot be measured, and the interference of output noise is handled.

Contributors

Feng LI designed the research. Feng LI and Hao YANG processed the data. Feng LI drafted the paper. Hao YANG helped organize the paper. All the authors revised and finalized the paper.

Conflict of interest

All the authors declare that they have no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Ase H, Katayama T, 2015. A subspace-based identification of Wiener–Hammerstein benchmark model. *Contr Eng Pract*, 44:126-137. https://doi.org/10.1016/j.conengprac.2015.07.011

de Moor B, de Gersem P, de Schutter B, et al., 1997. DAISY: a database for identification of systems. *Comput Sci*, 38(3):4-5.

dos Santos PL, Ramos JA, de Carvalho JLM, 2012. Identification of a benchmark Wiener–Hammerstein: a bilinear and Hammerstein–Bilinear model approach. *Contr Eng Pract*, 20(11):1156-1164. https://doi.org/10.1016/j.conengprac.2012.04.002

Falck T, Dreesen P, de Brabanter K, et al., 2012. Least-squares support vector machines for the identification of Wiener–Hammerstein systems. *Contr Eng Pract*, 20(11):1165-1174.

https://doi.org/10.1016/j.conengprac.2012.05.006

Ghanmi A, Elloumi M, Salhi H, et al., 2020. A recursive hierarchical parametric estimation algorithm for nonlinear systems described by Wiener–Hammerstein models. *Asian J Contr*, 22(3):1065-1074. https://doi.org/10.1002/asjc.2029

Hafsi S, Laabidi K, Ksouri-Lahmari M, 2012. Identification of Wiener–Hammerstein model with multisegment piecewise-linear characteristic. IEEE Mediterranean Electrotechnical Conf, p.5-10. https://doi.org/10.1109/MELCON.2012.6196367

Han Y, de Callafon RA, 2012. Identification of Wiener–Hammerstein benchmark model via rank minimization. *Contr Eng Pract*, 20(11):1149-1155. https://doi.org/10.1016/j.conengprac.2012.05.004

Janjanam L, Saha SK, Kar R, et al., 2022. Optimal design of cascaded Wiener–Hammerstein system using a heuristically supervised discrete Kalman filter with application on benchmark problems. *Expert Syst Appl*, 200:117065. https://doi.org/10.1016/j.eswa.2022.117065

Jia L, Xiong Q, Li F, 2017. Correlation analysis method based SISO neuro-fuzzy Wiener model. *J Process Contr*, 58: 73-89. https://doi.org/10.1016/j.jprocont.2017.08.002

Katayama T, Ase H, 2016. Linear approximation and identification of MIMO Wiener–Hammerstein systems. *Automatica*, 71:118-124. https://doi.org/10.1016/j.automatica.2016.04.040

Ławryńczuk M, 2016. Nonlinear predictive control of dynamic systems represented by Wiener–Hammerstein models. *Nonl Dynam*, 86(2):1193-1214. https://doi.org/10.1007/s11071-016-2957-0

Li F, Li J, Peng DG, 2017. Identification method of neuro-fuzzy-based Hammerstein model with coloured noise. *IET Contr Theory Appl*, 11(17):3026-3037. https://doi.org/10.1049/iet-cta.2017.0306

Li F, Zheng T, He NB, et al., 2022. Data-driven hybrid neural fuzzy network and ARX modeling approach to practical industrial process identification. *IEEE/CAA J Autom Sin*, 9(9):1702-1705.

Li F, Jia L, Gu Y, 2023a. Identification of nonlinear process described by neural fuzzy Hammerstein–Wiener model using multi-signal processing. *Adv Manuf*, 11:694-707. https://doi.org/10.1007/s40436-022-00426-w

Li F, Zhu XJ, He NB, et al., 2023b. Parameter learning for the nonlinear system described by Hammerstein model with output disturbance. *Asian J Contr*, 25(2):886-898. https://doi.org/10.1002/asjc.2829

Li LW, Ren XM, 2018. Identification of nonlinear Wiener–Hammerstein systems by a novel adaptive algorithm based on cost function framework. *ISA Trans*, 80:146-159. https://doi.org/10.1016/j.isatra.2018.07.015

Li LW, Ren XM, Guo FM, 2018. Modified multi-innovation stochastic gradient algorithm for Wiener–Hammerstein systems with backlash. *J Franklin Inst*, 355(9):4050-4075. https://doi.org/10.1016/j.jfranklin.2018.03.015

Li LW, Zhang HL, Ren XM, 2020. A modified multi-innovation algorithm to turntable servo system identification. *Circ Syst Signal Process*, 39(9):4339-4353. https://doi.org/10.1007/s00034-020-01392-z

Martin E, Lennart L, 2005. Linear approximations of nonlinear FIR systems for separable input processes. *Automatica*, 41(3):459-473. https://doi.org/10.1016/j.automatica.2004.11.016

Mu BQ, Chen HF, 2014. Recursive identification of errors-in-variables Wiener–Hammerstein systems. *Eur J Contr*, 20(1):14-23. https://doi.org/10.1016/j.ejcon.2013.10.005

Mzyk G, Wachel P, 2017. Kernel-based identification of Wiener–Hammerstein system. *Automatica*, 83:275-281. https://doi.org/10.1016/j.automatica.2017.06.038

Naitali A, Giri F, 2016. Wiener–Hammerstein system identification—an evolutionary approach. *Int J Syst Sci*, 47(1): 45-61. https://doi.org/10.1080/00207721.2015.1027758

Paduart J, Lauwers L, Pintelon R, et al., 2012. Identification of a Wiener–Hammerstein system using the polynomial nonlinear state space approach. *Contr Eng Pract*, 20(11): 1133-1139. https://doi.org/10.1016/j.conengprac.2012.06.006

Piroddi L, Farina M, Lovera M, 2012. Black box model identification of nonlinear input-output models: a Wiener–Hammerstein benchmark. *Contr Eng Pract*, 20(11):1109-1118. https://doi.org/10.1016/j.conengprac.2012.03.002

Rijlaarsdam D, Oomen T, Nuij P, et al., 2012. Uniquely connecting frequency domain representations of given order polynomial Wiener–Hammerstein systems. *Automatica*, 48(9): 2381-2384. https://doi.org/10.1016/j.automatica.2012.06.006

Ross S, 2014. Introduction to Probability Models (11th Ed.). Elsevier, Amsterdam, the Netherlands. https://doi.org/10.1016/C2012-0-03564-8

Shaikh MAH, Barbé K, 2019. Wiener–Hammerstein system identification: a fast approach through Spearman correlation. *IEEE Trans Instrum Meas*, 68(5):1628-1636. https://doi.org/10.1109/TIM.2019.2896366

Sjöberg J, Schoukens J, 2012. Initializing Wiener–Hammerstein models based on partitioning of the best linear approximation. *Automatica*, 48(2):353-359. https://doi.org/10.1016/j.automatica.2011.07.007

Sjöberg J, Lauwers L, Schoukens J, 2012. Identification of Wiener–Hammerstein models: two algorithms based on the best split of a linear model applied to the SYSID'09 benchmark problem. *Contr Eng Pract*, 20(11):1119-1125. https://doi.org/10.1016/j.conengprac.2012.07.001

Škrjanc I, 2021. An evolving concept in the identification of an interval fuzzy model of Wiener–Hammerstein nonlinear dynamic systems. *Inform Sci*, 581:73-87. https://doi.org/10.1016/j.ins.2021.09.004

Tiels K, Schoukens M, Schoukens J, 2014. Generation of initial estimates for Wiener–Hammerstein models via basis function expansions. *IFAC Proc Vol*, 47(3):481-486. https://doi.org/10.3182/20140824-6-ZA-1003.02292

Wang ZY, Zhang Y, Jin QB, et al., 2022. Wiener models robust identification of multi-rate process with time-varying delay using expectation-maximization algorithm. *J Process Contr*, 118:126-138. https://doi.org/10.1016/j.jprocont.2022.09.003

Weber D, Gühmann C, 2021. Non-autoregressive vs autoregressive neural networks for system identification. *IFAC-PapersOnLine*, 54(20):692-698. https://doi.org/10.1016/j.ifacol.2021.11.252

Zong TC, Li JH, Lu GP, 2021. Auxiliary model-based multi-innovation PSO identification for Wiener–Hammerstein systems with scarce measurements. *Eng Appl Artif Intell*, 106:104470. https://doi.org/10.1016/j.engappai.2021.104470