



# Hybrid identification method for fractional-order nonlinear systems based on the multi-innovation principle

Zhang Qian<sup>1</sup> · Wang Hongwei<sup>1,2</sup> · Liu Chunlei<sup>1</sup>

Accepted: 26 October 2022 / Published online: 25 November 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

Aiming at the identification problems arising for fractional-order Hammerstein-Wiener system parameter coupling, namely, the difficulty of estimating the fractional order, low algorithm accuracy and slow convergence, an alternate identification method based on the principle of multiple innovations is proposed. First, a discrete model of a fractional-order Hammerstein-Wiener system is constructed. Second, an information matrix composed of fractional-order variables is used as the system input, combined with the multi-innovation principle, and the multi-innovation recursive gradient descent algorithm and the multi-innovation Levenberg-Marquardt algorithm are used to alternately estimate the parameters and fractional order of the model. The algorithms are executed cyclically and alternately presuppose each other. Finally, the convergence of the overall algorithm is theoretically analyzed, and the fractional-order Hammerstein-Wiener nonlinear system model is used to carry out numerical simulation experiments to verify the effectiveness of the algorithm. Moreover, we apply the proposed algorithm to an actual flexible manipulator system and perform fractional-order modeling and identification with high accuracy. Compared with the methods proposed by other scholars, the method proposed in this paper is more effective.

**Keywords** Fractional-order nonlinear system · System identification · Convergence analysis · Multi-innovation principle

## 1 Introduction

Fractional calculus has become ubiquitous in an increasing number of physical systems and industrial processes, such as viscous materials [1], fluid mechanics [2], and advanced materials [3]. Many complex physical and industrial systems are difficult to explain with traditional integer calculus, but the introduction of fractional calculus can largely overcome this difficulty. Therefore, fractional-order systems have absolute advantages in modeling, identification and control compared to integer order systems [4–6].

However, a known system model and parameters are prerequisites for a system to be accurately controlled. Therefore, before controlling the system, we first need to obtain the model structure and parameters of the system by means of system identification. Modeling and identification are particularly

important for applications involving practical systems, such as batteries; hence, many scholars have explored and comprehensively summarized various modeling and identification methods for lithium-ion batteries [7, 8]. Actual systems often exhibit strong nonlinear characteristics, and block structure models such as the Hammerstein, Wiener, and Hammerstein-Wiener models connect the nonlinear dynamic part and the linear static part of a system to each other, so that the nonlinear characteristics of the system can be described.

Regarding the modeling and identification of fractional-order systems, a large number of researchers have addressed these topics [9–12]. Zhang Qian and other scholars studied the identification of fractional-order systems with colored noise, combined the multi-innovation principle with the Levenberg-Marquardt algorithm, and successfully applied their results to a fractional-order Hammerstein system [13]. Later, on the basis of this previous research, these researchers adopted the fractional-order Hammerstein method of separation identification, and used a neuro fuzzy model to fit the nonlinear part of the Hammerstein model, thus converting the identification of the nonlinear system into a completely linear problem [14]. Rahmani et al. combined the Lyapunov method with a linear optimization algorithm and applied it for the modeling and identification of a fractional-order Hammerstein model [15].

✉ Wang Hongwei  
wq272427yrx@163.com

<sup>1</sup> School of Electrical Engineering, Xinjiang University, Urumqi 830047, People's Republic of China

<sup>2</sup> School of Control Science and Engineering, Dalian University of Technology, Dalian 110024, People's Republic of China

At the same time, other scholars have also combined intelligent optimization algorithms with traditional classical identification algorithms for the identification of fractional-order systems [16–18]. Hammar et al. applied the particle swarm optimization algorithm for the parameter identification of the state space model of a fractional-order Hammerstein system [19]. Sersour et al. then extended this particle swarm optimization algorithm to propose an adaptive velocity particle swarm optimization algorithm, which can successfully identify fractional-order discrete Wiener systems [20].

Most of the literature has focused on fractional-order Hammerstein and Wiener systems. Compared with a Hammerstein-Wiener system, these two types of systems have simpler structures, and it is more difficult to fully express the strong nonlinearity of more complex systems. Therefore, it is very important to explore an identification method suitable for fractional-order Hammerstein-Wiener systems.

Through a comprehensive analysis of the identification methods for the transfer functions of fractional-order systems in the above literature, the main problems encountered in system identification are identified as follows: 1) There are many parameters and variables to be identified, and they are coupled with each other. 2) The estimation of fractional orders is difficult. 3) The algorithms converge slowly. 4) The algorithms may not always successfully converge. Therefore, this paper proposes a hybrid parameter identification algorithm based on the multi-innovation principle. Previous scholars [21–24] have proposed a multi-innovation least-mean-square algorithm based on the principle of multi-innovation identification. The basic principle of the multi-innovation identification method is to expand the scalar innovation into a multi-innovation vector, and the innovation vector into an innovation matrix, allowing both the current data and past data to be used. In this paper, it is shown that this multi-innovation series identification algorithm can achieve improved convergence and improved accuracy of parameter estimation; therefore, this paper introduces the multi-innovation principle into the proposed identification algorithm. The main contributions of this paper are as follows: 1) A fractional-order discrete Hammerstein-Wiener system model is constructed. 2) A multi-innovation recursive gradient descent algorithm is designed to estimate the model parameters with an information vector composed of fractional-order variables as input, thereby adapting the Levenberg-Marquardt algorithm to estimate the fractional order. The modeling efficiency is improved by using two multi-innovation algorithms interactively. 3) Based on the convergence theorem, the performance of the proposed algorithm is analyzed.

The main structure of this paper is as follows: Section 2 introduces background knowledge on fractional calculus and describes the fractional-order Hammerstein-Wiener system model. Section 3 introduces the multi-innovation gradient descent algorithm used to estimate the parameters of fractional-

order nonlinear models. Section 4 proposes the multi-innovation Levenberg-Marquardt algorithm to estimate the fractional orders of nonlinear systems. Section 5 applies the convergence theorem to analyze the performance of the proposed algorithm. Section 6 presents a simulation case study. Finally, a summary of this research and an outlook on future research are given.

## 2 Problem formulation and preliminary

### 2.1 Fractional calculus

From different perspectives, researchers have obtained several common forms of fractional calculus operators, including the Grünwald-Letnikov (GL) fractional operators [25], the Riemann-Liouville (RL) fractional operators [26] and the Caputo fractional operators [27]. The GL-type fractional operator for a discrete system that is used in this paper is defined as follows:

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

where  $0 < \bar{\alpha} < 1$  is the fractional order;  $k$  and  $h$  represent the number of sampling times and the sampling time, respectively; and  $\binom{\bar{\alpha}}{j}$  is defined as follows:

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

This can be written in recursive form as:

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

where  $\beta(j) = (-1)^j \binom{\bar{\alpha}}{j}$ . To facilitate simulation and concise expression, Eq. (1) can be written as follows according to Eqs. (2) and (3):

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

Under the assumption that the system sampling time is  $h = 1$ , Eq. (4) can be organized into the following equation:

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

In this paper, the fractional calculus operator expressed in Eq. (5) will be used.

### 2.2 Fractional-order linear model

For fractional-order systems, there are different linear model descriptions. This paper considers the following linear discrete transfer function model:

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

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

Model (6) can be written in the following form:

$$\sum_{i=1}^{n_a} a_i z^{-\alpha_i} y(k) = \sum_{i=1}^{n_b} b_i z^{-\gamma_i} u(k) \tag{7}$$

When the fractional orders of the denominator and numerator polynomials in (7) are completely different, the fractional-order model is called a nonhomogeneous system; when the fractional order of each polynomial is of the form  $\alpha_i = i - \bar{\alpha}$ ,  $\gamma_j = j - \bar{\alpha}$  ( $i = 1, 2, \dots, n_a; j = 1, 2, \dots, n_b$ ) ( $\bar{\alpha}$  is the order factor), the model is a same-dimensional system. The fractional-order same-dimensional case is considered in this paper.

The regression equation for Eq. (7) can be written as:

$$\sum_{i=1}^{n_a} a_i z^{-i+\bar{\alpha}} y(k) = \sum_{i=1}^{n_b} b_i z^{-i+\bar{\alpha}} u(k) \tag{8}$$

By introducing the fractional backward operator  $z^{-i+\bar{\alpha}}x(t) = \Delta^{\bar{\alpha}}x(t - i)$ , Eq. (7) can be written in the following form by means of the discrete fractional operator  $\Delta$ :

$$\sum_{i=1}^{n_a} a_i \Delta^{\bar{\alpha}} y(k-i) = \sum_{i=1}^{n_b} b_i \Delta^{\bar{\alpha}} u(k-i) \tag{9}$$

We will use Eq. (9) to describe the linear part of the fractional-order Hammerstein-Wiener model.

### 2.3 Problem description

In a block-structure nonlinear model, the dynamic linear and static nonlinear blocks are connected in series, in parallel or in a feedback structure. Such a model can well describe the

nonlinear system of an actual process. The general structure of a Hammerstein-Wiener system is shown in Fig. 1, where a linear dynamic module is surrounded by two static nonlinear modules at its input and output.

Specifically, the input-output relationship of the Hammerstein-Wiener system can be expressed as:

$$y(k) = A(z)g(y(k)) + B(z)f(u(k)) + v(k) \tag{10}$$

In this equation,  $u(k)$  and  $y(k)$  are the input and output of the system, respectively, and  $v(k)$  is the external noise of the system. The nonlinear links can be represented by two static nonlinear functions  $f(\cdot)$  and  $g(\cdot)$ . The linear elements are represented by the polynomials  $A(z)$  and  $B(z)$  containing the backshift operator  $z^{-1}$ , i.e.,

$$A(z) = a_1z^{-\alpha_1} + a_2z^{-\alpha_2} + \dots + a_{n_a}z^{-\alpha_{n_a}} = \sum_{i=1}^{n_a} a_i z^{-\alpha_i}$$

$$B(z) = b_1z^{-\gamma_1} + b_2z^{-\gamma_2} + \dots + b_{n_b}z^{-\gamma_{n_b}} = \sum_{i=1}^{n_b} b_i z^{-\gamma_i} \tag{11}$$

The two static nonlinear functions  $f(\cdot)$  and  $g(\cdot)$  are nonlinear functions composed of several known basis functions, as follows:

$$f(u(k)) = p_1f_1(u(k)) + p_2f_2(u(k)) + \dots + p_{n_p}f_{n_p}(u(k))$$

$$= \sum_{i=1}^{n_p} p_i f_i(u(k)) \tag{12}$$

$$g(y(k)) = q_1g_1(y(k)) + q_2g_2(y(k)) + \dots + q_{n_q}g_{n_q}(y(k))$$

$$= \sum_{i=1}^{n_q} q_i g_i(y(k)) \tag{13}$$

where  $f_1(\cdot), \dots, f_{n_p}(\cdot)$  are  $n_p$  known basis functions and  $g_1(\cdot), \dots, g_{n_q}(\cdot)$  are  $n_q$  known basis functions.

Substituting the above two equations into Eq. (10), we obtain:

$$y(k) = \sum_{i=1}^{n_a} a_i z^{-\alpha_i} g(y(k)) + \sum_{i=1}^{n_b} b_i z^{-\gamma_i} f(u(k)) + v(k) \tag{14}$$

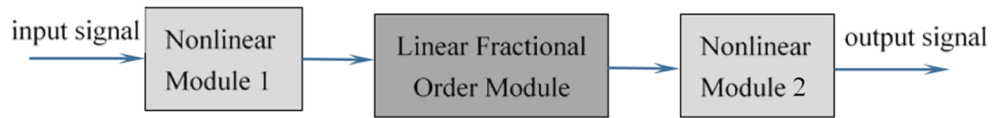
Substituting Eqs. (12) and (13) into Eq. (14) yields the following expression for the description of the entire system:

$$y(k) = \sum_{i=1}^{n_a} a_i z^{-\alpha_i} \sum_{j=1}^{n_q} q_j g_j(y(k)) + \sum_{i=1}^{n_b} b_i z^{-\gamma_i} \sum_{j=1}^{n_p} p_j f_j(u(k))$$

$$+ v(k) \tag{15}$$

In this paper, the fractional-order discrete system is considered to be a symmetric system, that is,  $\alpha_i = i - \bar{\alpha}, \gamma_j = j - \bar{\alpha}$ ,

**Fig. 1** Hammerstein-Wiener nonlinear system structure diagram



with  $\Delta$  representing the differential operator. Thus, Eq. (15) can be written as:

$$\begin{aligned}
 y(k) &= \sum_{i=1}^{n_a} a_i \sum_{j=1}^{n_q} q_j \Delta^{\bar{\alpha}} g_j(y(k-i)) + \sum_{i=1}^{n_b} b_i \sum_{j=1}^{n_p} p_j \Delta^{\bar{\alpha}} f_j(u(k-i)) + v(k) \\
 &= q_1 \sum_{i=1}^{n_a} a_i \Delta^{\bar{\alpha}} g_1(y(k-i)) + \dots + q_{n_q} \sum_{i=1}^{n_a} a_i \Delta^{\bar{\alpha}} g_{n_q}(y(k-i)) \\
 &\quad + p_1 \sum_{i=1}^{n_b} b_i \Delta^{\bar{\alpha}} f_1(u(k-i)) + \dots + p_{n_p} \sum_{i=1}^{n_b} b_i \Delta^{\bar{\alpha}} f_{n_p}(u(k-i)) + v(k)
 \end{aligned} \tag{16}$$

The parameter vectors are defined as:

$$\begin{aligned}
 \mathbf{a} &= [a_1 \in \mathbb{R}^{n_a}, \mathbf{b} = [b_1 \in \mathbb{R}^{n_b}, \\
 \mathbf{p} &= [p_1 \in \mathbb{R}^{n_p}, \mathbf{q} = [q_1 \in \mathbb{R}^{n_q}
 \end{aligned} \tag{17}$$

To obtain unique model parameters, the parameters of the model are normalized. For this purpose, the first coefficients of the two nonlinear modules are fixed; that is, the first elements in the parameter vectors  $\mathbf{p}$  and  $\mathbf{q}$  take values of 1,  $p_1 = 1$  and  $q_1 = 1$ . On this basis, Eq. (16) can be rewritten as:

$$\begin{aligned}
 y(k) &= \sum_{i=1}^{n_a} a_i \Delta^{\bar{\alpha}} g_1(y(k-i)) + q_2 \sum_{i=1}^{n_a} a_i \Delta^{\bar{\alpha}} g_2(y(k-i)) + \\
 &\quad \dots + q_{n_q} \sum_{i=1}^{n_a} a_i \Delta^{\bar{\alpha}} g_{n_q}(y(k-i)) + \sum_{i=1}^{n_b} b_i \Delta^{\bar{\alpha}} f_1(u(k-i)) + \\
 &\quad p_2 \sum_{i=1}^{n_b} b_i \Delta^{\bar{\alpha}} f_2(u(k-i)) + \dots + p_{n_p} \sum_{i=1}^{n_b} b_i \Delta^{\bar{\alpha}} f_{n_p}(u(k-i)) + v(k)
 \end{aligned} \tag{18}$$

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

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

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

$$\boldsymbol{\varphi}^T(k, \bar{\alpha}) = \begin{bmatrix} \boldsymbol{\psi}(k, \bar{\alpha}) \\ \boldsymbol{\zeta}(k, \bar{\alpha}) \end{bmatrix} \in \mathbb{R}^n, n = n_q \times n_a + n_p \times n_b \tag{20}$$

where where  $\boldsymbol{\psi}(k, \bar{\alpha}) = [\boldsymbol{\psi}_1^T(k, \bar{\alpha}), \boldsymbol{\psi}_2^T(k, \bar{\alpha}), \dots, \boldsymbol{\psi}_{n_q}^T(k, \bar{\alpha})]^T \in \mathbb{R}^{n_q \times n_a}$ ;

$$\boldsymbol{\psi}_i(k, \bar{\alpha}) = [\Delta^{\bar{\alpha}} g_i(y(k-1)), \dots, \Delta^{\bar{\alpha}} g_i(y(k-n_a))]^T,$$

$i = 1, 2, \dots, n_q$ ;

$$\boldsymbol{\zeta}(k, \bar{\alpha}) = [\boldsymbol{\zeta}_1^T(k, \bar{\alpha}), \boldsymbol{\zeta}_2^T(k, \bar{\alpha}), \dots, \boldsymbol{\zeta}_{n_p}^T(k, \bar{\alpha})]^T \in \mathbb{R}^{n_p \times n_b};$$

and

$$\boldsymbol{\zeta}_j(k, \bar{\alpha}) = [\Delta^{\bar{\alpha}} f_j(u(k-1)), \dots, \Delta^{\bar{\alpha}} f_j(u(k-n_b))]^T,$$

$\boldsymbol{\zeta}_j, j = 1, 2, \dots, n_p$ .

$\boldsymbol{\theta}$  is the unknown parameter vector, which is

$$\begin{aligned}
 \boldsymbol{\theta} &= [\mathbf{a}, q_2 \mathbf{a}, \dots, q_{n_q} \mathbf{a}, \mathbf{b}, p_2 \mathbf{b}, \dots, p_{n_p} \mathbf{b}]^T \\
 &= [\mathbf{q} \otimes \mathbf{a}, \mathbf{p} \otimes \mathbf{b}]^T \in \mathbb{R}^n,
 \end{aligned}$$

where  $\otimes$  is the Kronecker product or direct product, defined as follows: given  $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$  and  $\mathbf{B} = [b_{ij}] \in \mathbb{R}^{p \times q}$ ,  $\mathbf{A} \otimes \mathbf{B} = [a_{ij} \mathbf{B}] \in \mathbb{R}^{(mp) \times (nq)}$ .

In the following sections, an identification method is designed in accordance with Eq. (19) to estimate the unknown parameter vectors  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{p}$ ,  $\mathbf{q}$  and the fractional order  $\bar{\alpha}$  in a fractional-order model.

### 3 Model parameter identification based on the multi-innovation identification principle

To enable the estimation of the parameters of Model (19), an objective function is first given:

$$J(\boldsymbol{\theta}) = [y(k) - \boldsymbol{\varphi}^T(k, \bar{\alpha}) \boldsymbol{\theta}]^2 \tag{21}$$

According to the minimum value of the objective function (21), it is necessary to take the extreme value of the estimated parameter, and the following stochastic gradient descent algorithm can be obtained:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \mu(k)\varphi(k, \bar{\alpha}) \left[ y(k) - \varphi^T(k, \bar{\alpha}) \hat{\theta}(k-1) \right] \quad (22)$$

where  $\hat{\theta}(k)$  is the estimated parameter vector at the  $k$ th instance of sampling. Notably, the fractional order  $\bar{\alpha}$  in (22) is unknown, and the information vector  $\varphi(k, \bar{\alpha})$  cannot be obtained, meaning that algorithm (22) cannot be used. To overcome this problem, the real fractional order  $\bar{\alpha}$  is replaced with the fractional order estimate  $\hat{\alpha}$ ; then, by taking  $\mu(k) = 1/r(k)$ , with  $r(k) = r(k-1) + \|\hat{\varphi}(k, \hat{\alpha})\|^2$ , the gradient descent algorithm can be written as:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{\hat{\varphi}(k, \hat{\alpha})}{r(k)} \left[ y(k) - \hat{\varphi}^T(k, \hat{\alpha}) \hat{\theta}(k-1) \right] \quad (23)$$

$$r(k) = r(k-1) + \|\hat{\varphi}(k, \hat{\alpha})\|^2, r(0) = 1 \quad (24)$$

In this equation,  $\hat{\varphi}(k, \hat{\alpha})$  is the  $k$ th estimated information vector, which is:

$$\hat{\varphi}(k, \hat{\alpha}) = \begin{bmatrix} \hat{\psi}(k, \hat{\alpha}) \\ \hat{\zeta}(k, \hat{\alpha}) \end{bmatrix} \in R^n, n = n_q \times n_a + n_p \times n_b \quad (25)$$

where  $\psi(k, \hat{\alpha}) = [\hat{\psi}_1^T(k, \hat{\alpha}), \hat{\psi}_2^T(k, \hat{\alpha}), \dots, \hat{\psi}_{n_q}^T(k, \hat{\alpha})]^T \in R^{n_q \times n_a}$ ;

$$\hat{\psi}_i(k, \hat{\alpha}) = \left[ \Delta^{\hat{\alpha}} g_i(y(k-1)), \dots, \Delta^{\hat{\alpha}} g_i(y(k-n_a)) \right]^T, i = 1, 2, \dots, n_q;$$

$$\hat{\zeta}(k, \hat{\alpha}) = \left[ \hat{\zeta}_1^T(k, \hat{\alpha}), \hat{\zeta}_2^T(k, \hat{\alpha}), \dots, \hat{\zeta}_{n_p}^T(k, \hat{\alpha}) \right]^T; \text{ and}$$

$$\hat{\zeta}_j(k, \hat{\alpha}) = \left[ \Delta^{\hat{\alpha}} f_j(u(k-1)), \dots, \Delta^{\hat{\alpha}} f_j(u(k-n_b)) \right]^T, j = 1, 2, \dots, n_p.$$

$\hat{\theta}(k)$  is the  $k$ th estimated parameter vector,  $\hat{\theta} = [\hat{\mathbf{q}} \otimes \hat{\mathbf{a}}, \hat{\mathbf{p}} \otimes \hat{\mathbf{b}}]^T \in R^n$ .

For the gradient descent algorithm of Eq. (23), its main disadvantage is its slow convergence rate. To improve its performance, the information length is introduced, and the multi-innovation identification principle is adopted to improve the identification performance. The principle of multi-innovation identification is to enhance single-innovation correction technology by proposing a multi-innovation correction technique

in order to establish an identification method with multi-innovation correction, which can significantly improve the convergence speed of the identification algorithm. For identification, the single innovation  $e(k) = y(k) - \hat{\varphi}^T(k, \hat{\alpha}) \hat{\theta}(k-1)$  is expanded to a  $P$ -dimensional multi-innovation,

$$\mathbf{E}(P, k, \hat{\alpha}) = \left[ y(k) - \hat{\varphi}^T(k, \hat{\alpha}) \hat{\theta}(k-1), y(k-1) - \hat{\varphi}^T(k-1, \hat{\alpha}) \hat{\theta}(k-1), \dots, y(k-P+1) - \hat{\varphi}^T(k-P+1, \hat{\alpha}) \hat{\theta}(k-1) \right]^T \in R^P$$

The input-output innovation matrix  $\hat{\Phi}(P, k, \hat{\alpha})$  and the stacked output vector  $\mathbf{Y}(P, k)$  are defined as:

$$\hat{\Phi}(P, k, \hat{\alpha}) = \begin{bmatrix} \hat{\varphi}(k, \hat{\alpha}) \\ \hat{\varphi}(k-1, \hat{\alpha}) \\ \dots \\ \hat{\varphi}(k-P+1, \hat{\alpha}) \end{bmatrix} \in R^{n \times P} \quad (26a)$$

$$\mathbf{Y}(P, k) = [y(k), y(k-1), \dots, y(k-P+1)]^T \in R^P \quad (26b)$$

The  $P$ -dimensional multi-innovation error vector  $\mathbf{E}(P, k, \hat{\alpha})$  is expressed as:

$$\mathbf{E}(P, k, \hat{\alpha}) = \mathbf{Y}(P, k) - \hat{\Phi}^T(P, k, \hat{\alpha}) \hat{\theta}(k-1) \quad (27)$$

When the multi-innovation satisfies  $P = 1$ , because  $\hat{\Phi}(1, k, \hat{\alpha}) = \hat{\varphi}(k, \hat{\alpha})$  and  $\mathbf{E}(1, k, \hat{\alpha}) = y(k) - \hat{\varphi}^T(k, \hat{\alpha}) \hat{\theta}(k-1)$ , Eq. (23) can be equivalently expressed as:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{\hat{\Phi}(1, k, \hat{\alpha})}{r(k)} \mathbf{E}(1, k, \hat{\alpha}) \quad (28)$$

By replacing the 1-dimensional information vectors  $\hat{\Phi}(1, k, \hat{\alpha})$  and  $\mathbf{E}(1, k, \hat{\alpha})$  with a  $P$ -dimensional information matrix and multi-innovation vector and taking  $r(k) = r(k-1) + \|\hat{\Phi}(P, k, \hat{\alpha})\|^2$ , the fractional order estimate  $\hat{\alpha}$  can be obtained by means of the multi-innovation Levenberg-Marquardt algorithm discussed in Section 4. At this time, the multi-innovation recursive gradient descent algorithm (MIGD) is expressed as follows:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{\hat{\Phi}(P, k, \hat{\alpha})}{r(k)} \mathbf{E}(P, k, \hat{\alpha}) \quad (29)$$

$$\mathbf{E} \left( P, k, \widehat{\alpha} \right) = \mathbf{Y}(P, k) - \widehat{\Phi}^T \left( P, k, \widehat{\alpha} \right) \widehat{\theta}(k-1) \quad (30)$$

$$r(k) = r(k-1) + \left\| \widehat{\Phi} \left( P, k, \widehat{\alpha} \right) \right\|^2, r(0) = 1 \quad (31)$$

$$\begin{aligned} \widehat{\Phi} \left( P, k, \widehat{\alpha} \right) \\ = \left[ \widehat{\varphi} \left( k, \widehat{\alpha} \right), \widehat{\varphi} \left( k-1, \widehat{\alpha} \right), \dots, \widehat{\varphi} \left( k-P+1, \widehat{\alpha} \right) \right] \in \mathbb{R}^{n \times P} \end{aligned} \quad (32)$$

$$\mathbf{Y}(P, k) = [y(k), y(k-1), \dots, y(k-P+1)]^T \in \mathbb{R}^P \quad (33)$$

$$\begin{aligned} \widehat{\varphi}^T \left( k, \widehat{\alpha} \right) &= \begin{bmatrix} \widehat{\psi} \left( k, \widehat{\alpha} \right) \\ \widehat{\zeta} \left( k, \widehat{\alpha} \right) \end{bmatrix} \in \mathbb{R}^n, n \\ &= n_q \times n_a + n_p \times n_b \end{aligned} \quad (34)$$

$$\widehat{\psi} \left( k, \widehat{\alpha} \right) = \left[ \widehat{\psi}_1^T \left( k, \widehat{\alpha} \right), \widehat{\psi}_2^T \left( k, \widehat{\alpha} \right), \dots, \widehat{\psi}_{n_q}^T \left( k, \widehat{\alpha} \right) \right]^T \in \mathbb{R}^{n_q \times n_a} \quad (35)$$

$$\begin{aligned} \widehat{\psi}_i \left( k, \widehat{\alpha} \right) &= \left[ \Delta^{\widehat{\alpha}} g_i(y(k-1)), \dots, \Delta^{\widehat{\alpha}} g_i(y(k-n_a)) \right]^T, i \\ &= 1, 2, \dots, n_q \end{aligned} \quad (36)$$

$$\widehat{\zeta} \left( k, \widehat{\alpha} \right) = \left[ \widehat{\zeta}_1^T \left( k, \widehat{\alpha} \right), \widehat{\zeta}_2^T \left( k, \widehat{\alpha} \right), \dots, \widehat{\zeta}_{n_p}^T \left( k, \widehat{\alpha} \right) \right]^T \quad (37)$$

$$\begin{aligned} \widehat{\zeta}_j \left( k, \widehat{\alpha} \right) &= \left[ \Delta^{\widehat{\alpha}} f_j(u(k-1)), \dots, \Delta^{\widehat{\alpha}} f_j(u(k-n_b)) \right]^T, j \\ &= 1, 2, \dots, n_p \end{aligned} \quad (38)$$

For the initial values of the parameters to be identified,  $\widehat{\theta}(0) = \mathbf{1}_{n_0}/p_0$  is taken, where  $\mathbf{1}_{n_0}$  is a column vector consisting entirely of 1 s.

Once the estimated parameter vector  $\widehat{\theta}$  has been obtained, the first  $n_a$  elements in  $\widehat{\theta}$  are estimates of the vector  $\widehat{\mathbf{a}}$ , and the  $(n_a n_q + 1)$ th element to  $(n_a n_q + n_b)$ th elements in  $\widehat{\theta}$  are estimates of the vector  $\widehat{\mathbf{b}}$ . In the parameter vector  $\widehat{\theta}$ , there are  $n_a$  estimates of  $\widehat{q}_j$  and  $n_b$  estimates of  $\widehat{p}_j$ . The final estimates are calculated by averaging:

$$\widehat{q}_j = \frac{1}{n_a} \sum_{i=1}^{n_a} \frac{\theta_{n_a(j-1)+i}}{\widehat{a}_i}, \quad j = 2, 3, \dots, n_q \quad (39)$$

$$\widehat{p}_j = \frac{1}{n_b} \sum_{i=1}^{n_b} \frac{\theta_{n_a n_q + (j-1)n_b + i}}{\widehat{b}_i}, \quad j = 2, 3, \dots, n_p \quad (40)$$

## 4 An estimation method for the fractional order $\widehat{\alpha}$ based on the multi-innovation principle

For the identification of a fractional-order system, the main parameters to be identified are the fractional order  $\widehat{\alpha}$  and the model parameters  $\{a_i, b_i, p_i, q_i\}$ . Fractional order estimation and model parameter identification are two stages in the system identification process. The identification results provide the initial conditions for another phase of the algorithm to proceed. For the model parameter identification algorithm, the estimation of the fractional order provides the conditions for identification; for the fractional order estimation algorithm, the identification of the model parameters provides the initial premise for estimation. Figure 2 shows a description of the interactive identification process of the two multi-innovation identification methods.

On this basis, the design of the fractional order estimation algorithm is the key step for the success of the whole algorithm. According to the objective function of Eq. (21), the entire identification objective function is:

$$J = \frac{1}{N} \sum_{k=1}^N [y(k) - \widehat{y}(k)]^2 = \frac{1}{N} \sum_{k=1}^N \left[ y(k) - \widehat{\Phi}^T \left( k, \widehat{\alpha} \right) \widehat{\theta} \right]^2 \quad (41)$$

where  $\widehat{y}(k) = \widehat{\Phi}^T \left( k, \widehat{\alpha} \right) \widehat{\theta}$  and  $N$  is the total number of samples.

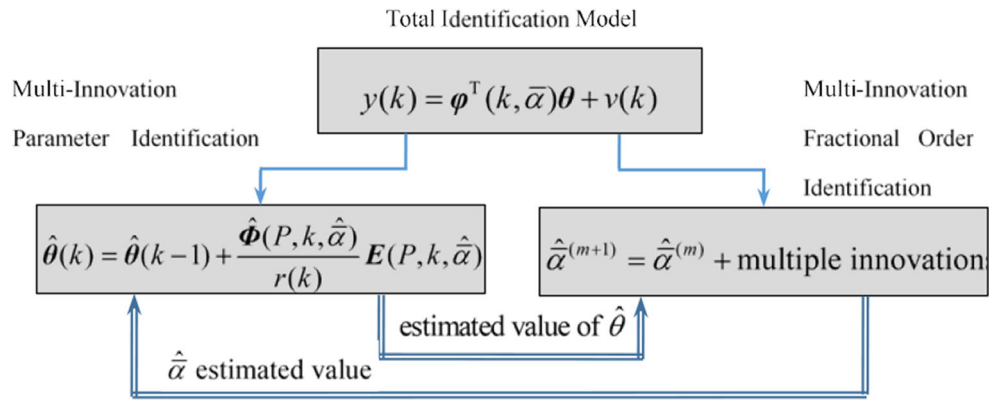
The purpose is to find a suitable  $\widehat{\alpha}$  for the parameters  $\widehat{\theta}$  estimated via the multi-innovation parameter identification algorithm such that  $J$  is as small as possible. The Levenberg-Marquardt algorithm iterates as follows:

$$\widehat{\alpha}^{(m+1)} = \widehat{\alpha}^{(m)} - \left\{ \left[ J'' + \lambda I \right]^{-1} J' \right\}_{\widehat{\alpha} = \widehat{\alpha}^{(m)}} \quad (42)$$

The update of the fractional order  $\widehat{\alpha}$  is based on the calculation of the gradient  $J'$  and the Hessian matrix  $J''$  corresponding to each  $\widehat{\alpha}$ , and  $\lambda$  is an adjustment parameter.  $J'$  and  $J''$  are calculated as follows:

$$\begin{aligned} J'_{\widehat{\alpha}} &= -\frac{2}{N} \left[ \frac{\partial \widehat{\varphi}^T \left( k, \widehat{\alpha} \right) \widehat{\theta}}{\partial \widehat{\alpha}} \right] \\ &= -\frac{2}{N} \left[ \frac{\partial \widehat{y}(k)}{\partial \widehat{\alpha}} \right]^T \left[ y(k) - \widehat{\varphi}^T \left( k, \widehat{\alpha} \right) \widehat{\theta} \right] \\ &= -\frac{2}{N} \left[ \sigma \widehat{y}(k) / \widehat{\alpha} \right]^T \left[ y(k) - \widehat{\varphi}^T \left( k, \widehat{\alpha} \right) \widehat{\theta} \right] \end{aligned} \quad (43)$$

**Fig. 2** The interactive identification process of two multi-innovation identification methods



where  $\sigma\hat{y}(k)/\hat{\alpha} = \frac{\partial\hat{y}(k)}{\partial\hat{\alpha}}$  is the sensitivity function with respect to  $\hat{\alpha}$ , and its calculation process is:

$$\sigma\hat{y}(k)/\hat{\alpha} \approx \frac{\hat{y}(k, \hat{\alpha} + \delta\hat{\alpha}) - \hat{y}(k, \hat{\alpha})}{\delta\hat{\alpha}} \tag{44}$$

where  $\delta\hat{\alpha}$  is the variation of  $\hat{\alpha}$ .

The equation for calculating the second derivative  $J''_{\hat{\alpha}}$  is:

$$\begin{aligned} J'_{\hat{\alpha}} &= \frac{2}{N} \left( \frac{\partial\hat{y}(k)}{\partial\hat{\alpha}} \right)^T \left( \frac{\partial\hat{y}(k)}{\partial\hat{\alpha}} \right) \\ &= \frac{2}{N} \left( \sigma\hat{y}(k)/\hat{\alpha} \right)^T \left( \sigma\hat{y}(k)/\hat{\alpha} \right) \end{aligned} \tag{45}$$

Thus,  $J'_{\hat{\alpha}}$  and  $J''_{\hat{\alpha}}$  are calculated as follows:

$$J'_{\hat{\alpha}} = -\frac{2}{N} \left[ \sigma\hat{y}(k)/\hat{\alpha} \right]^T \left[ y(k) - \hat{\Phi}^T(k, \hat{\alpha})\hat{\theta} \right] \tag{46a}$$

$$J''_{\hat{\alpha}} = \frac{2}{N} \left( \sigma\hat{y}(k)/\hat{\alpha} \right)^T \left( \sigma\hat{y}(k)/\hat{\alpha} \right) \tag{46b}$$

For the Levenberg-Marquardt algorithm of Eq. (42), its essence is to use the single-innovation  $y(k) - \hat{\Phi}^T(k, \hat{\alpha})\hat{\theta}$  estimation algorithm. The following will introduce the multi-innovation Levenberg-Marquardt algorithm, the basic idea of which is to expand the scalar innovation into an innovation

vector or matrix using both current and historical data. Many research results show that multi-innovation identification can effectively improve algorithm convergence and the accuracy of parameter estimation.

The stacked sensitivity function vector  $\Xi(P, k, \hat{\alpha})$  is defined as:

$$\Xi(P, k, \hat{\alpha}) = \left[ \sigma\hat{y}(k)/\hat{\alpha}, \sigma\hat{y}(k-1)/\hat{\alpha}, \dots, \sigma\hat{y}(k-P+1)/\hat{\alpha} \right]^T \in R^P \tag{47}$$

Based on the definitions of  $\hat{\Phi}(P, k, \hat{\alpha})$  and  $\mathbf{Y}(P, k)$ , the multi-innovation Levenberg-Marquardt algorithm can be expressed as follows:

$$\begin{aligned} \hat{\alpha}^{(m+1)} &= \hat{\alpha}^{(m)} + \frac{2}{N} \left\{ \left[ \frac{2}{N} \Xi^T(P, k, \hat{\alpha}) \Xi(P, k, \hat{\alpha}) + \lambda \right]^{-1} \right. \\ &\quad \left. \Xi^T(P, k, \hat{\alpha}) \left[ Y(P, k) - \hat{\Phi}^T(P, k, \hat{\alpha})\hat{\theta} \right] \right\}_{\hat{\alpha} = \hat{\alpha}^{(m)}} \end{aligned} \tag{48}$$

In this research, the multi-innovation gradient descent method is used to identify the parameters of the fractional-order nonlinear system, and the multi-innovation Levenberg-Marquardt identification algorithm is used to estimate the fractional order of the system. In accordance with interactive estimation theory and the hierarchical identification principle, two algorithms alternately perform parameter identification and fractional order estimation. In each iteration, the parameter estimates rely on previous fractional order

estimates, while fractional order estimation is performed based on the parameter estimation of the previous iteration, and the two steps form a complete hierarchical interactive calculation process. The whole algorithm can be summarized as follows:

**Step 1:** Let  $m = 1$  and  $k = 1$  and set the initial parameters  $\hat{\theta}^0$ ,  $\hat{\alpha}^0$ ,  $p$  and  $\delta\hat{\alpha}$ .

**Step 2:** Collect input/output data  $\{u(k), y(k)\}$ ,  $k = 1, 2, \dots, N$ .

**Step 3:** Construct the innovation matrix  $\hat{\Phi}(P, k, \hat{\alpha})$ , stack the output vector  $Y(P, k)$ , and calculate the fractional order sensitivity function  $\sigma\hat{y}(k)/\hat{\alpha}$  according to Eq. (44).

**Step 4:** Build a vector of stacked sensitivity functions,  $\Xi(P, k, \hat{\alpha})$ .

**Step 5:** Use Eqs. (46a) and (46b) to calculate  $J'_{\hat{\alpha}}$  and  $J''_{\hat{\alpha}}$ , and use Eq. (48) to obtain the estimated value  $\hat{\alpha}$  of the fractional order.

**Step 6:** Use Eqs. (29)–(38) to identify the parameters  $\hat{\theta}^{(m)}(k)$ , and use Eqs. (39) and (40) to calculate  $\hat{p}_j$  and  $\hat{q}_j$  ( $j = 2, 3, \dots, m$ ).

**Step 7:** Let  $k = k + 1$ . If  $k \leq N$ , go to **Step 6**; otherwise, go to **Step 8**.

**Step 8:** Use Eq. (41) to calculate the objective function.

**Step 9:** If  $\frac{|J(\hat{\theta}^{(m+1)}) - J(\hat{\theta}^{(m)})|}{|J(\hat{\theta}^{(m)})|} \leq \xi$ , then  $\hat{\theta} = \hat{\theta}^{(m)}$ ,  $\hat{\alpha} = \hat{\alpha}^{(m)}$ , and  $J(\hat{\theta}) = J(\hat{\theta}^{(m)})$ ; otherwise, set

$m = m + 1$  and return to **Step 3**.

To help readers understand the logic and innovation of this paper more clearly, we summarize the proposed algorithm in the form of pseudocode, as shown in Table 1.

## 5 Performance analysis

To better illustrate the performance of the algorithm, some mathematical notation is first introduced. The symbols  $\lambda_{\max}[\mathbf{X}]$  and  $\lambda_{\min}[\mathbf{X}]$  represent the largest and smallest eigenvalues, respectively, of the matrix  $\mathbf{X}$ . For  $g(k) \geq 0$ ,  $f(k) = O(g(k))$  or  $f(k) \sim O(g(k))$  represents that there exists a constant  $\delta > 0$  such that  $f(k) \leq \delta g(k)$ . In addition, the following lemma is given.

**Lemma 1** Suppose that  $\{x(k)\}$ ,  $\{a_k\}$  and  $\{b_k\}$  are sequences of nonnegative random variables and satisfy the following relation:

$$x(k) \leq (1 - a_k)x(k) + b_k, k \geq 0,$$

where  $a_k \in [0, 1)$  and  $x(0) < \infty$ ; then, it holds that

$$\lim_{k \rightarrow \infty} x(k) \leq \lim_{k \rightarrow \infty} \frac{b_k}{a_k}.$$

**Lemma 2** For system (19) and the single-innovation gradient descent algorithm given by Eqs. (23) and (24), there is a constant  $0 < \bar{\alpha} < \beta < \infty$ , and the fractional order estimate  $\hat{\alpha}$  causes the input fractional order information vector  $\hat{\varphi}(k, \hat{\alpha})$  of the system to satisfy the following continuous excitation conditions:

$$(A1) \quad \hat{\alpha} \mathbf{I}_n \leq \frac{1}{N} \sum_{i=0}^{N-1} \hat{\varphi}(k+i, \hat{\alpha}) \hat{\varphi}^T(k+i, \hat{\alpha}) \leq \beta \mathbf{I}_n, \text{ a.s. } k > 0$$

Then,  $r(k)$  in Eq. (24) satisfies the following inequality:  $n\bar{\alpha}(k-N+1)+1 \leq r(k) \leq n\beta(k+N-1)+1$ , a.s.  $k > 0$ .



**Table 1** Pseudocode of the algorithm in this paper

<b>Algorithm 1</b>	
1:	<b>Input:</b> Input and output $\{u, y\}$ , multiple innovations $L$ , nonlinear terms $g(\bullet)$ and $f(\bullet)$
2:	<b>Output:</b> System parameter estimates $\hat{\theta}$ , fractional order $\hat{\alpha}$
3:	Initializ $m \leftarrow 1$ , $k \leftarrow 1$ , $r(0) \leftarrow 1$ , $\lambda \leftarrow 1$ , $\delta\hat{\alpha} \leftarrow 0.001$ , and $\hat{\theta}^{(0)}$ , $\hat{\alpha}^{(0)}$ ;
4:	$N \leftarrow \text{length}(y)$ ;
5:	<b>while</b> $k \leq N$
6:	Collect $\{u(k), y(k)\}$ ;
7:	Initialize $\hat{\psi}(k, \hat{\alpha}^{(m)}) \leftarrow []$ and $\hat{\zeta}(k, \hat{\alpha}^{(m)}) \leftarrow []$ ;
8:	<b>for</b> $i \leftarrow 1$ to $n_p$
9:	<b>do</b> $\hat{\psi}(k, \hat{\alpha}^{(m)}) \leftarrow [\hat{\psi}(k, \hat{\alpha}^{(m)}), [\Delta^{\hat{\alpha}} g_i(y(k-1)), \dots, \Delta^{\hat{\alpha}} g_i(y(k-n_a))]^T]$ ;
10:	<b>end</b>
11:	<b>for</b> $i \leftarrow 1$ to $n_f$
12:	<b>do</b> $\hat{\zeta}(k, \hat{\alpha}^{(m)}) \leftarrow [\hat{\zeta}(k, \hat{\alpha}^{(m)}), [\Delta^{\hat{\alpha}} f_j(u(k-1)), \dots, \Delta^{\hat{\alpha}} f_j(y(k-n_b))]^T]$ ;
13:	<b>end</b>
14:	Build an information vector $\hat{\phi}(k, \hat{\alpha}^{(m)}) \leftarrow [\hat{\psi}(k, \hat{\alpha}^{(m)}); \hat{\zeta}(k, \hat{\alpha}^{(m)})]^T$ ;
15:	Construct multi-innovation vectors $\hat{\Phi}(L, k, \hat{\alpha}^{(m)}) \leftarrow [\hat{\phi}(k, \hat{\alpha}^{(m)}), \hat{\phi}(k-1, \hat{\alpha}^{(m)}), \dots, \hat{\phi}(k-L+1, \hat{\alpha}^{(m)})]$ ,
16:	$Y(L, k) \leftarrow [y(k), y(k-1), \dots, y(k-L+1)]^T$ ,
17:	$E(L, k, \hat{\alpha}) \leftarrow Y(L, k) - \hat{\Phi}^T(L, k, \hat{\alpha})\hat{\theta}^{(m)}$ ,
18:	$r(k) \leftarrow r(k-1) + \ \hat{\Phi}(L, k, \hat{\alpha})\ ^2$ ;
19:	Update system parameters $\hat{\theta}^{(m)} \leftarrow \hat{\theta}^{(m-1)} + (\hat{\Phi}(L, k, \hat{\alpha})/r(k))E(L, k, \hat{\alpha})$ ;
20:	<b>for</b> $l \leftarrow 0$ to $L-1$
21:	<b>do</b> $\hat{y}(k-l, \hat{\alpha}^{(m)} + \delta\hat{\alpha}) \leftarrow \hat{\phi}^T(k-l, \hat{\alpha}^{(m)} + \delta\hat{\alpha})\hat{\theta}^{(m)}$ ;
22:	$\hat{y}(k-l, \hat{\alpha}^{(m)}) \leftarrow \hat{\phi}^T(k-l, \hat{\alpha}^{(m)})\hat{\theta}^{(m)}$ ;
23:	$\sigma\hat{y}(k-l)/\hat{\alpha}^{(m)} \leftarrow \hat{y}(k-l, \hat{\alpha}^{(m)} + \delta\hat{\alpha}) - \hat{y}(k-l, \hat{\alpha}^{(m)})/\delta\hat{\alpha}$ ;
24:	<b>end</b>
25:	$\Xi(L, k, \hat{\alpha}) \leftarrow [\sigma\hat{y}(k)/\hat{\alpha}, \sigma\hat{y}(k-1)/\hat{\alpha}, \dots, \sigma\hat{y}(k-L+1)/\hat{\alpha}]^T$ ;
26:	$\hat{\alpha}^{(m+1)} \leftarrow \hat{\alpha}^{(m)} + \frac{2}{N} \left\{ \begin{aligned} & \left[ \frac{2}{N} \Xi^T(L, k, \hat{\alpha})\Xi(L, k, \hat{\alpha}) + \lambda I \right]^{-1} \cdot \\ & \Xi^T(L, k, \hat{\alpha})[Y(L, k) - \hat{\Phi}^T(L, k, \hat{\alpha})\hat{\theta}] \end{aligned} \right\}_{\hat{\alpha}=\hat{\alpha}^{(m)}} ;$
27:	$J^{(m)} \leftarrow 1/N \sum_{k=1}^N [y(k) - \hat{\phi}^T(k, \hat{\alpha}^{(m)})\hat{\theta}^{(m)}]^2$ ;
28:	<b>if</b> $ J(\hat{\theta}^{(m+1)}) - J(\hat{\theta}^{(m)}) / J(\hat{\theta}^{(m)})  \leq \xi$
29:	<b>then</b> output $\hat{\theta} \leftarrow \hat{\theta}^{(m)}$ , $\hat{\alpha} \leftarrow \hat{\alpha}^{(m)}$ and $J(\hat{\theta}) \leftarrow J^{(m)}$ ;
30:	<b>else</b> $m \leftarrow m+1$ ;
31:	<b>end</b>
32:	$k \leftarrow k+1$ ;
33:	<b>end</b>

**Proof** By tracing both sides of condition (A1), we can obtain:

$$nN\bar{\alpha} \leq \frac{1}{N} \sum_{i=0}^{N-1} \left\| \widehat{\varphi}(k+i, \widehat{\alpha}) \right\|^2 \leq nN\beta, \text{ a.s. } k > 0.$$

Let  $[x]$  be the largest integer sum  $nN\beta = \delta_1$  no greater than  $x$ ; it holds that  $\left\| \widehat{\varphi}(k+i, \widehat{\alpha}) \right\|^2 \leq \delta_1$ , a.s., and for Eq. (24), continuous iterative calculation is performed:

$$\begin{aligned} r(k) &= r(k-1) + \left\| \widehat{\varphi}(k, \widehat{\alpha}) \right\|^2 = \sum_{j=1}^k \left\| \widehat{\varphi}(j, \widehat{\alpha}) \right\|^2 + r(0) \\ &\leq \sum_{j=0}^{\lfloor \frac{k-1}{N} \rfloor} \sum_{i=1}^N \left\| \widehat{\varphi}(jN+i, \widehat{\alpha}) \right\|^2 + r(0) \\ &\leq \sum_{j=0}^{\lfloor \frac{k-1}{N} \rfloor} \delta_1 + r(0) \\ &\leq \left( \left\lfloor \frac{k-1}{N} \right\rfloor + 1 \right) \delta_1 + 1 \\ &\leq n\beta(k+N-1) + 1 \end{aligned} \tag{49}$$

In addition:

$$\begin{aligned} r(k) &\geq \sum_{j=0}^{\lfloor \frac{k}{N} \rfloor - 1} \sum_{i=1}^N \left\| \widehat{\varphi}(jN+i, \widehat{\alpha}) \right\|^2 + r(0) \\ &\geq \sum_{j=0}^{\lfloor \frac{k}{N} \rfloor - 1} nN\bar{\alpha} + r(0) \\ &\geq \left( \left\lfloor \frac{k}{N} \right\rfloor \right) nN\bar{\alpha} + 1 \\ &\geq n\bar{\alpha}(k-N+1) + 1 \end{aligned} \tag{50}$$

The proof of the lemma is complete.

**Theorem** For system (19) and the multi-innovation gradient descent algorithm given by Eqs. (29)–(38), the fractional order estimate  $\widehat{\alpha}$  causes the input fractional order information vector  $\widehat{\varphi}(k, \widehat{\alpha})$  of the system to satisfy the continuous excitation condition (A1), under the assumption that the noise signal  $\{v(k)\}$  is an independent random signal that satisfies.

$$(A2) \ E[v(k)] = 0, \ E[v(k)v(i)] = 0, \ k \neq i, \ E[(v(k))^2] = \sigma_v^2,$$

where  $E(\cdot)$  is the mathematical expectation. Then, the parameter estimation error vector  $\widetilde{\theta}(k) := \widehat{\theta}(k) - \theta$  satisfies:

$$\lim_{k \rightarrow \infty} E \left[ \left\| \widetilde{\theta}(k) \right\|^2 \right] \leq \lim_{k \rightarrow \infty} \frac{N\beta\sigma_v^2[n\beta(k-N+1) + 1]}{\bar{\alpha} [n\bar{\alpha}(k-N+1) + 1]^2}.$$

**Proof** To simplify the calculation, let the innovation length be  $P = N$ , and define the noise vector as:

$$\mathbf{V}(P, k) = [v(k), v(k-1), \dots, v(k-P+1)]^T \in R^P,$$

By subtracting  $\theta$  from both sides of Eq. (29) and using Eqs. (30), (32), (33) and (19), we obtain:

$$\begin{aligned} \widetilde{\theta}(k) &= \widetilde{\theta}(k-1) + \frac{\widehat{\Phi}(P, k, \widehat{\alpha})}{r(k)} \left[ -\widehat{\Phi}^T(P, k, \widehat{\alpha}) \widetilde{\theta}(k-1) + V(P, k) \right] \\ &= \left[ I - \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) \widehat{\Phi}^T(P, k, \widehat{\alpha})}{r(k)} \right] \widetilde{\theta}(k-1) + \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) V(P, k)}{r(k)} \end{aligned} \tag{51}$$

Taking the norm on both sides of the above equation yields:

$$\begin{aligned} \left\| \widetilde{\theta}(k) \right\|^2 &\leq \left\| \left[ I - \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) \widehat{\Phi}^T(P, k, \widehat{\alpha})}{r(k)} \right] \widetilde{\theta}(k-1) \right\|^2 + \dots \\ &\quad 2\widetilde{\theta}^T(k-1) \left[ I - \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) \widehat{\Phi}^T(P, k, \widehat{\alpha})}{r(k)} \right] \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) V(P, k)}{r(k)} \\ &\quad + \frac{\left\| \widehat{\Phi}(P, k, \widehat{\alpha}) V(P, k) \right\|^2}{r^2(k)} \\ &\leq \lambda_{\max} \left[ I - \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) \widehat{\Phi}^T(P, k, \widehat{\alpha})}{r(k)} \right] \left\| \widetilde{\theta}(k-1) \right\|^2 + \dots \\ &\quad 2\widetilde{\theta}^T(k-1) \left[ I - \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) \widehat{\Phi}^T(P, k, \widehat{\alpha})}{r(k)} \right] \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) V(P, k)}{r(k)} \\ &\quad + \frac{\left\| \widehat{\Phi}(P, k, \widehat{\alpha}) V(P, k) \right\|^2}{r^2(k)} \end{aligned} \tag{52}$$

Using Lemma 1 and condition (A1), we can obtain:

$$I - \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) \widehat{\Phi}^T(P, k, \widehat{\alpha})}{r(k)} \leq \left[ 1 - \frac{N\bar{\alpha}}{n\beta(k-N+1)+1} \right] I, \text{ a.s.}$$

$$\begin{aligned} & \mathbb{E} \left[ \left\| \widehat{\Phi}(P, k, \widehat{\alpha}) V(P, k) \right\|^2 \right] \\ & \leq \mathbb{E} \left\{ \lambda_{\max} \left[ \widehat{\Phi}(P, k, \widehat{\alpha}) \widehat{\Phi}^T(P, k, \widehat{\alpha}) \right] \left\| V(P, k) \right\|^2 \right\} \\ & \leq L\beta \mathbb{E} \left[ \left\| V(P, k) \right\|^2 \right] \leq P^2 \beta \sigma_v^2 = N^2 \beta \sigma_v^2 \end{aligned} \tag{53}$$

Taking mathematical expectations on both sides of Eq. (52) at the same time, and noting that  $V(P, k)$  and  $\tilde{\theta}(k-1)$ ,  $\widehat{\Phi}(P, k, \widehat{\alpha})$ ,  $I - \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) \widehat{\Phi}^T(P, k, \widehat{\alpha})}{r(k)}$  are not linearly correlated, using conditions (A1) and (A2), we have

$$\begin{aligned} \mathbb{E} \left[ \left\| \tilde{\theta}(k) \right\|^2 \right] & \leq \left[ 1 - \frac{N\bar{\alpha}}{n\beta(k-N+1)+1} \right] \mathbb{E} \left[ \left\| \tilde{\theta}(k-1) \right\|^2 \right] + \dots \\ & 2\mathbb{E} \left\{ \tilde{\theta}(k-1) \left[ I - \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) \widehat{\Phi}^T(P, k, \widehat{\alpha})}{r(k)} \right] \frac{\widehat{\Phi}(P, k, \widehat{\alpha}) V(P, k)}{r(k)} \right\} \\ & \quad + \frac{N^2 \beta \sigma_v^2}{\left[ n\bar{\alpha}(k-N+1)+1 \right]^2} \\ & \leq \left[ 1 - \frac{N\bar{\alpha}}{n\beta(k-N+1)+1} \right] \mathbb{E} \left[ \left\| \tilde{\theta}(k-1) \right\|^2 \right] + \frac{N^2 \beta \sigma_v^2}{\left[ n\bar{\alpha}(k-N+1)+1 \right]^2} \end{aligned} \tag{54}$$

Using Lemma 1, we obtain:

$$\lim_{k \rightarrow \infty} \mathbb{E} \left\{ \left\| \tilde{\theta}(k) \right\|^2 \right\} \leq \lim_{k \rightarrow \infty} \frac{N\beta\sigma_v^2[n\beta(k-N+1)+1]}{\bar{\alpha} \left[ n\bar{\alpha}(k-N+1)+1 \right]^2} \tag{55}$$

The theorem is proven.

The meaning of the above theorem is that since the two multi-innovation algorithms are interactive, if the fractional order estimate can ensure that the input vector of the system is continuously excited, then the parameter identification error of the system can be bounded.

## 6 Experimental study

### 6.1 Academic example

To verify the effectiveness of the proposed algorithm, the following fractional-order nonlinear system is considered:

$$y(k) = A(z)g(y(k)) + B(z)f(u(k)) + v(k) \tag{56}$$

where  $A(z) = a_1 z^{-\bar{\alpha}} + a_2 z^{-2\bar{\alpha}}$ ,  $B(z) = b_1 z^{-\bar{\alpha}} + b_2 z^{-2\bar{\alpha}}$ ,  $f(u(k)) = \sum_{j=1}^2 p_j f_j(u(k))$ , and  $g(y(k)) = \sum_{j=1}^3 q_j g_j(y(k))$ .

The entire output of the system is:

$$\begin{aligned} y(k) & = \sum_{j=1}^3 q_j \sum_{i=1}^2 a_i \Delta^{\bar{\alpha}} g_j(y(k-i)) + \dots \\ & + \sum_{j=1}^2 p_j \sum_{i=1}^2 b_i \Delta^{\bar{\alpha}} f_j(u(k-i)) + v(k) \end{aligned} \tag{57}$$

where  $f_1(u(k-i)) = u(k-i)$ ,  $f_2(u(k-i)) = u^2(k-i)$ ,  $g_1(y(k-i)) = y(k-i)$ ,  $g_2(y(k-i)) = y^2(k-i)$ , and  $g_3(y(k-i)) = y^3(k-i)$ .

When the fractional order is taken to be  $\bar{\alpha} = 0.6$ , the entire output of the system is:

$$\begin{aligned} y(k) & = a_1 \Delta^{0.6} y(k-1) + a_2 \Delta^{0.6} y(k-2) + q_2 a_1 \Delta^{0.6} y^2(k-1) + \\ & q_2 a_2 \Delta^{0.6} y^2(k-2) + q_3 a_1 \Delta^{0.6} y^3(k-1) + \\ & q_3 a_2 \Delta^{0.6} y^3(k-2) + b_1 \Delta^{0.6} u(k-1) + b_2 \Delta^{0.6} u(k-2) \\ & + p_2 b_1 \Delta^{0.6} u^2(k-1) + p_2 b_2 \Delta^{0.6} u^2(k-2) + v(k) \end{aligned}$$

The parameter vector is:

$$\begin{aligned} \mathbf{a} & = [a_1 \ a_2]^T = [0.1 \ 0.2]^T, \mathbf{b} = [b_1 \ b_2]^T = [-0.4 \ - \\ & 0.2]^T, \mathbf{p} = [p_1 \ p_2]^T = [1 \ 0.5]^T, \mathbf{q} = [q_1 \ q_2 \ q_3]^T = \\ & [1 \ 0.7 \ 0.35]^T. \end{aligned}$$

$$\begin{aligned} \theta & = [a_1 \ a_2 \ q_2 a_1 \ q_2 a_2 \ q_3 a_1 \ q_3 a_2 \ b_1 \ b_2 \ p_2 b_1 \ p_2 b_2]^T \\ & = [0.1 \ 0.2 \ 0.07 \ 0.14 \ 0.035 \ 0.07 \ -0.4 \ -0.2 \\ & -0.2 \ -0.1]^T \end{aligned}$$

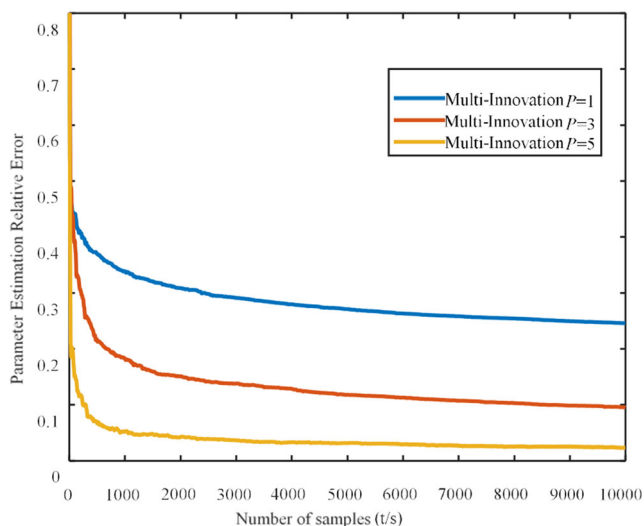
During the simulation, the input signal is a random signal with zero mean and unit variance, and sampling is performed 10,000 times. The noise signal is an independent random signal with zero mean and a variance of  $\sigma^2 = 0.01$ . In this study, multi-innovation lengths of  $P = 1, 3, 5$  are selected. To verify the effectiveness of the proposed method, the relative parameter estimation error  $\delta = \left\| \widehat{\theta}(k) - \theta \right\| / \left\| \theta \right\|$  is used as the evaluation indicator for verification. The parameter identification results are given in Table 2 and Table 2 (continued). Figure 3 shows the parameter estimation error curves with different multi-innovations; Fig. 4 shows the estimation results for the fractional order with different multi-innovations. From the analysis and comparison of these results, it is obvious that as the multi-innovation length increases, the convergence speed of identification becomes faster and the identification accuracy becomes higher. Figure 5 compares the output of the identification model with the actual output when the multi-innovation length is  $P = 5$ .

To illustrate the superiority of the proposed method, the method proposed in this paper is compared with the single-

**Table 2** Partial parameter identification results of fractional-order nonlinear systems ( $P = 1, 3, 5$ )

$P$	$k$	$\hat{a}_1$	$\hat{a}_2$	$\hat{b}_1$	$\hat{b}_2$	$\hat{\alpha}$	
1	2000	0.1259	0.2212	-0.4203	-0.1725	0.4576	
	5000	0.1138	0.1705	-0.4121	-0.1842	0.4585	
	10,000	0.1045	0.2070	-0.4102	-0.1901	0.4587	
3	2000	0.0744	0.2145	-0.4155	-0.1767	0.5170	
	5000	0.1250	0.2103	-0.4100	-0.1973	0.5179	
	10,000	0.1035	0.2057	-0.3929	-0.2046	0.5187	
5	2000	0.1010	0.1810	-0.4096	-0.1886	0.5987	
	5000	0.0882	0.1892	-0.3825	-0.1915	0.6031	
	10,000	0.0982	0.1959	-0.3977	-0.1986	0.6026	
True values		0.1	0.2	-0.4	-0.2	0.6	
$P$	$k$	$\hat{q}_2 \hat{a}_1$	$\hat{q}_2 \hat{a}_2$	$\hat{q}_3 \hat{a}_1$	$\hat{q}_3 \hat{a}_2$	$\hat{p}_2 \hat{b}_1$	$\hat{p}_2 \hat{b}_2$
1	2000	0.0729	0.1298	0.3422	0.6835	-0.1871	-0.0933
	5000	0.0722	0.1345	0.3454	0.6937	-0.1896	-0.0941
	10,000	0.0715	0.1355	0.0358	0.0697	-0.2013	-0.1018
3	2000	0.0657	0.1298	0.0372	0.7311	-0.1884	-0.1819
	5000	0.0661	0.1311	0.0367	0.0726	-0.1939	-0.0727
	10,000	0.0668	0.1367	0.0361	0.0714	-0.1966	-0.1012
5	2000	0.6672	0.1537	0.0371	0.0751	-0.1834	-0.1046
	5000	0.06933	0.1524	0.0363	0.0745	-0.1953	-0.0985
	10,000	0.072	0.1432	0.0358	0.072	-0.1969	-0.0991
True values		0.07	0.14	0.035	0.07	-0.2	-0.1

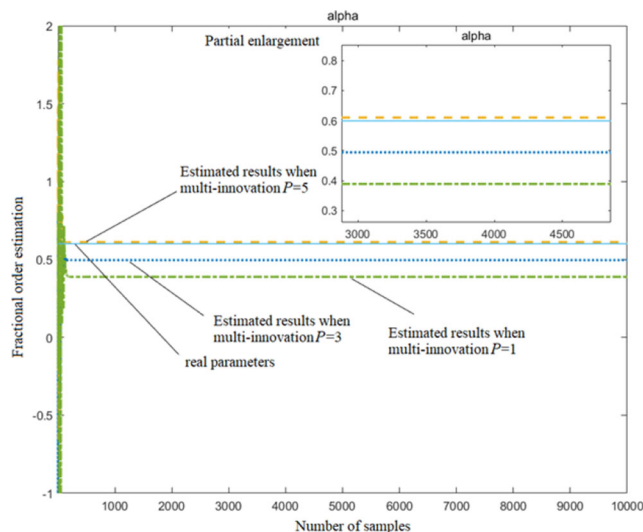
innovation Levenberg-Marquardt algorithm and the single-innovation gradient descent method proposed in references [28, 29], etc. Figure 6 shows the results for the method in this paper ( $P = 5$ ) in terms of the relative error curves with respect to the methods of [28, 29]. From a comparative analysis, it is obvious that the method in this paper offers a better identification convergence speed and higher identification accuracy.



**Fig. 3** Parameter estimation error curves for different multi-innovation

### 6.2 Actual system

To further illustrate the applicability of this method in a practical system, we model and identify a flexible manipulator system from the Manufacturing and Automation Laboratory of Katolik University Leuven. In this system, a robotic arm is mounted on a motor; the obtained system input is the reaction



**Fig. 4** Estimation results of fractional order of different multi-innovation

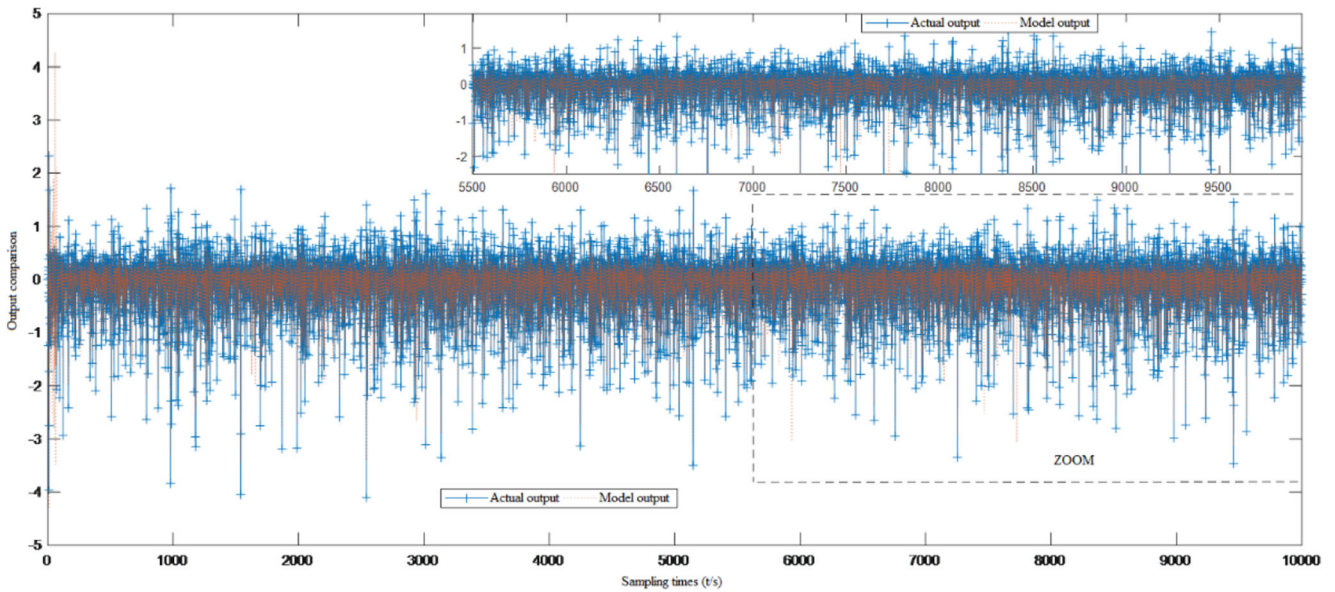


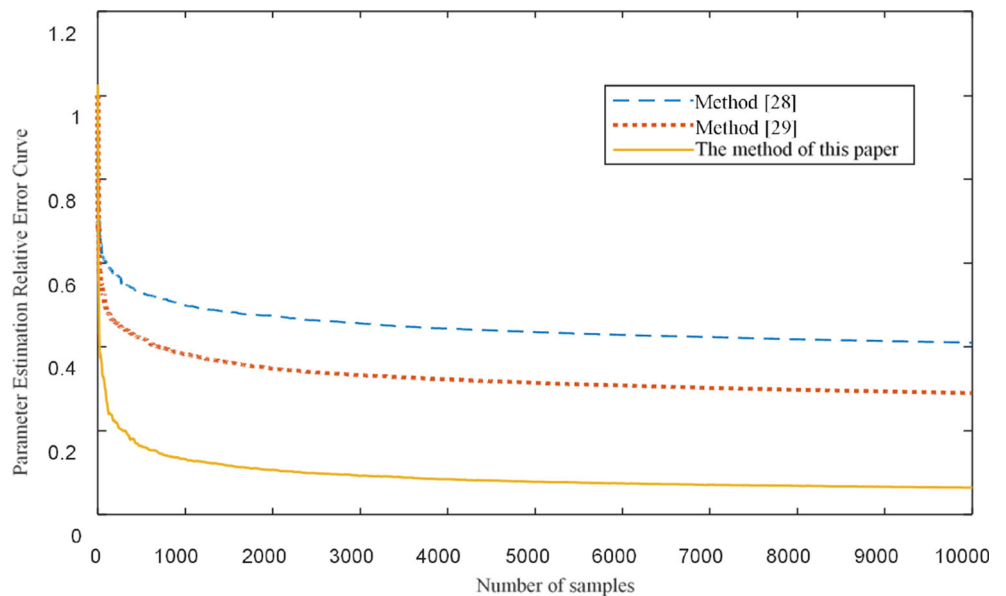
Fig. 5 Comparison of the output of the identification model ( $P = 5$ ) with the actual output

torque of the structure, and the output is the acceleration of the flexible arm [30]. To model this system using the method proposed in this paper, the structure of the system needs to be determined first. Reference [28] has carried out Hammerstein-Wiener modeling for this system; therefore, we adopt the system structure determined in the literature [2, 3] to directly apply the method proposed in this paper. Unlike in the academic example of the previous subsection, where the real parameters of the system are known and the modeling effect can be evaluated simply by comparing the accuracy of the parameters, in an actual system, we do not know the real system parameters and so we need to compare the estimated output of the system with the actual output fitting effect.

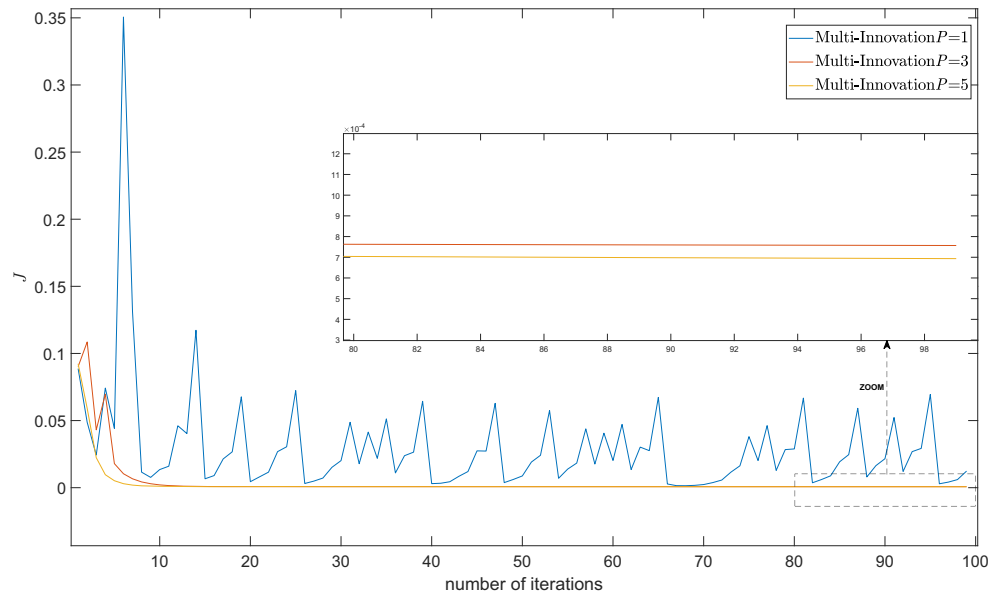
First, on the premise that the system structure is determined, we use the method proposed in this paper to model the robotic arm system and obtain the objective function values under different innovation numbers. Figure 7 shows the objective functions for innovation numbers of  $P = 1, 3, 5$ . It can be seen from Fig. 7 that with an increase in the number of iterations, the objective function gradually converges. By zooming in on the relevant part of the figure, it becomes obvious that the greater the number of innovations is, the higher the estimation accuracy and the smaller the objective function. When the number of innovations is  $P = 5$ , the objective function is the smallest.

Therefore, an innovation number of  $P = 5$  is selected to model the robotic arm system. Figure 8 shows the output

Fig. 6 Relative error curve of parameter estimation between the method in this paper ( $P = 5$ ) and the method in the literature



**Fig. 7** Objective function  $J$  under different multi-innovation



**Table 3** Comparison of the error range of the robot arm estimation with the literature ( $P = 5$ )

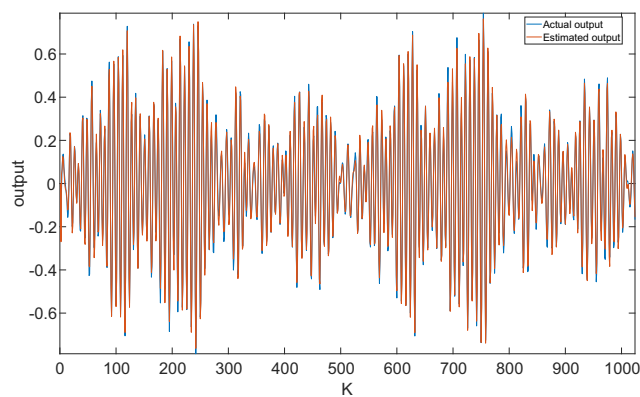
Method	[28]	[14]	The method of this paper
Error range	$[-0.2,0.2]$	$[-0.15,0.15]$	$[-0.1,0.1]$

Table 3. It can be clearly seen that the error of the method proposed in this paper is smaller and its modeling accuracy is higher.

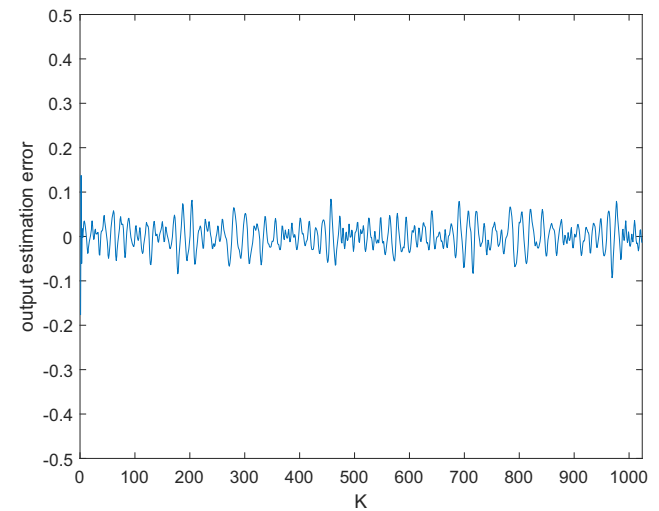
fitting diagram when the innovation number is 5. The actual output and the estimated output are basically in agreement. Figure 9 shows the system estimation error. It can be clearly seen that this error is small and the modeling accuracy is high. To further verify the superiority of the method in this paper, we have added comparisons with other methods from the literature. References [14, 28] proposed different methods for the modeling and identification of this manipulator system and presented corresponding output error figures. We compare the error with the results of these studies as shown in

### 7 Conclusion

To overcome the difficulties in identifying fractional-order nonlinear systems, a hybrid parameter identification algorithm based on the principle of multiple innovations is proposed. A multi-innovation recursive gradient descent algorithm and a multi-innovation Levenberg-Marquardt algorithm are designed based on the principle of multi-innovation



**Fig. 8** System actual output and estimated output ( $P = 5$ )



**Fig. 9** System estimated output error ( $P = 5$ )

identification to estimate the model parameters and the fractional order of the system. As seen through simulation and verification based on an academic example, proper use of the multi-innovation principle can increase the convergence speed of the proposed identification algorithm and improve the identification accuracy. In addition, the proposed algorithm is applied to an actual flexible manipulator system to verify the applicability of the multi-innovation principle in solving practical problems and confirm the practicability of the algorithm. In this paper, the proposed algorithm is verified in both academic and practical applications, and the model identification results are compared with those of algorithms proposed by other scholars. The method proposed in this paper has higher modeling accuracy and lower error.

From our simulation study, we obtain the following conclusions:

- 1) With an appropriate increase in the number of innovations, the convergence speed of the system identification algorithm and the identification accuracy can be improved.
- 2) By introducing the multi-innovation principle, a multi-innovation gradient descent algorithm and a multi-innovation L-M algorithm are designed. The two algorithms take turns estimating the model parameters and fractional order, and the overall algorithm is simple and convenient.

The modeling of MIMO fractional-order nonlinear systems is still a challenging topic, and this is also a future direction of study for researchers.

**Funding** This work was supported by the National Natural Science Foundation of China (Grant number 61863034).

## Declarations

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

## References

1. Matlob MA, Jamali Y (2019) The concepts and applications of fractional order differential calculus in modeling of viscoelastic systems: a primer. *Crit Rev Biomed Eng* 47(4):249–276
2. Yavuz M, Sene N, Yıldız M (2022) Analysis of the influences of parameters in the fractional second-grade fluid dynamics. *Mathematics* 10(7):1125
3. Failla G, Zingales M (2020) Advanced materials modelling via fractional calculus: challenges and perspectives. *Phil Trans R Soc A* 378(2172):20200050
4. Qureshi S, Yusuf A, Shaikh AA, Inc M, Baleanu D (2019) Fractional modeling of blood ethanol concentration system with real data application. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29(1):013143
5. Shulin L, Naxin C, Yan L et al (2017) Modeling and state-of-charge estimation of automotive lithium-ion batteries based on fractional-order theory. *Chin J Electrotechnical Technol* 32(4):189–195
6. Jajarmi A, Baleanu D (2021) On the fractional optimal control problems with a general derivative operator. *Asian Journal of Control* 23(2):1062–1071
7. Wang S, Takyi-Aninakwa P, Jin S et al (2022) An improved feedforward-long short-term memory modeling method for the whole-life-cycle state of charge prediction of lithium-ion batteries considering current-voltage-temperature variation. *Energy* 124224
8. Wang S, Jin S, Bai D, Fan Y, Shi H, Fernandez C (2021) A critical review of improved deep learning methods for the remaining useful life prediction of lithium-ion batteries. *Energy Rep* 7:5562–5574
9. Zou C, Zhang L, Hu X, Wang Z, Wik T, Pecht M (2018) A review of fractional-order techniques applied to lithium-ion batteries, lead-acid batteries, and supercapacitors. *J Power Sources* 390:286–296
10. Zúñiga-Aguilar CJ, Gómez-Aguilar JF, Romero-Ugalde HM, Jahanshahi H, Alsaadi FE (2022) Fractal-fractional neuro-adaptive method for system identification. *Eng Comput* 38(4):3085–3108
11. Xiong R, Tian J, Shen W et al (2018) A novel fractional order model for state of charge estimation in lithium ion batteries. *IEEE Trans Veh Technol* 68(5):4130–4139
12. Burkovska O, Glusa C, D'elia M (2022) An optimization-based approach to parameter learning for fractional type nonlocal models. *Comput Math Appl* 116:229–244
13. Zhang Q, Wang H, Liu C (2021) Identification of fractional-order Hammerstein nonlinear ARMAX system with colored noise. *Nonlinear Dynamics* 106(4):3215–3230
14. Zhang Q, Wang H, Liu C (2022) MILM hybrid identification method of fractional order neural-fuzzy Hammerstein model. *Nonlinear Dynamics*:1–15
15. Rahmani MR, Farrokhi M (2018) Identification of neuro-fractional Hammerstein systems: a hybrid frequency–time-domain approach. *Soft Comput* 22(24):8097–8106
16. Ze D, Ning M (2019) Parameter estimation of fractional-order chaotic systems by differential quantum particle swarm optimization. *Journal of System. Simulation* 31(08):1664–1673. <https://doi.org/10.16182/j.issn1004731x.joss.17-0265>
17. Guo H, Gu W, Khayatnezhad M, Ghadimi N (2022) Parameter extraction of the SOFC mathematical model based on fractional order version of dragonfly algorithm. *Int J Hydrog Energy* 47(57):24059–24068
18. Lai X, He L, Wang S, Zhou L, Zhang Y, Sun T, Zheng Y (2020) Co-estimation of state of charge and state of power for lithium-ion batteries based on fractional variable-order model. *J Clean Prod* 255:120203
19. Hammar K, Djamah T, Bettayeb M (2015) Fractional hammerstein system identification using particle swarm optimization //2015 7th international conference on modelling, identification and control (ICMIC). IEEE, 1–6
20. Sersour L, Djamah T, Bettayeb M (2015) Identification of wiener fractional model using self-adaptive velocity particle swarm optimization//2015 7th international conference on modelling, identification and control (ICMIC). IEEE, 1–6
21. Jin QB, Wang Z, Liu XP (2015) Auxiliary model-based inter-varying multi-innovation least squares identification for multivariable OE-like systems with scarce measurements. *J Process Control* 35(1):154–168
22. Chaudhary NI, Raja MAZ, He Y, Khan ZA, Tenreiro Machado JA (2021) Design of multi innovation fractional LMS algorithm for parameter estimation of input nonlinear control autoregressive systems. *Appl Math Model* 93:412–425

23. Xu L, Ding F, Zhu Q (2022) Separable synchronous multi-innovation gradient-based iterative signal modeling from on-line measurements. *IEEE Trans Instrum Meas* 71:1–13
24. Ding F (2016) *System identification: multi-innovation identification theory and method*. Science Press, Beijing
25. Mohammad JM, Hamed M, Mohammad T (2018) Recursive identification of multiple-input single-output fractional-order Hammerstein model with time delay. *Appl Soft Comput* 70(3): 486–500
26. Qi ZD, Sun Q, Ge WP, He YK (2020) Nonlinear modeling of PEMFC based on fractional order subspace identification. *Asian J Control* 22(5):1892–1900
27. Wang JC, Wei YH, Liu TY, Li A, Wang Y (2020) Fully parametric identification for continuous time fractional order Hammerstein systems. *J Franklin Inst* 357(1):651–666
28. Hammar K, Djamah T, Bettayeb M (2019, Published online) Nonlinear system identification using fractional Hammerstein-wiener models. *Nonlinear Dyn* 98:2327–2338
29. Feng D (2014) *System identification-performance analysis of identification methods*. Science Press. (in Chinese)
30. De Moor B, De Gerssem P, De Schutter B et al (1997) DAISY: a database for identification of systems. *Journal A* 38:4–5

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



**Zhang Qian** was born in 1995. She is a PhD candidate in the School of Electrical Engineering, Xinjiang University. Her research direction: Research on Fractional Order Identification and Control Method and Its Application in PEMFC.



**Wang Hongwei** was born in 1969, doctor of engineering, professor. Graduated from Harbin Industry in 1999 University of Automatic Control Department. Now working in the School of Control Science and Engineering, Dalian University of Technology. Has published 110 SCI, EI and other search papers, and 2 invention patents. Host and participate in 5 national and local projects. Research direction: network control system, multi-sampling rate system control, switching system control.



**Liu Chunlei** was born in 1996. He is a PhD candidate in the School of Electrical Engineering, Xinjiang University. His research direction: Research on Life Prediction and Life Extension Control Strategy of PEMFC Power Generation System Based on Fractional Order Theory.