



MILM hybrid identification method of fractional order neural-fuzzy Hammerstein model

Qian Zhang · Hongwei Wang · Chunlei Liu

Received: 23 September 2021 / Accepted: 13 February 2022 / Published online: 12 March 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract Aiming at the difficult identification of fractional order Hammerstein nonlinear systems, including many identification parameters and coupling variables, unmeasurable intermediate variables, difficulty in estimating the fractional order, and low accuracy of identification algorithms, a multiple innovation Levenberg–Marquardt algorithm (MILM) hybrid identification method based on the fractional order neuro-fuzzy Hammerstein model is proposed. First, a fractional order discrete neuro-fuzzy Hammerstein system model is constructed; secondly, the neuro-fuzzy network structure and network parameters are determined based on fuzzy clustering, and the self-learning clustering algorithm is used to determine the antecedent parameters of the neuro-fuzzy network model; then the multiple innovation principle is combined with the Levenberg–Marquardt algorithm, and the MILM hybrid algorithm is used to estimate the

linear module parameters and fractional order. Finally, the academic example of the fractional order Hammerstein nonlinear system and the example of a flexible manipulator are identified to prove the effectiveness of the proposed algorithm.

Keywords Fractional order Hammerstein model · Neuro-fuzzy network · Levenberg–Marquardt algorithm · Principle of multiple innovation

1 Introduction

In recent years, many viscoelastic materials (including viscoelastic solid and fluid substances) [1], “anomalous” diffusion (plasma motion at high temperature and high pressure) [2], thermal conduction of soft substances (also known as complex fluid, a complex state between ideal solids and liquids) [3] have emerged in complex physics, mechanics, chemistry, biology and engineering, which cannot be well described physically by integer order models, and the introduction of fractional order models can better describe the dynamics of the actual above-mentioned systems characteristic. These phenomena involve the memory and inheritance of physical and mechanical processes, paths and dependencies, and global correlations. The use of fractional calculus in the modeling process has different geometric interpretations or physical meanings.

Q. Zhang (✉) · H. Wang · C. Liu (✉)
School of Electrical Engineering, Xinjiang University,
Urumqi 830047, People’s Republic of China
e-mail: 1072592647@qq.com

C. Liu
e-mail: 993486933@qq.com

H. Wang
e-mail: wq272427yrx@163.com

H. Wang
School of Control Science and Engineering Dalian,
University of Technology, Dalian 110024, People’s
Republic of China

Both the fractional order model and the integer order model can describe actual work such as physics, chemistry, and mechanics. However, the physical meanings of fractional differential equations and integer differential equation models are fundamentally different. In terms of time, the integer order differential equation is characterized by a change or a certain property at a certain moment in a physical, chemical or mechanical process, while the property represented by the fractional order differential equation is related to the entire dynamic development and change process of the system; in terms of space, integer order spatial differential equations describe the local properties of a physical process at a certain position in space, while the physical process described by fractional order spatial differential equations is related to the changes in the entire space [4]. In terms of model generalization ability, the fractional order differential equation model overcomes the shortcomings of the traditional integer order differential equation model that the theoretical model does not match the experimental results when describing some complex physical and mechanical processes, and can better simulate the dynamic characteristics of the actual system, and can achieve better consistency between experimental effects and research models [5].

Based on actual field sampled data, the establishment of a system process model is generally called a system identification method. Practice has proved that some systems in actual engineering, such as viscoelastic systems, electrochemical diffusion, and fractal systems, are often difficult to accurately describe with traditional integer order models. This requires the use of system identification methods to accurately describe the actual systems with fractional order models. In the field of control research, the use of system identification methods to obtain system fractional order model descriptions is a common and effective modeling method.

According to the different forms of fractional order models used, the identification models can be divided into fractional order transfer function models and fractional order state-space models. In terms of fractional order system identification based on transfer function model, Victor S et al. used a fractional Poisson filter to provide a time domain identification method based on equation error and output error [6]; Based on the auxiliary variable model, Fahim SM extended this method to fractional order systems, and

gave the optimal algorithm of the improved auxiliary variable method [7]; Karima Hammar et al. studied the modeling problem of Hammerstein–Wiener nonlinear fractional order system on the basis of discrete transfer function. Using the Levenberg–Marquardt identification algorithm, the parameter identification and fractional parameter estimation in the model are given [8]. This paper adopts a gradient descent identification algorithm similar to single innovation, which has the problem of low identification accuracy; In addition, Shalaby R and Malti R proposed identification methods for fractional order continuous systems based on orthogonal basis functions and block impulse functions, respectively [9, 10].

Based on the identification of fractional state-space equations, Jonscher and Westerlund established a fractional order state-space model containing inductance and capacitance based on fractional calculus [11, 12]; Based on the fractional order state space model of the inductor-capacitor and the improved Oustaloup method, Tan Cheng et al. used the fractional order Lyapunov stability theory to design a nonlinear controller to control the boost circuit in pseudo-continuous mode [13]. In addition, there are also references that give short-term memory and frequency domain identification methods for fractional order continuous system state-space models [14–17].

The Hammerstein model is a model that combines a dynamic linear model with a static nonlinear model. It can describe a large class of industrial nonlinear processes. Therefore, its identification research has attracted widespread attention from researchers. Classic two-stage identification methods such as over-parameterization and iterative methods, random methods, frequency domain methods, and blind identification methods are often used [18–24], which separate the static nonlinear module from the dynamic linear module, but it has the shortcoming that if the system has process noise, the product term of the identified parameter has a large deviation; and it can separate the original system parameters more accurately only when the rank of the parameter product matrix is 1.

2 Related work and innovations

The fractional order Hammerstein nonlinear model can better describe the dynamic characteristics of the

original actual system, but its identification has the following difficulties: (1) There are many identification parameters and coupling variables; (2) the intermediate variables cannot be measured; (3) the fractional order is difficult to estimate; (4) the accuracy of the identification algorithm is low. The methods proposed in the references are all based on polynomial fractional order Hammerstein nonlinear systems. For multivariable functions, this type of system requires a large number of parameters and high-order numbers, and the nonlinearity is difficult to parameterize. However, for the mathematical model description of the nonlinear module of the nonlinear Hammerstein system, data-based neural networks, fuzzy systems, and neuro-fuzzy systems can also be used; these models are very suitable for processing complex nonlinear systems. Since these models have been proven to have the universal ability to approximate nonlinear systems, they are often used to build nonlinear system models. Therefore, this paper proposes a multi-innovation Levenberg–Marquardt (MILM) hybrid identification method based on the fractional order neuro-fuzzy Hammerstein model. The main innovations are as follows:

- (1) Use the neuro-fuzzy network to fit the static nonlinear module of the fractional order Hammerstein nonlinear model;
- (2) A self-learning neuro-fuzzy network method based on fuzzy clustering is proposed to determine the structure and parameters of the neuro-fuzzy network;
- (3) The MILM hybrid identification algorithm is proposed. The L-M identification algorithm is used to estimate the dynamic linear module parameters, and the scalar innovation is expanded into a vector of multiple innovations, and the information vector is expanded to an information matrix, which improves the accuracy of the single innovation L-M identification algorithm.

The algorithm proposed in this paper with these innovative points can improve the convergence and accuracy of the identification algorithm, and it has reference significance for the author and other scholars for the modeling and identification of fractional order Hammerstein models.

The overall structure of this paper is arranged as follows: Sect. 2 expounds the related work and

innovations of this paper; Sect. 3 describes the fractional calculus operator and the fractional order linear model, and constructs a fractional order discrete neuro-fuzzy Hammerstein system model; Sect. 4 proposes the self-learning neuro-fuzzy network method and MILM hybrid identification algorithm; Sect. 5 uses the fractional order Hammerstein nonlinear academic example of the discontinuous function and the example of a flexible manipulator for verification, and compares the accuracy of the identification result of the example of the flexible manipulator with the single innovation L-M algorithm in [8].

3 Description of the problem

This section introduces the basic concepts related to fractional calculation and describes the fractional order linear model and the fractional order neuro-fuzzy Hammerstein model. First, the basic principles of fractional calculus used in this paper are given. Secondly, the linear model of the fractional order transfer function is described and defined in detail. Finally, the fractional order neuro-fuzzy Hammerstein model is defined, and the nonlinearity of the fractional order Hammerstein model is fitted with a neuro-fuzzy network.

3.1 Fractional Calculus

The calculation and application of fractional order systems have attracted the attention of many researchers, and they have been applied in the fields of identification, control, filtering, and fault diagnosis. Fractional integral and differential calculations have been mentioned in many references [25–27]. Fractional calculus operators have different definitions in different references, mainly including GL fractional calculus operators, RL fractional calculus operators and Caputo fractional calculus operators. The GL calculus operator is defined as follows:

$$\Delta^\alpha x(kh) = \frac{1}{h^\alpha} \sum_{j=0}^k (-1)^j \binom{\alpha}{j} x((k-j)h). \quad (1)$$

In Eq. (1), Δ^α represents a fractional difference operator with order α , the initial value is 0, $x(kh)$ represents a function of $t = kh$, k is the number of

samplings, the system sampling time is h , and $\binom{\alpha}{j}$ is defined by the following equation:

$$\binom{\alpha}{j} = \begin{cases} 1 & \text{for } j = 0 \\ \frac{\alpha(\alpha - 1) \cdots (\alpha - j + 1)}{j!} & \text{for } j > 0 \end{cases} \quad (2)$$

According to Eq. (2), the following recurrence relation can be obtained,

$$\begin{cases} \beta(0) = 1 \\ \beta(j) = \beta(j - 1) \frac{(j - \alpha - 1)}{j} \end{cases} \text{ for } j = 1, \dots, k \quad (3)$$

Among them, $\beta(j) = (-1)^j \binom{\alpha}{j}$, combining Eqs. (2) and (3), Eq. (1) can be written as:

$$\Delta^\alpha x(kh) = \frac{1}{h^\alpha} \sum_{j=0}^k \beta(j)x((k - j)h) \quad (4)$$

Assuming that the system sampling time $h = 1$, Eq. (4) can be sorted into the following equation:

$$\Delta^\alpha x(k) = \sum_{j=0}^k \beta(j)x(k - j) \quad (5)$$

In the following chapters, Eq. (5) is used as a fractional order operator to study the modeling of fractional order nonlinear systems.

3.2 Description of the fractional order linear model

In this paper, consider the form of the following discrete linear transfer function:

$$y(k) = G(z)u(k) = \frac{B(z)}{A(z)}u(k) \quad (6)$$

where $u(k)$ and $y(k)$ are the system input and output, respectively. $A(z)$ and $B(z)$ are the denominator and numerator polynomial of the transfer function, $A(z) = 1 + a_1z^{-\alpha_1} + a_2z^{-\alpha_2} + \dots + a_{n_a}z^{-\alpha_{n_a}}$, $B(z) = b_1z^{-\gamma_1} + b_2z^{-\gamma_2} + \dots + b_{n_b}z^{-\gamma_{n_b}}$; α_i and γ_j ($i = 1, 2, \dots, n_a$, $j = 1, 2, \dots, n_b$) are the fractional orders of the corresponding polynomial, $\alpha_i \in \mathbb{R}^+$, $\gamma_j \in \mathbb{R}^+$, and z^{-1} is the backshift operator, that is, $z^{-1}y(k) = y(k - 1)$.

For Eq. (6), when the fractional order is completely different, the fractional order model of Eq. (6) is defined as a general non-identical (disproportionate)

fractional order system; otherwise, each fractional order is an integer multiple of the base order α (α is the order factor), that is, $\alpha_i = i\alpha, \gamma_j = j\alpha$ ($i = 1, 2, \dots, n_a; j = 1, 2, \dots, n_b$), at this time the model is defined as a homogeneous (proportionate) order system. In this paper, considering the proportional fractional order system, Eq. (6) can be written as:

$$\begin{aligned} y(k) &= \frac{B(z)}{A(z)}u(k) \\ &= \frac{b_1z^{-\alpha} + b_2z^{-2\alpha} + \dots + b_{n_b}z^{-n_b\alpha}}{1 + a_1z^{-\alpha} + a_2z^{-2\alpha} + \dots + a_{n_a}z^{-n_a\alpha}}u(k) \end{aligned} \quad (7)$$

Using the discrete fractional operator Δ of Eqs. (5), (7) can be written as the following equation:

$$\begin{aligned} y(k) &= [1 - A(z)]y(k) + B(z)u(k) \\ &= -a_1\Delta^\alpha y(k - 1) - a_2\Delta^\alpha y(k - 2) - \dots - a_{n_a}\Delta^\alpha y(k - n_a) \\ &\quad + b_1\Delta^\alpha u(k - 1) + b_2\Delta^\alpha u(k - 2) + \dots \\ &\quad + b_{n_b}\Delta^\alpha u(k - n_b) \\ &= -\sum_{i=1}^{n_a} a_i\Delta^\alpha y(k - i) + \sum_{i=1}^{n_b} b_i\Delta^\alpha u(k - i) \end{aligned} \quad (8)$$

Equation 8 is used as the model of the fractional order linear module part.

3.3 Description of the fractional order Neural-Fuzzy Hammerstein Model

Consider a Hammerstein model with external noise, which is described as follows:

$$S_c: \begin{cases} \bar{u}(k) = f(u(k)) & \text{Static nonlinear module part} \\ A(z)y(k) = B(z)\bar{u}(k) + v(k) & \text{Dynamic linear module part} \end{cases} \quad (9)$$

where $u(k)$ and $y(k)$, respectively, represent the input and output of the system at the k -th time; $\bar{u}(k)$ represents the output of the static nonlinear module at the k -th time; $v(k)$ is a white noise sequence with a mean value of 0 and a variance of σ^2 ; $f(\bullet)$ represents a nonlinear function, $A(z)$ and $B(z)$ are fractional polynomials of the same element, and their expressions are given as follows:

$$\begin{aligned}
 A(z) &= 1 + a_1z^{-\alpha} + a_2z^{-2\alpha} + \dots + a_{n_a}z^{-n_a\alpha} \\
 B(z) &= b_1z^{-\alpha} + b_2z^{-2\alpha} + \dots + b_{n_b}z^{-n_b\alpha}
 \end{aligned}
 \tag{10}$$

For model (9), its structure can be represented by Fig. 1.

In Fig. 1, the static nonlinear module is fitted with a neural-fuzzy network model, which is specifically the following T-S fuzzy model:

$$R^l : \text{ If } u(kT) \text{ is } F^l, \text{ then } \bar{u}(kT) = w_l, \quad l = 1, 2, \dots, c
 \tag{11}$$

Among them, R^l represents the l -th rule; $u(k)$ is the input of the model; w_l is the subsequent parameter of the l -th rule; c is the total number of rules; F^l is the fuzzy set, which is described by the Gaussian membership function μ_{F^l} .

The abovementioned T-S fuzzy model can be described by a neuro-fuzzy network model. Figure 2 shows the specific structure of the network, which has four levels. The specific conditions are as follows:

- Level 1 (The input level). This level is directly transferred to the next level by the input signal $u(k)$;
- Level 2 (Membership function level). Receive the signal from the input level and calculate the membership function of the input signal. The degree of membership in this level is taken as a following Gaussian function form:

$$\mu_{F^l}(u(k)) = \exp\left(-\frac{(u(k) - c_l)^2}{\delta_l^2}\right), \tag{12}$$

$$l = 1, 2, \dots, c.$$

Among them, c_l and δ_l^2 are the center and width of the Gaussian membership function, respectively, and they and the rule number c are the antecedent parameters of the neural-fuzzy network.

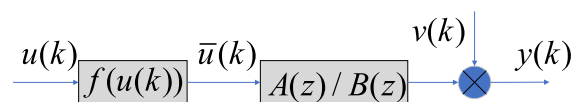


Fig. 1 The structure diagram of the fractional order neuro-fuzzy Hammerstein model

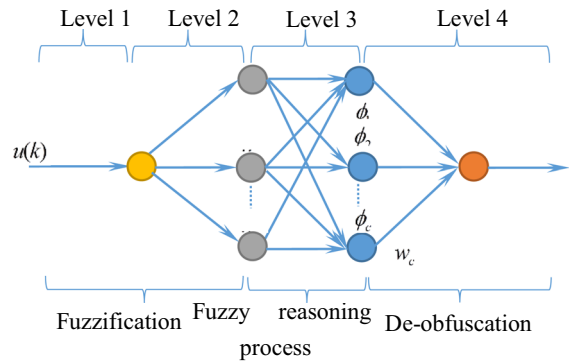


Fig. 2 The structure of the neuro-fuzzy network

- Level 3 (Fuzzy rules level). The number of nodes in this level is equal to the number of rules.
- Level 4 (The output level). This level performs de-obfuscation operation, and its output is given as follows:

$$\bar{u}(k) = \sum_{l=1}^c w_l \phi_l(u(k)) \tag{13}$$

Among them, $\phi_l(u(k)) = \frac{\mu_{F^l}(u(k))}{\sum_{l=1}^c \mu_{F^l}(u(k))}$ and w_l is the

subsequent parameter of the neuro-fuzzy network.

Specifically, the input–output relationship of the neuro-fuzzy Hammerstein system can be expressed by the following equation,

$$y(k) = [1 - A(z)]y(k) + B(z)f(u(k)) + v(k) \tag{14}$$

The static nonlinear function $f(\cdot)$ is a nonlinear function composed of several fuzzy basis functions. They are given as follows:

$$\begin{aligned}
 \bar{u}(k) &= f(u(k)) \\
 &= w_1 \phi_1(u(k)) + w_2 \phi_2(u(k)) + \dots + w_c \phi_c(u(k)) \\
 &= \sum_{l=1}^c w_l \phi_l(u(k))
 \end{aligned}
 \tag{15}$$

Among them, $\phi_1(u(k)), \dots, \phi_c(u(k))$ is c fuzzy basis functions, $\phi_l(u(k)) = \frac{\mu_{F^l}(u(k))}{\sum_{l=1}^c \mu_{F^l}(u(k))}$, and μ_{F^l} is the

Gaussian membership function of the l -th rule.

Substituting Eq. (15) into Eq. (14), we have

$$\begin{aligned}
 y(k) &= - \sum_{i=1}^{n_a} a_i z^{-i\alpha} y(k) + \sum_{j=1}^{n_b} b_j z^{-j\alpha} f(u(k)) + v(k) \\
 &= - \sum_{i=1}^{n_a} a_i z^{-i\alpha} y(k) + \sum_{j=1}^{n_b} b_j z^{-j\alpha} \\
 &\quad \sum_{l=1}^c w_l \phi_l(u(k)) + v(k)
 \end{aligned}$$

Using the discrete fractional operator Δ of Eqs. (5), (15) can be written as:

$$\begin{aligned}
 y(k) &= - \sum_{i=1}^{n_a} a_i \Delta^\alpha y(k-i) \\
 &\quad + \sum_{j=1}^{n_b} b_j \sum_{l=1}^l w_l \Delta^\alpha \phi_l(u(k-j)) + v(k) \\
 &= - \sum_{i=1}^{n_a} a_i \Delta^\alpha y(k-i) + w_1 \sum_{j=1}^{n_b} b_j \Delta^\alpha \phi_1(u(k-j)) + \dots \\
 &\quad + w_c \sum_{j=1}^{n_b} b_j \Delta^\alpha \phi_c(u(k-j)) + v(k)
 \end{aligned} \tag{16}$$

Define the parameter vector as:

$$\begin{aligned}
 \mathbf{a} &= [a_1 a_2 \dots a_{n_a}]^T \in \mathbb{R}^{n_a}, \quad \mathbf{b} = [b_1 b_2 \dots b_{n_b}]^T \\
 &\in \mathbb{R}^{n_b}, \quad \mathbf{w} = [w_1 w_2 \dots w_c]^T \in \mathbb{R}^c
 \end{aligned} \tag{17}$$

In order to obtain the unique parameters of the model, the model parameters are standardized. For this reason, set the first element in the parameter \mathbf{w} to be 1, $w_1 = 1$. On this basis, Eq. (16) can be rewritten as:

$$\begin{aligned}
 y(k) &= - \sum_{i=1}^{n_a} a_i \Delta^\alpha y(k-i) + \sum_{j=1}^{n_b} b_j \Delta^\alpha \phi_1(u(k-j)) + \dots \\
 &\quad + w_2 \sum_{j=1}^{n_b} b_j \Delta^\alpha \phi_2(u(k-j)) + \dots \\
 &\quad + w_c \sum_{j=1}^{n_b} b_j \Delta^\alpha \phi_c(u(k-j)) + v(k)
 \end{aligned} \tag{18}$$

According to the definition of each parameter vector of Eq. (17), Eq. (18) can be written in the form of linear regression,

$$y(k) = \varphi^T(k, \alpha) \boldsymbol{\theta} + v(k) \tag{19}$$

where $\varphi^T(k, \alpha)$ is the information vector, which is

$$\varphi^T(k, \alpha) = \begin{bmatrix} \Psi(k, \alpha) \\ \zeta(k, \alpha) \end{bmatrix} \in \mathbb{R}^n, n = n_a + c \times n_b \tag{20}$$

Among them, $\Psi(k, \alpha) = [\Delta^\alpha y(k-1), \dots, \Delta^\alpha y(k-n_a)]^T \in \mathbb{R}^{n_a}$;

$\zeta(k, \alpha) = [\zeta_1^T(k, \alpha), \zeta_2^T(k, \alpha), \dots, \zeta_c^T(k, \alpha)]^T \in \mathbb{R}^{c \times n_b}$;

$\zeta_j(k, \alpha) = [\Delta^\alpha \phi_j(u(k-1)), \dots, \Delta^\alpha \phi_j(u(k-n_b))]^T$, $j = 1, 2, \dots, c$

$\boldsymbol{\theta}$ is an unknown parameter vector, which is $\boldsymbol{\theta} = [\mathbf{a}, \mathbf{b}, w_2 \mathbf{b}, \dots, w_c \mathbf{b}]^T = [\mathbf{a}, \mathbf{w} \otimes \mathbf{b}]^T \in \mathbb{R}^n$. In the Eq, \otimes is the Kronecker product or direct product, if $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$, $\mathbf{B} = [b_{ij}] \in \mathbb{R}^{p \times q}$, $\mathbf{A} \otimes \mathbf{B} = [a_{ij} \mathbf{B}] \in \mathbb{R}^{(mp) \times (nq)}$.

For the fractional order neural-fuzzy Hammerstein model described in Fig. 1, the purpose of identification is to use sampled data $\{u(k), y(k)\}$, $k = 1, 2, \dots$, to determine the neural-fuzzy network parameters c_l , δ_l , w_l and the number of rules c of the nonlinear module, and the dynamic fractional order linear module parameters, including a_i , b_i and the fractional order α . To this end, this paper divides the identification process into two processes: one process first determines the antecedent parameters of the nonlinear module neuro-fuzzy network; the other process is to identify the parameters of the linear module, the fractional order and the conclusion parameters of the neuro-fuzzy network. In Sect. 4, the two identification processes will be described in detail.

4 Model parameter identification

This section describes the identification method of model parameters in the previous section. First, the antecedent parameters of the neuro-fuzzy network are determined. Secondly, the multi-innovation L-M identification algorithm adopted by the parameters of the dynamic linear module is described. Finally, the process of the entire identification algorithm is summarized.

4.1 Determining the antecedent parameters of the neural-fuzzy network

For the neuro-fuzzy network model, it is necessary to determine c_l , δ_l and the number of rules c , which is actually a nonlinear optimization process. In this paper, a self-learning clustering algorithm is used to

determine the antecedent parameters c_l , δ_l of the neuro-fuzzy network model and the number of rules c . The specific learning process is as follows:

- (1) Collect the input data $u(k)$, $k = 1, 2, \dots$, take the input data $u(1)$ as the first cluster, and set the cluster center $c_1 = u(1)$. At this time, the number of clusters $\bar{N} = 1$, and the number of data pairs belonging to the first category is $\bar{N}_1 = 1$.
- (2) For the k -th input data $u(k)$, calculate the similarity between the k -th training data and each cluster center $c_l (l = 1, 2, \dots, c)$ according to the similarity criterion, and find the cluster L with the greatest similarity, that is, find the class to which $u(k)$ belongs (fuzzy rules). Define the similarity criterion as follows:

$$S_L = \max_{1 \leq l \leq c} \sqrt{e^{-\|u(k) - c_l\|_2^2}} \quad (21)$$

- (3) Determine whether to add new clusters according to the following criteria.

If $S_L < S_0$ (S_0 represents a preset threshold), it indicates that the k -th input data does not belong to an existing cluster, and a new cluster must be established. Let $c_{\bar{N}+1} = u(k)$, and let $\bar{N} = \bar{N} + 1$, $\bar{N}_{\bar{N}} = 1$, ($\bar{N}_{\bar{N}}$ represents the number of data pairs belonging to the \bar{N} -th cluster).

If $S_L \geq S_0$, it indicates that the k -th input data belongs to the existing L -th cluster. Let $\bar{N}_L = \bar{N}_L + 1$ and adjust the center of the L -th cluster as follows:

$$c_L = c_L + \frac{\lambda}{\bar{N}_L + 1} (u(k) - c_L) \quad (22)$$

In Eq. (22), $\lambda \in [0, 1][0, 1]$. It can be seen from Eq. (15) that as more and more data points belong to this cluster, the adjustment rate $\frac{\lambda}{\bar{N}_L + 1}$ of the cluster center decreases. In general, the larger the parameter λ , the larger the update range of the cluster centers. The parameter λ cannot be too large, otherwise it will destroy the existing classification.

(4) Let $k = k + 1$, repeat the above steps (2) and (3) until all the input data are assigned to the corresponding clusters. Thus, the number of clusters (fuzzy rules) is \bar{N} , and the number of rules is finally determined to be $c = \bar{N}$. The width of the membership function is given as follows:

$$\delta_l = \min_{\substack{j=1,2,\dots,c \\ j \neq l}} \frac{\|c_l - c_j\|_2}{\rho} \quad (23)$$

In Eq. (23), ρ is the overlap coefficient, generally $1 \leq \rho \leq 2$.

Through the abovementioned self-learning clustering process, the input data can be clustered to obtain the parameters c_l , δ_l and the rule number c of the neuro-fuzzy antecedent structure. Through self-learning clustering to determine the centers and widths of various types, the fuzzy basis function $\phi_1(u(k)), \dots, \phi_c(u(k))$ in Eq. (15) can be determined, the parameters a_i , b_j of the fractional order linear module, the fractional order α , and the conclusion parameter w_l of the neural-fuzzy network are completed by the multi-innovation L-M identification algorithm.

4.2 Estimation of dynamic linear module parameters

The estimation of the linear dynamic module parameters mainly determines the linear model parameters and the fractional order. This paper uses the L-M gradient descent algorithm to achieve. The L-M gradient descent algorithm is essentially a nonlinear optimization method, which includes Gauss–Newton optimization and gradient descent methods. However, the disadvantage of this method is that the accuracy is not high. For this reason, this paper applies the principle of multiple innovations to the L-M algorithm and proposes a multiple innovation L-M algorithm to solve the parameter identification problem of the dynamic linear module. The principle of multi-innovation identification has been applied in many references. The principle is to expand the scalar innovation into the multi-innovation of the vector and the information vector into the information matrix; therefore, both the previous sampling information and the current information are used in the identification algorithm, thereby improving the identification accuracy.

To this end, consider the following identification objective function,

$$J = \frac{1}{N} \sum_{k=1}^N e^2(k) \tag{24}$$

where N is the total number of samples; $e(k)$ is the prediction error or innovation, which is

$$e(k) = y(k) - \hat{y}(k) = y(k) - \varphi^T(k, \hat{\alpha})\hat{\theta} \tag{25}$$

Among them, $\hat{y}(k)$, $\hat{\theta}$ and $\hat{\alpha}$ are the output of the estimation model, the estimation of the parameter vector θ and the estimation of the fractional order α , respectively. If the objective function (24) has a minimum value, then there is the following iterative L-M algorithm:

$$\theta^{(i+1)} = \theta^{(i)} + \left\{ [J'' + \lambda I]^{-1} J' \right\}_{\hat{\theta}=\theta^{(i)}} \tag{26}$$

$$\alpha^{(i+1)} = \alpha^{(i)} + \left\{ [J'' + \lambda I]^{-1} J' \right\}_{\hat{\alpha}=\alpha^{(i)}} \tag{27}$$

Among them, λ is the adjustment operator that guarantees the convergence of the algorithm. Based on the prediction error of Eq. (25), the first derivative J'_θ and the second derivative J''_θ of the objective function with respect to the parameter θ are given as follows:

$$J'_\theta = -\frac{2}{N} \varphi(k, \hat{\alpha}) [y(k) - \varphi^T(k, \hat{\alpha})\hat{\theta}] \tag{28}$$

$$J''_\theta = \frac{2}{N} \varphi(k, \hat{\alpha}) \varphi^T(k, \hat{\alpha}) \tag{29}$$

Based on the prediction error of Eq. (25), the first derivative J'_α of the objective function with respect to the fractional order α is given as follows:

$$\begin{aligned} J'_\alpha &= -\frac{2}{N} \left[\frac{\partial \varphi^T(k, \hat{\alpha})\hat{\theta}}{\partial \hat{\alpha}} \right]^T [y(k) - \varphi^T(k, \hat{\alpha})\hat{\theta}] \\ &= -\frac{2}{N} \left[\frac{\partial \hat{y}(k)}{\partial \hat{\alpha}} \right]^T [y(k) - \varphi^T(k, \hat{\alpha})\hat{\theta}] \\ &= -\frac{2}{N} [\sigma \hat{y}(k)/\hat{\alpha}]^T [y(k) - \varphi^T(k, \hat{\alpha})\hat{\theta}] \end{aligned}$$

where $\sigma \hat{y}(k)/\hat{\alpha} = \frac{\partial \hat{y}(k)}{\partial \hat{\alpha}}$ is the sensitivity function with respect to the fractional order. The sensitivity function is calculated as follows:

$$\sigma \hat{y}(k)/\hat{\alpha} \approx \frac{\hat{y}(k, \hat{\alpha} + \delta \hat{\alpha}) - \hat{y}(k, \hat{\alpha})}{\delta \hat{\alpha}} \tag{30}$$

where $\delta \hat{\alpha}$ is the small change of the fractional order $\hat{\alpha}$.

Based on the prediction error of Eq. (25), the second derivative J''_α of the objective function with respect to the fractional order α is given as follows:

$$\begin{aligned} J''_\alpha &= \frac{2}{N} \left(\frac{\partial \hat{y}(k)}{\partial \hat{\alpha}} \right)^T \left(\frac{\partial \hat{y}(k)}{\partial \hat{\alpha}} \right) \\ &= \frac{2}{N} (\sigma \hat{y}(k)/\hat{\alpha})^T (\sigma \hat{y}(k)/\hat{\alpha}). \end{aligned} \tag{31}$$

Therefore, J'_θ , J''_θ , J'_α and J''_α are organized as:

$$\begin{bmatrix} J'_\theta \\ J'_\alpha \end{bmatrix} = -\frac{2}{N} \begin{bmatrix} \varphi(k, \hat{\alpha}) [y(k) - \varphi^T(k, \hat{\alpha})\hat{\theta}] \\ [\sigma \hat{y}(k)/\hat{\alpha}]^T [y(k) - \varphi^T(k, \hat{\alpha})\hat{\theta}] \end{bmatrix} \tag{32}$$

$$\begin{bmatrix} J''_\theta \\ J''_\alpha \end{bmatrix} = \frac{2}{N} \begin{bmatrix} \varphi(k, \hat{\alpha}) \varphi^T(k, \hat{\alpha}) \\ [\sigma \hat{y}(k)/\hat{\alpha}]^T [\sigma \hat{y}(k)/\hat{\alpha}] \end{bmatrix}. \tag{33}$$

In the algorithms (26) and (27), the single innovation $y(k) - \varphi^T(k, \hat{\alpha})\hat{\theta}$ is mainly used. For the L-M algorithm, its identification accuracy is not high. To this end, this paper proposes a multi-innovation L-M identification algorithm (MILM).

First, define the L -dimensional multi-innovation vector $\mathbf{E}(L, k, \hat{\alpha})$, the input and output information matrix $\hat{\Phi}(L, k, \hat{\alpha})$, and the stacked output vector $\mathbf{Y}(L, k)$ as follows:

$$\begin{aligned} \mathbf{E}(L, k, \hat{\alpha}) &= [y(k) - \varphi^T(k, \hat{\alpha})\hat{\theta}, y(k-1) - \varphi^T(k-1, \hat{\alpha})\hat{\theta}, \dots, \\ &\quad y(k-L+1) - \varphi^T(k-L+1, \hat{\alpha})\hat{\theta}]^T \in \mathbb{R}^L \end{aligned} \tag{34}$$

$$\begin{aligned} \hat{\Phi}(L, k, \hat{\alpha}) &= [\varphi(k, \hat{\alpha}), \varphi(k-1, \hat{\alpha}), \dots, \\ &\quad \varphi(k-L+1, \hat{\alpha})] \in \mathbb{R}^{n \times L} \end{aligned} \tag{35}$$

$$\begin{aligned} \mathbf{Y}(L, k) &= [y(k), y(k-1), \dots, y(k-L+1)]^T \in \mathbb{R}^L, \quad n \\ &= n_a + cn_b. \end{aligned} \tag{36}$$

The L -dimensional multi-innovation vector $\mathbf{E}(L, k, \hat{\alpha})$ can be expressed as

$$\mathbf{E}(L, k, \hat{\alpha}) = \mathbf{Y}(L, k) - \hat{\Phi}^T(L, k, \hat{\alpha})\hat{\theta}. \tag{37}$$

Define the stacking sensitivity function vector $\Xi(L, k, \hat{\alpha})$ as

$$\Xi(L, k, \hat{\alpha}) = [\sigma\hat{y}(k)/\hat{\alpha}, \sigma\hat{y}(k-1)/\hat{\alpha}, \dots, \sigma\hat{y}(k-L+1)/\hat{\alpha}]^T \in \mathbb{R}^L. \tag{38}$$

In this way, Eqs. (32) and (33) can be written as gradients with multiple innovations,

$$\begin{bmatrix} J''_{\hat{\theta}} \\ J''_{\hat{\alpha}} \end{bmatrix} = -\frac{2}{N} \begin{bmatrix} \Phi(L, k, \hat{\alpha}) [Y(k) - \Phi^T(L, k, \hat{\alpha})\hat{\theta}] \\ [\Xi(L, k, \hat{\alpha})]^T [Y(k) - \Phi^T(L, k, \hat{\alpha})\hat{\theta}] \end{bmatrix} \tag{39}$$

$$J''_{\hat{\alpha}} = \frac{2}{N} \begin{bmatrix} \Phi(L, k, \hat{\alpha})\Phi^T(L, k, \hat{\alpha}) \\ [\Xi(L, k, \hat{\alpha})]^T[\Xi(L, k, \hat{\alpha})] \end{bmatrix}. \tag{40}$$

After obtaining the estimated parameter vector $\hat{\theta}$, the first n_a element in $\hat{\theta}$ is the estimate of the vector $\hat{\mathbf{a}}$, and the element n_a+1 to the element n_a+n_b in $\hat{\theta}$ is the estimate of the vector $\hat{\mathbf{b}}$. By observing the parameter vector $\hat{\theta}$, there are n_b valuations of \hat{w}_j . For this reason, they are calculated by the average method,

$$\hat{w}_j = \frac{1}{n_b} \sum_{i=1}^{n_b} \frac{\theta_{n_a+(j-1)n_b+i}}{\hat{b}_i}, j = 2, 3, \dots, c. \tag{41}$$

In this paper, a hybrid method based on fuzzy clustering to determine the structure and network parameters of the neuro-fuzzy network and using multiple innovations to estimate the parameters of the linear module is proposed. The former clustering method is the premise and foundation of the hybrid algorithm, and the latter multi-innovation method is the key to the hybrid identification algorithm. The following summarizes the hybrid algorithm as follows:

- Step 1 Set $m = 1, k = 1, c = 1$, set the initial parameter values $\hat{\theta}^0, \hat{\alpha}^0, p$ and $\delta\hat{\alpha}$;
- Step 2 Collect input data $\{u(k)\}, k = 1, 2, \dots, N$, and determine the neuro-fuzzy antecedent parameters $\hat{c}_i, \hat{\delta}_i$ and the number of rules \hat{c} according to self-learning clustering;
- Step 3 According to the sampled data $\{u(k), y(k)\}, k = 1, 2, \dots, N$, perform the identification of the linear module parameter $\hat{\theta}$ and the identification of the neural-fuzzy network follow-up parameter \hat{w}_i ;

- Step 4 After obtaining the parameters of the neural-fuzzy network and the linear module, perform the identification of the fractional order $\hat{\alpha}$;
- Step 5 Let $k = k + 1$, if $s \leq P$, go to step 3, otherwise go to step 6;
- Step 6 If $\frac{|J(\hat{\theta}^{(m+1)}) - J(\hat{\theta}^{(m)})|}{|J(\hat{\theta}^{(m)})|} \leq \zeta$, then set $\hat{\theta} = \hat{\theta}^{(m)}, \hat{\alpha} = \alpha^{(m)}$ and $J(\hat{\theta}) = J(\hat{\theta}^{(m)})$, algorithm end. Otherwise $m = m + 1, s = 1$, go to step 3.

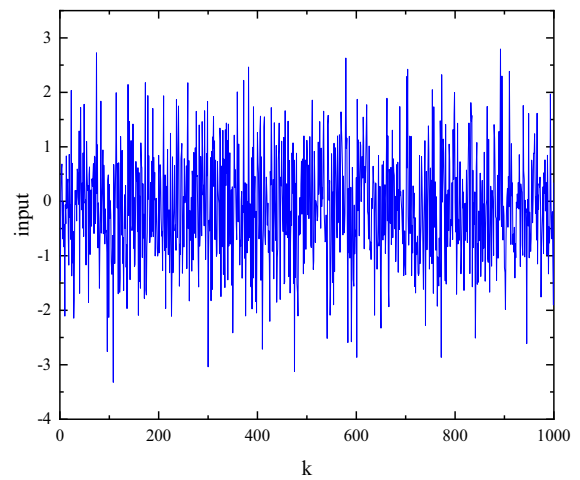


Fig. 3 System input signal

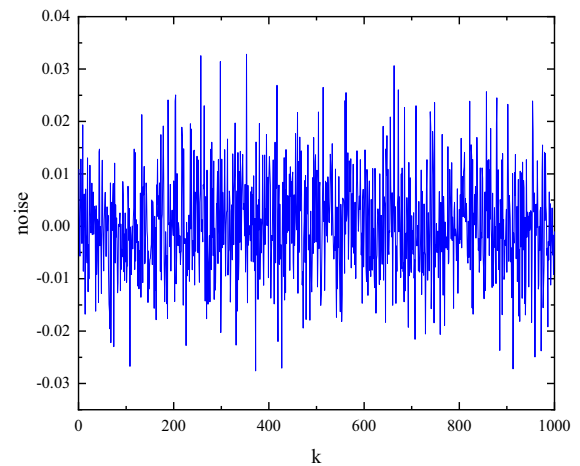


Fig. 4 System noise signal

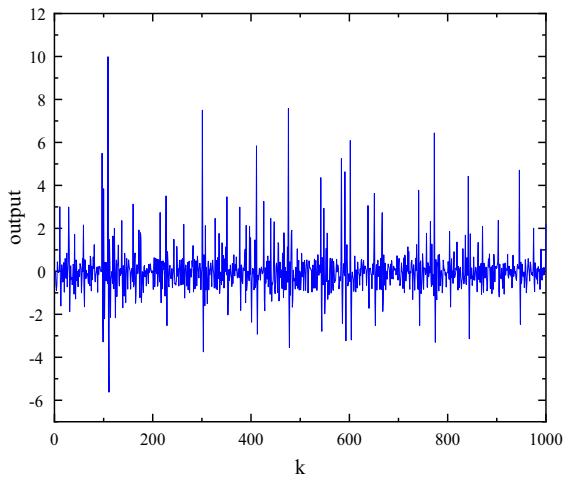


Fig. 5 The actual output signal of the system

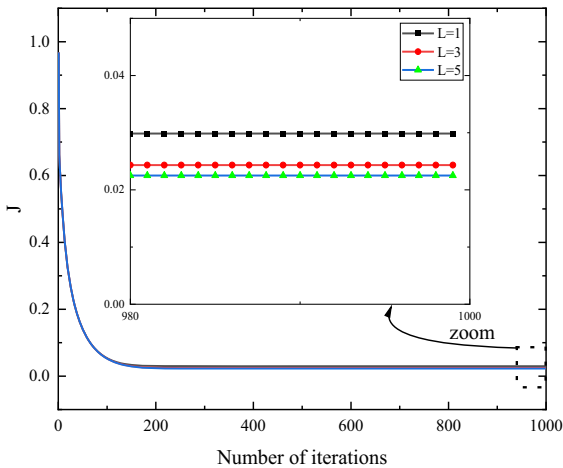


Fig. 6 Objective function J under different innovation lengths ($L = 1, 3, 5$)

5 Simulation examples

This section uses two examples to verify the effectiveness of the proposed method. The first example is an academic example of a discontinuous fractional order Hammerstein nonlinear system, which proves the theoretical feasibility of the method in this paper. Secondly, the method in this paper is applied to an actual system example of a flexible manipulator experiment, and the effect is compared with the results of reference [8], which proves that the method in this paper is more closely fitting with the actual system.

5.1 Academic example

In order to verify the effectiveness of the proposed algorithm, consider the Hammerstein fractional order nonlinear system with discontinuous functions:

$$\bar{u}(k) = \begin{cases} 2 - \cos(3u(k)) - \exp(-u(k)), & u(k) \leq 3.15 \\ 3, & u(k) > 3.15 \end{cases}$$

$$y(k) = (1 - A(z))y(k) + B(z)\bar{u}(k) + v(k)$$

where $u(k)$ and $y(k)$ are the system input and the system output, respectively. $A(z)$ and $B(z)$ are the denominator polynomial and the numerator polynomial, respectively,

$$A(z) = a_1z^{-\alpha} + a_2z^{-2\alpha},$$

$$B(z) = b_1z^{-\alpha} + b_2z^{-2\alpha},$$

$$\bar{u}(k) = \sum_{j=1}^c w_j\phi_j(u(k)).$$

The entire output of the system is given as follows:

Table 1 Estimated value of system parameters and fractional order ($L = 5$)

a_1	a_2	b_1	b_2	w_2b_1	w_2b_2	w_3b_1	w_3b_2	w_4b_1
- 0.03089	- 0.23291	- 0.43352	- 0.13023	- 0.61408	- 0.30428	0.272017	- 0.01413	- 0.99972

Table 2 Partial estimated value of system parameters and fractional order ($L = 5$)

w_4b_2	w_5b_1	w_5b_2	w_6b_1	w_6b_2	w_7b_1	w_7b_2	w_8b_1	w_8b_2
- 0.31443	- 0.03659	0.005282	- 1.55709	- 0.60673	- 0.32837	- 0.1862	- 2.30642	- 0.80718

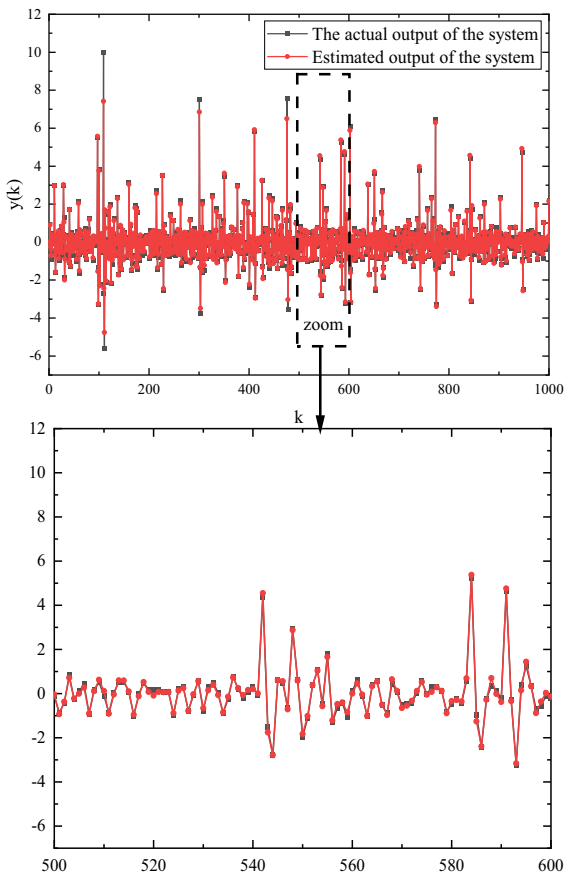


Fig. 7 The actual output and estimated output of the system ($L = 5$)

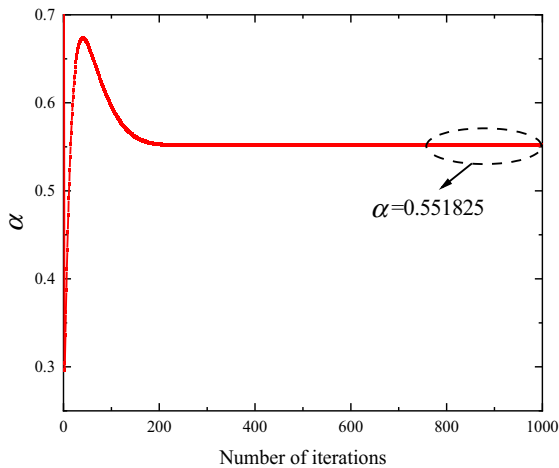


Fig. 8 Estimated value of system fractional order

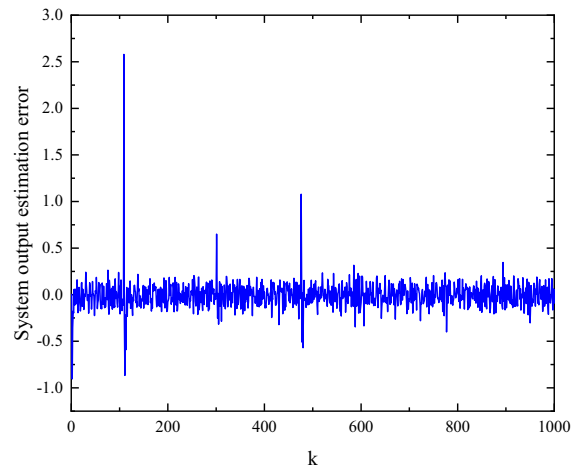


Fig. 9 System output estimation error

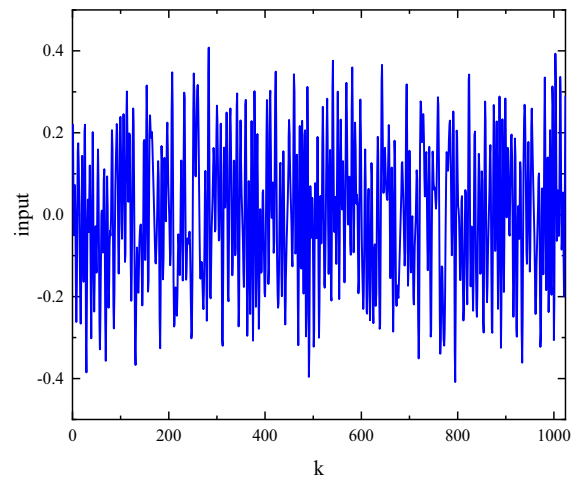


Fig. 10 System input signal

$$y(k) = \sum_{i=1}^2 a_i \Delta^\alpha y(k-i) + \sum_{j=1}^c w_j \sum_{i=1}^2 b_i \Delta^\alpha \phi_j(u(k-i)) + v(k)$$

Taking the fractional order $\alpha = 0.6$, the entire output of the system is given as follows:

$$y(k) = a_1 \Delta^{0.6} y(k-1) + a_2 \Delta^{0.6} y(k-2) + b_1 \Delta^{0.6} \phi(u(k-1)) + \dots + b_2 \Delta^{0.6} \phi(u(k-2)) + w_2 b_1 \Delta^{0.6} \phi(u(k-1)) + \dots + w_2 b_2 \Delta^{0.6} \phi(u(k-2)) + \dots + w_c b_1 \Delta^{0.6} \phi(u(k-1)) + \dots + w_c b_2 \Delta^{0.6} \phi(u(k-2)) + v(k)$$

The parameter vector is given as follows:

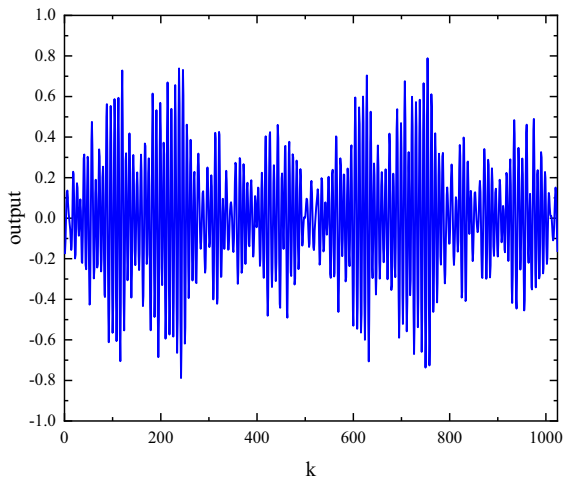


Fig. 11 The actual output signal of the system

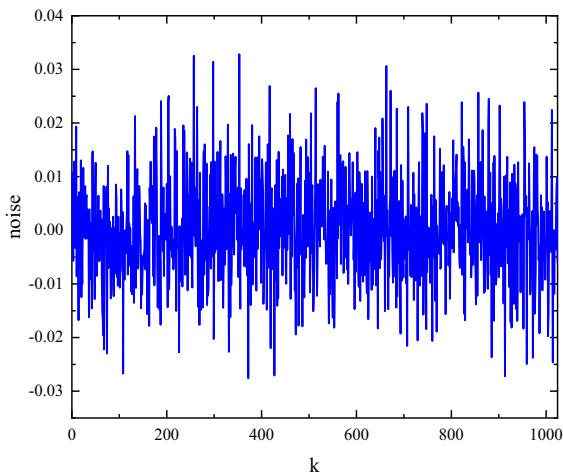


Fig. 12 System noise signal

$$\mathbf{a} = [a_1 \ a_2]^T = [0.1 \ 0.2]^T, \quad \mathbf{b} = [b_1 \ b_2]^T = [-0.4 \ -0.2]^T.$$

In the simulation, the input signal is a random signal with zero mean and unit variance, and the number of sampling is 1000, as shown in Fig. 3. The noise signal is an independent random signal with zero mean value and variance $\sigma^2=0.01$. Figure 4 shows the noise signal with 1000 sampling times. In the research, the length of multi-innovation $L = 1, 3, 5$ is selected. The neural-fuzzy network is used to approximate the static nonlinear system, and the self-learning fuzzy clustering is used to learn the antecedent structure. The initial value of the rule number is $c = 2$ and each rule adopts

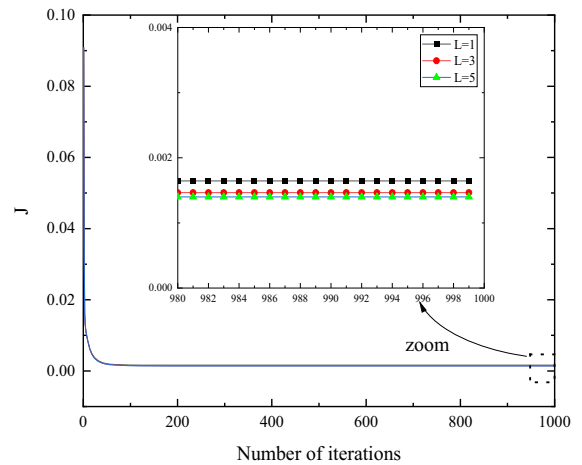


Fig. 13 Objective function J under different innovation lengths ($L = 1, 3, 5$)

an exponential clustering function. In order to verify the effectiveness of the proposed method, the objective function J in Eq. (24) is used to verify.

Figure 5 shows the actual output signal of the system; Fig. 6 is the objective functions J of different multiple innovations. Through comparison and analysis, it can be seen that multiple innovations are more advantageous than single innovations. As the information length increases, the objective function J gradually decreases, and the identification accuracy is significantly improved. Using self-organizing learning fuzzy clustering method, the number of system rules is 8, so the parameter vector to be estimated of the system is $\theta = [\mathbf{a}, \mathbf{b}, w_2 \mathbf{b}, \dots, w_8 \mathbf{b}]^T \in \mathbb{R}^n$, and the estimated values of system parameters are shown in Tables 1, 2. (Note: The specific value is calculated by Eq. (41), for example: $\hat{w}_2 = \left(\frac{w_2 b_1}{b_1} + \frac{w_2 b_2}{b_2} \right) / 2$). Figure 7 is the comparison result between the output of the identification model and the actual output of the system when the multi-innovation length $L = 5$, and the figure is partially. Fig. 8 is the system fractional order estimation curve, and finally converges to $\alpha = 0.551825$; Fig. 9 is the system output estimation error curve. Through analysis and comparison, the output of the system identification model and the actual output of the system have a good fitting effect and basically coincide.

Table 3 Estimated value of system parameters and fractional order ($L = 5$)

a_1	a_2	b_1	b_2	w_2b_1	w_2b_2
0.985457	0.326518	- 0.06353	- 0.19711	0.0956	0.016165

Table 4 Partial estimated value of system parameters and fractional order ($L = 5$)

w_3b_1	w_3b_2	w_4b_1	w_4b_2	w_5b_1	w_5b_2
- 0.08677	- 0.22514	0.127764	0.016702	0.200552	0.112353

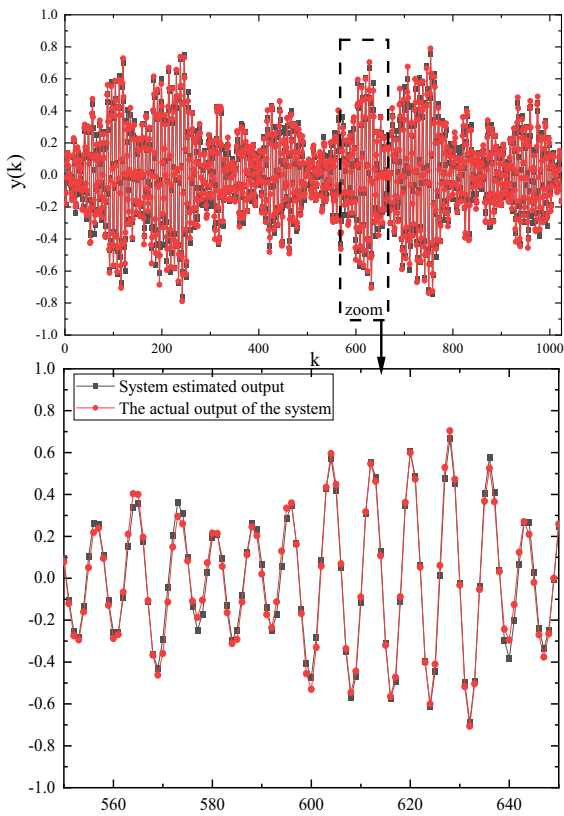


Fig. 14 The actual output and estimated output of the system ($L = 5$)

5.2 Example of a flexible manipulator

In order to further show the effectiveness of the proposed method, this paper also applies the proposed method to a flexible robotic arm system. The system consists of a flexible arm mounted on a motor. The input is the reaction torque of the structure on the

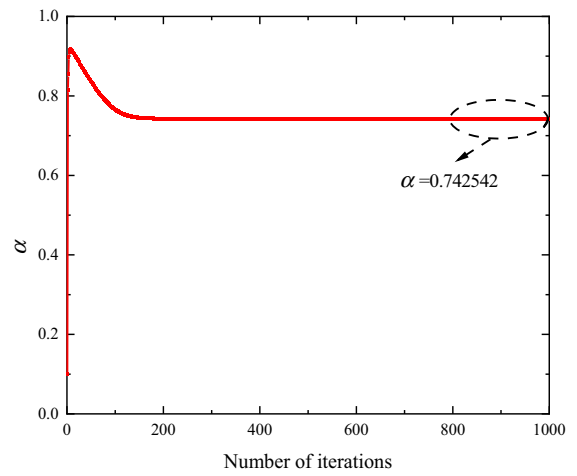


Fig. 15 Estimated value of system fractional order

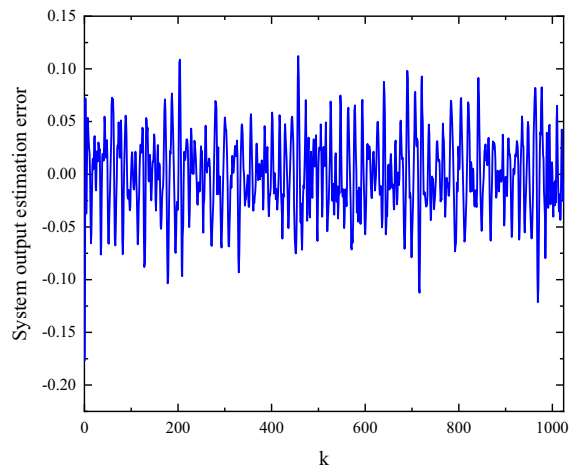


Fig. 16 System output estimation error

ground, and the output is the acceleration of the flexible arm. The benchmark data set of the flexible

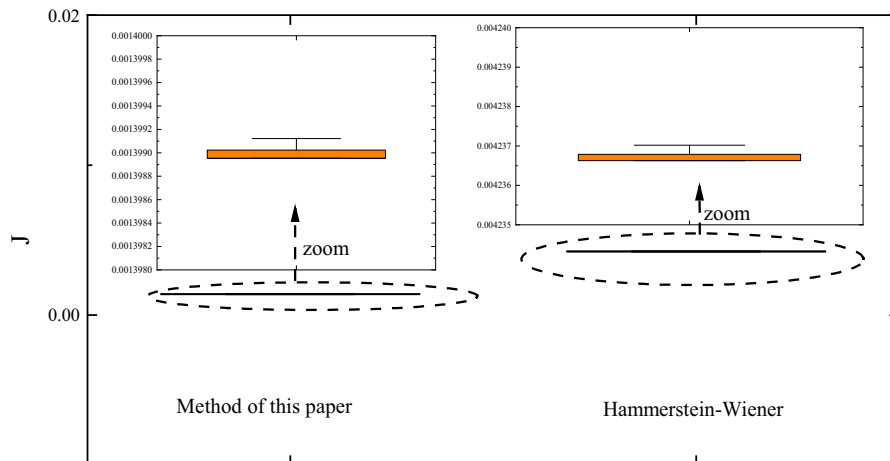


Fig. 17 Objective function J under different methods

manipulator obtained by DAISY database [28] (system identification database) is further verified. The measured data set contains 1024 input and output samples as shown in Figs. 10, 11. The noise signal is an independent random signal with zero mean and variance $\sigma^2 = 0.01$, and Fig. 12 shows the noise signal with 1024 sampling times.

In the research, select the multi-innovation length $L = 1, 3, 5$. Figure 13 is the objective functions of different multiple innovations. Through comparison and analysis, it can be seen that multiple innovations have more advantages than single innovations. Obviously, as the length of multiple innovations increases, the objective function J gradually decreases. The recognition accuracy is significantly improved. Using self-organizing learning fuzzy clustering method, the number of system rules is 5, so the system parameter vector to be estimated is $\theta = [\mathbf{a}, \mathbf{b}, w_2 \mathbf{b}, \dots, w_5 \mathbf{b}]^T \in \mathbb{R}^n$, and the system parameter estimated value is shown in Tables 3, 4. (Note: The specific value is calculated by Eq. (41), for example: $\hat{w}_2 = (\frac{w_2 b_1}{b_1} + \frac{w_2 b_2}{b_2})/2$). Figure 14 shows the comparison result between the output of the identification model and the actual output of the system when the multi-innovation length $L = 5$, and it is partially enlarged; Fig. 15 is the system fractional order estimation curve, which finally converges to $\alpha = 0.742542$; Fig. 16 is the system output estimation error curve. Through analysis and comparison, the output of the system identification model and the

actual output of the system have a good fitting effect and basically coincide.

In order to show the effectiveness of the proposed method, the method in this paper is also compared with the single innovation L-M algorithm proposed in the reference [8]. Figure 17 shows the objective function J curve between the method in this paper ($L = 5$) and the reference method. Through comparative analysis and comparison, it is obvious that the method in this paper has small errors and high identification accuracy.

6 Conclusion

Aiming at the difficulty in identifying fractional order Hammerstein nonlinear systems, a MILM hybrid identification method based on neuro-fuzzy principle is proposed. The static nonlinear module of the Hammerstein model is fitted by neuro-fuzzy, the self-learning clustering algorithm determines the antecedent parameters and the number of rules of the neuro-fuzzy network model, and the multi-innovation Levenberg–Marquardt algorithm is designed to estimate the system parameters and the fractional order of the system. Through simulation verification, proper use of multiple innovations in the algorithm can improve the identification accuracy. But for the self-learning method, the selection of similarity is empirical, and this will also become our research focus in the future. The modeling research of multi-input multi-output fractional order nonlinear systems is still a

challenging topic, which is also the research direction of researchers in the future.

Funding This work was supported by the National Natural Science Foundation of China (Grant number 61863034) and the Autonomous Region Graduate Research and Innovation Project (Grant number XJ2019G060).

Data availability At present, this paper completely describes the theoretical research and does not analyze the data set during the research period. The collection of data set is random according to different readers, but some codes can be provided by contacting the corresponding author.

Declarations

Conflict of interest The authors declare that there is no conflict of interest with respect to the research, authorship, and/or publication of this paper.

References

1. Abu Arqub, O.: Solutions of time-fractional Tricomi and Keldysh equations of Dirichlet functions types in Hilbert space. *Numer. Methods Partial Differ.* **34**(5), 1759–1780 (2018)
2. Del-Castillo-Negrete, D., Carreras, B.A., Lynch, V.E.: Front dynamics in reaction-diffusion systems with Levy flights: a fractional diffusion approach. *Phys. Rev. Lett.* **91**(1), 018302 (2003)
3. Written, T.A.: *Insights from Soft Condensed Matter*. Springer, Berlin (1999)
4. Dingyu, X.: *Fractional Calculus and Fractional Control*. Science Press, Beijing (2018)
5. Cheng, J.: *Theory of Fractional Difference Equations*. Xiamen University Press, Xiamen (2011)
6. Victor, S., Malti, R.: Parameter and differentiation order estimation in fractional models. *Automatica* **49**(4), 926–935 (2013)
7. Fahim, S.M., Ahmed, S., Imtiaz, S.A.: Fractional order model identification using the sinusoidal input. *ISA Trans.* **83**(1), 35–41 (2018)
8. Karima, H., Tounsia, D., Maamar, B.: Identification of Fractional Hammerstein systems with application to a heating process. *Nonlinear Dyn.* **2**, 10058 (2019)
9. Shalaby, R., Mohammad, E.H., Belal, A.Z.: Fractional order modeling and control for under-actuated inverted pendulum. *Commun. Nonlinear Sci. Numer. Simul.* (2019). <https://doi.org/10.1016/j.cnsns.2019.02.023>
10. Malti, R., Aoun, M., Oustaloup, A.: Systhesis of fractional kautz-like basis with two periodically repeating complex conjugate modes. In: *Control, Communication and Signal Processing, First International Symposium* (2004)
11. Jonscher, A.K.: Dielectric relaxation in solids. *J. Phys. D Appl. Phys.* **32**(14), 57–69 (1999)
12. Westerlund, S., Ekstam, L.: Capacitor theory. *IEEE Trans. Dielectrics Electr. Insulation.* **1**(5), 826–839 (1994)
13. Cheng, T., Zhishan, L., Juqiu, Z.: Nonlinear control of fractional boost converter in pseudo-continuous mode of inductor current. *Acta Physica Sinica.* **25**(1), 63–69 (2014)
14. Abuaisa, T., Kertzscher, J.: Fractional-order modeling and parameter identification of electrical coils. *Fract. Calc. Appl. Anal.* **22**(1), 193–216 (2019)
15. Boubaker, S.: Identification of nonlinear Hammerstein system using mixed integer-real coded particle swarm optimization: application to the electric daily peak-load forecasting. *Nonlinear Dyn.* **90**(2), 797–814 (2017)
16. Hu, Y.S., Fan, Y., Wei, Y.H., Wang, Y., Liang, Q.: Subspace-based continuous-time identification for fractional order systems from non-uniformly sampled data. *Int. J. Syst. Sci.* **47**(1), 122–134 (2016)
17. Wang, J., Wei, Y., Liu, T., Li, A., Wang, Y.: Fully parametric identification for continuous time fractional order Hammerstein systems. *J. Franklin Inst.* **357**, 651–666 (2020)
18. Ding, F., Chen, T.W.: Identification of Hammerstein nonlinear ARMAX systems. *Automatica* **41**(9), 1479–1489 (2005)
19. Gomez, J.C., Baeyens, E.: Subspace-based identification algorithms for Hammerstein and Wiener models. *Eur. J. Control.* **11**(2), 127–136 (2005)
20. Bai, E.W., Li, D.: Convergence of the iterative Hammerstein system identification algorithm. *IEEE Trans. Autom. Control* **49**(11), 1929–1940 (2004)
21. Narendra, K.S., Gallman, P.G.: An iterative method for the identification of nonlinear systems using a Hammerstein model. *IEEE Trans. Autom. Control* **11**(3), 546–550 (1966)
22. Chen, H.F.: Pathwise convergence of recursive identification algorithms for Hammerstein systems. *IEEE Trans. Autom. Control* **49**(10), 1641–1649 (2004)
23. Jing, X.J.: Frequency domain analysis and identification of blockoriented nonlinear systems. *J. Sound Vib.* **330**(22), 5427–5442 (2011)
24. Yu, C.P., Zhang, C.S., Xie, L.H.: Blind identification of non-minimum phase ARMA systems. *Automatica* **49**(6), 1846–1854 (2013)
25. Zennir, Y., Elhadi, G., Riad, B.: Robust fractional multi-controller design of inverted pendulum system. In: *2016 20th International Conference on System Theory, Control and Computing (ICSTCC)*. IEEE. (2016)
26. Caputo, M.: Linear models of dissipation whose Q is almost frequency independent. Part II. *Geophys. J. Int.* **13**(5), 529–539 (1967)
27. Monje, C.A., Chen, Y.Q., Vinagre, B.M., et al.: *Fractional-Order Systems and Controls: Fundamentals and Applications*. Springer, Berlin (2010)
28. De Moor, B., De Gerssem, P., De Schutter, B., et al.: DAISY: a database for identification of systems. *JOURNAL A.* **38**, 4–5 (1997)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.