Identification of Fractional Order System with Scarce Measurement Data Based on Multi-Innovation Estimation Algorithm

Wang Hongwei  Zhang Qian  Zha Qin  Mutalifu Ahemaide

Abstract  Aiming at the modeling issues of fractional order Hammerstein system with scarce measurements, a novel multi-innovation hybrid identification algorithm is proposed to deal with them. Firstly, a multi-innovation estimation algorithm based on auxiliary model is presented to estimate the parameters of the nonlinear fractional order system, and a multi-innovation Levenberg-Marquardt algorithm are derived to confirm the fractional orders. Secondly, the convergence properties of the proposed algorithm are analyzed using the lemmas and theorems. Finally, in order to illustrate the effectiveness of the proposed algorithm, two fractional order nonlinear systems with scarce measurements are studied to prove the validity.

Keywords: fractional order nonlinear systems; system identification; multi-innovation principle; scarce measurements

Mutalifu Ahemaide(✉)  
School of Electrical Engineering, Xinjiang University, Urumqi 830047, P.R. China  
email:wq272427yrx@163.com

1 Introduction

In recent years, the researchers have been interested in the modeling and control of the real fractional order systems. Fractional order systems (FOSs) is more common than the traditional integer order systems, especially for the application with long memory characteristics, such as supercapacitors [1], thermal diffusion processes [2,3], and viscoelastic structures [4], etc. More and more researchers began to pay attention to fractional models, and found that fractional model has more accurate description method for many physical processes [5,6]. In order to deal with the modeling problems of fractional order system, some modeling methods are presented to deal with parameters identification and fractional orders estimation, including poisson moment functions (PMF) [7,8], modulation functions [9,10], auxiliary variable method [11], orthogonal basis functions [12], and block pulse functions [13], etc.

In practice, due to some reasons, including network communication, sensor failure, and hardware limitation, etc, some sampled data in the system are lost. For network control systems, there are data packet loss, data disorder and communication time delay to cause missing data. Due to manual operation, sensor failure and hardware limitation, some sampled data in the system are lost in industrial processes, chemical processes and biological processes. Multi-rate sampled systems with different input and output sampling periods which are very common in process industries can be viewed as a class of missing-data systems. Similarly, the fractional order systems in network environment, industrial processes, chemical processes and biological processes can also encounter the occurrence of scarce measurements or missing measurements, e.g., electroencephalographic
recorders [14], electric individual-wheel system on
an autonomous platform [15], etc.

For the system with scarce measurements or
missing measurements, the reconstruction methods
of scarce measurements or missing measurements
have received widespread attention in the control
field. One method is to substitute for missing data
points by interpolating calculation among available
data points, such as linear interpolation, parabolic
interpolation, cubic interpolation, spline function
interpolation, or piecewise constant interpolation
and so on. Another method is to use a dynamic model to
reconstruct the missing measurement data. Basing on
auxiliary model modeling principle, Ding et al.
presented a multi-innovation gradient-descent
algorithms to deal with the missing data of the scarce
measurement systems [16-17]. Broersen et al.
proposed a finite interval likelihood algorithm for
AR model spectral estimation using the conditional
log likelihoods method [18]. In [19], in order to
handle the scarce measurement systems, an auxiliary
model-based algorithm is proposed to deal with the
identification of multi-variable output error model
with scarce measurements. The effectiveness of the
proposed method is proved by comparing with other
methods in the literature.

In this paper, we attempt to study the identification
of fractional order Hammerstein systems (FOHS)
with scarce measurements. For the identification of
these systems, it is difficult how to estimate the
unknown parameters by traditional identification
methods. The main difficult factors embody: 1) There are many parameters of FOHS to be identified
by scarce measurements, including the model
parameters and the corresponding fractional orders;
2) Convergence of the identification algorithm; 3)
Influence of external interference and outside noise.

For the above-mentioned issues, we proposed a
new identification algorithm for FOHS with scarce
measurements. The innovation and contributions of
this paper are shown as follows: 1) Build a fractional
order model with scarce sampling data; 2) Based on
the theory of interactive estimation and the principle
of hierarchical identification, a multi-innovation
gradient descent (MIGD) algorithm and a
multi-innovation Levenberg-Marquardt (MILM)
algorithm are derived to estimate model parameters
and fractional orders, respectively; 3) Give the
convergence analysis of the presented algorithm
using the lemmas and theorems.

The rest of our paper is organized as follows.
Section 2 introduces the problem formulation of
fractional order system with scarce measurements.
Section 3 discusses the MIGD algorithm for
fractional order system with scarce measurements.
Section 4 gives the estimation of fractional order by
the MILM algorithm. Section 5 proves the
corversgence of the presented algorithms using the
stochastic identification theorems. Section 6 studies
two fractional order systems to show the
effectiveness of the proposed algorithms. Finally, we
offer some concluding remarks in Section 7.

2 Problem formulation
2.1 Calculation of fractional order system

In the past decades in system modeling and control,
fractional order systems have attracted continually an
increasing interesting among researchers [20-22].
The most commonly used in discrete cases are GL
fractional calculus [20], RL fractional calculus [21]
and Caputo fractional calculus [22], which is used in
this paper is the definition of GL calculus is
expressed as follows:

\[ \Delta^\alpha x(kh) = \frac{1}{h^\alpha} \sum_{j=0}^{k} (-1)^j \binom{\alpha}{j} x((k-j)) \]  \hspace{1cm} (1)

where \( \Delta^\alpha \) is the fractional order difference operator
of order \( \alpha \); \( x(kh) \) denotes a function of \( t = kh \),
which \( k \) is the \( k \)-th sampling, and \( h \) is assumed to be
equal to 1. The term \( \binom{\alpha}{j} \) is the binomial term
defined by:

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

According to Eqs. (1)-(2), we give the following recurrence equation:
\[
\begin{align*}
\begin{cases}
  w(0) = 1 \\
  w(j) = \left(1 - \frac{\alpha + 1}{j}\right) w(j-1) \quad \text{for } j = 1, 2, \ldots, k
\end{cases}
\end{align*}
\]

where
\[
w(j) = (-1)^j \left(\frac{\alpha}{j}\right)
\]

According to Eq. (4), Eq. (1) can be written as the following equation:
\[
\Delta^m x(k) = \sum_{j=0}^{k} w(j)x((k - j)).
\]

In this paper, we use Eq. (5) as the fractional calculation to study the modeling of the FOHS in subsequent sections.

### 2.2 Description of nonlinear fractional order model

In Fig. 1, we consider a nonlinear fractional order system which consists of a static nonlinear link in series with a linear fractional order system.

![Fig.1](image)

**Fig.1** The structure diagram of the nonlinear fractional order stochastic system.

In Fig. 1, \(u(k)\) and \(y(k)\) are the system input and the system output, respectively, \(u(k) \in R\) and \(y(k) \in R\); \(v(k)\) is the outside noise. The nonlinear system in Fig.1 is also called a fractional order Hammerstein system (FOHS), and its linear model is shown as follows:

\[
y(k) = \frac{B(z)}{A(z)} \bar{u}(k) + v(k)
\]

where \(\bar{u}(k)\) is the output of the static nonlinear link, and it can be represented as:

\[
\bar{u}(k) = \sum_{j=1}^{m} c_j f_j(u(k)) = f^T(u(k)) c = c^T f(u(k)),
\]

\[
f(k) = f(u(k)) = [f_1(u(k)), f_2(u(k)), \ldots, f_m(u(k))]^T \in R^n,
\]

where \(f_j(u(k))\) (\(j = 1, 2, \ldots, m\)) are the nonlinear basis functions and can be selected as polynomial functions, trigonometric functions, exponential functions, piecewise linear functions, etc. In addition, \(f(k) \in R^n\) and \(c \in R^m\) are information vector and parameter vector of nonlinear link, respectively.

According to Fig. 1, the noiseless output of the system can be written as:

\[
y_0(k) = \frac{B(z)}{A(z)} \bar{u}(k)
\]

where \(A(z)\) and \(B(z)\) are the denominator polynomial and the numerator polynomial, respectively,

\[
A(z) = 1 + a_1 z^{-\alpha} + a_2 z^{-2\alpha} + \cdots + a_n z^{-n\alpha},
\]

\[
B(z) = b_1 z^{-\beta_1} + b_2 z^{-\beta_2} + \cdots + b_m z^{-m\beta},
\]

\(\alpha_i\) and \(\gamma_j\) (\(i = 1, 2, \ldots, n_u; j = 1, 2, \ldots, n_y\)) are the corresponding fractional orders of the polynomials, \(\alpha_i \in R^+\) and \(\gamma_j \in R^+\), and \(z^{-1}\) is a unit backward shift operator with \(z^{-1} y(k) = y(k - 1)\).

When the fractional orders of the denominator polynomial and the numerator polynomial in (8) are completely different, the fractional order models of Eq. (8) are generally non-identical (disproportionate) order systems; Otherwise, each fractional order is an integer multiple of the base order (\(\alpha\) is order factor), \(\alpha_i = i\alpha\), \(\gamma_j = j\alpha\) \((i = 1, 2, \ldots, n_u; j = 1, 2, \ldots, n_y)\), such a model is defined as a homogeneous (proportionate) order system. In our paper, consider a proportional fractional order system. Then, Eq. (8) can be written as:

\[
y_0(k) = \frac{B(z)}{A(z)} u(k) = \frac{b_1 z^{-\beta} + b_2 z^{-2\beta} + \cdots + b_m z^{-m\beta}}{1 + a_1 z^{-\alpha} + a_2 z^{-2\alpha} + \cdots + a_n z^{-n\alpha}} u(k) \quad (9)
\]

By means of the discrete fractional order operator \(\Delta\) and the derivation, Eq. (9) can be derived as follows:

\[
y_0(k) = [1 - A(z)] y_0(k) + B(z) \bar{u}(k)
\]

\[
= -a_1 \Delta^\alpha y_0(k - 1) - a_2 \Delta^\alpha y_0(k - 2) - \cdots
\]

\[
- a_n \Delta^\alpha y_0(k - n_u) + b_1 \Delta^\alpha \bar{u}(k - 1) + \cdots
\]

\[
+ b_m \Delta^\alpha \bar{u}(k - n_y)
\]

\[
= -a_1 \Delta^\alpha y_0(k - 1) - a_2 \Delta^\alpha y_0(k - 2) - \cdots
\]

\[
- a_n \Delta^\alpha y_0(k - n_u) + c_1 \sum_{i=1}^{\infty} b_1 \Delta^\alpha f_1(u(k - i)) + \cdots
\]

\[
+ c_m \sum_{i=1}^{\infty} b_m \Delta^\alpha f_m(u(k - i))
\]

From Fig.1, the nonlinear function \(f(*)\) is connected in series with the linear dynamic fractional-order subsystem \(G(z) := \frac{B(z)}{A(z)}\). The first term coefficient of \(A(z)\) is 1. If the inner variable \(\bar{u}(k)\) is unmeasurable, then the parameters of \(f(*)\) and \(B(z)\) cannot be uniquely confirmed by \(u(k)\) and \(y(k)\). For this reason, let \(c_1 = 1\). Eq. (10) can be arranged as:
\[ y_i(k) = -\sum_{i=1}^{n} a_i \Delta^\alpha y_i(k-1) + \ldots \]
\[ + \sum_{i=1}^{n} b_i \Delta^\alpha f_1(u(k-i)) + \ldots \]
\[ + c_1 \sum_{i=1}^{n} b_i \Delta^\alpha f_2(u(k-i)) + \ldots \]
\[ + c_n \sum_{i=1}^{n} b_i \Delta^\alpha f_n(u(k-i)) \]

(11)

Herein, the overall output of FOHS with outside noise is arranged as follows,
\[ y(k) = -\sum_{i=1}^{n} a_i \Delta^\alpha y_i(k-1) + \ldots \]
\[ + \sum_{i=1}^{n} b_i \Delta^\alpha f_1(u(k-i)) + \ldots \]
\[ + c_1 \sum_{i=1}^{n} b_i \Delta^\alpha f_2(u(k-i)) + \ldots \]
\[ + c_n \sum_{i=1}^{n} b_i \Delta^\alpha f_n(u(k-i)) + \nu(k) \]

where \( \nu(k) \) is a white noise with zeros mean and variance \( \sigma^2 \). In our paper, the identification goal is to study the identification method of system parameters \( a, b, c \) and fractional order \( \alpha \) by measurement data points.

2.3 Fractional order systems with scarce measurements

In practice, due to some reasons, including sensor failure, hardware limitation and network transmission, etc., some sampled data in the system are lost. This system is called loss data system, such as multi-rate sampled data system. In [16], the loss data system is divided into two cases: one is missing data system, another is scarce measurement data system. In a certain time interval, compared with all sampled data, when the lost data points account for the majority, it is called the scarce measurement data system which is described as Fig. 2. When the lost data points account for a small part, it is called the missing data system which is described as Fig. 3.

For loss data, we design a mathematical description method to deal with them. First, define an integer sequence \( \{k_s: s=0,1,2,\ldots\} \) to satisfy,
\[ 0<k_0<k_1<k_2<\ldots<k_{s-1}<k_s<\ldots, \]
\[ k_s = k_s-1 + 1 \]

The observation data \( y(k_s) \) can be collected, that is, \( y(k_0), y(k_1), y(k_2), \ldots \) can be obtained and they are available data points. Under this situation, the observable data set is built as \( \{y(k_s): s=0,1,2,\ldots\} \), and the unobservable data set is built as \( \{y(k_s+1), y(k_s+2), \ldots y(k_s-1): s=0,1,2,\ldots\} \). The dynamic diagram of the system with scarce measurement data is shown in Fig. 2, where there is more missing data, the observable data set:
\[ \{y(k_0), y(k_1), y(k_2), y(k_3), y(k_4), y(k_5), y(k_6), y(k_7), \ldots\} \]
the unobservable data set:
\[ \{y(k_1), y(k_1), y(k_2), y(k_3), y(k_4), y(k_5), y(k_6), y(k_7), \ldots\} \]

The dynamic diagram of the system with missing sampling data is shown in Fig. 3, and the missing data is less, the observable data set:
\[ \{y(k_0), y(k_1), y(k_2), \ldots\} \]
the unobservable data set:
\[ \{y(k_1), y(k_1), y(k_2), \ldots\} \]

Fig. 2 The dynamic diagram of the system with scarce measurement data

Fig. 3 The dynamic diagram of the system with missing sampling data.

3. The identification of FOHS based on multi-innovation estimation algorithm

For convenience, we define the total parameter vector \( \theta \) and the information vector as \( \phi(k_s,\alpha) \) as follows:
\[ \theta := [a \ b \ c \otimes b]^T \in R^n, \quad n = n_s + mn_v, \]
\[ \phi(k_s,\alpha) := [\phi^1_1(k_s,\alpha), \phi^2_1(k_s,\alpha), \ldots, \phi^m_1(k_s,\alpha)]^T \in R^n, \]
where
\[ a := [a_1 \ a_2 \ a_n]^T \in R^n, \]
\[ b := [b_1 \ b_2 \ b_v]^T \in R^n, \]
\[ c := [c_1 \ c_2 \ c_m]^T \in R^n, \]
\[ \phi_0(k_s,\alpha) := [-\Delta^\alpha y_0(k_s-1), -\Delta^\alpha y_0(k_s-2), \ldots, -\Delta^\alpha y_0(k_s-n_v)]^T \in R^n \]
\[ \phi_j(k_\alpha, \alpha) := [\Delta^\alpha f_j(u(k_n - 1)), \Delta^\alpha f_j(u(k_n - 2)), \ldots, \Delta^\alpha f_j(u(k_n - n_j))]^T \in \mathbb{R}^{n_j}, \quad j = 1, 2, \ldots, m. \]

The symbol \( \otimes \) is the Kronecker product, if 
\[ A = [a_{ij}] \in \mathbb{R}^{p \times q}, \quad B = [b_{ij}] \in \mathbb{R}^{r \times s}, \]
then 
\[ A \otimes B = [a_{ij}b_{rs}] \in \mathbb{R}^{pq \times rs}. \]

Thus, Eqs. (11) and (12) are written as:
\[ y_0(k) = \phi^T(k, \alpha) \theta \]
\[ y(k) = y_0(k) + v(k), \]
\[ = \phi^T(k, \alpha) \theta + v(k). \]

The identification difficulty is that \( \phi(k, \alpha) \) in (14) contains the unknown inner variable \( y_0(k-i) \) and the unknown fractional order \( \alpha \). The solution is to replace \( y_0(k-i) \) and \( \alpha \) with their estimates \( \hat{y}_0(k-i) \) and \( \hat{\alpha} \). The estimation calculation of fractional order \( \alpha \) will be introduced in Section 4. Here, we give the following vector definitions:
\[ \phi_0(k, \hat{\alpha}) := [\Delta^\alpha \hat{y}_0(k-1), \Delta^\alpha \hat{y}_0(k-2), \ldots, \Delta^\alpha \hat{y}_0(k-n_\alpha)]^T \in \mathbb{R}^{n_\alpha}, \]
\[ n = n_\alpha + m \]
\[ \phi_0(k, \hat{\alpha}) := [-\Delta^\alpha \hat{y}_0(k-1), -\Delta^\alpha \hat{y}_0(k-2), \ldots, -\Delta^\alpha \hat{y}_0(k-n_\alpha)]^T \in \mathbb{R}^{n_\alpha}, \]
\[ \hat{\phi}_0(k, \hat{\alpha}) := [\Delta^\alpha f_j(u(k_n - 1)), -\Delta^\alpha f_j(u(k_n - 2)), \ldots, -\Delta^\alpha f_j(u(k_n - n_j))]^T \in \mathbb{R}^{n_j}. \]

Then the estimation of the parameter vector \( \theta \) can be solved by \( u(k_n - i) \), \( \hat{y}_0(k_n - i) \) instead of \( y_0(k_n - i) \). If the unknown \( \phi(k, \alpha) \) and \( \theta \) in Eq.(14) are replaced with \( \hat{\phi}_0(k, \hat{\alpha}) \) and \( \hat{\theta}(k) \), respectively, then the estimate of \( y_0(k_n - i) \) can be computed by:
\[ \hat{y}_0(k_n - i) = \hat{\phi}^T(k_n - i, \hat{\alpha}) \hat{\theta}(k). \]

Let \( \mu(k_n) \) be the step-size. According to the identification model in Eq. (14), a negative gradient search is used to minimize the following cost function,
\[ J(\theta) = [y(k) - \hat{\phi}^T(k, \hat{\alpha}) \theta]^2. \]

We can get the following gradient descent algorithm:
\[ \hat{\theta}(k) = \hat{\theta}(k_n) + \mu(k_n) [y(k) - \hat{\phi}^T(k, \hat{\alpha}) \hat{\theta}(k_n)] \]
\[ \hat{\theta}(k) = \hat{\theta}(k_n), \quad k \in T_e = \{k_n, k_n + 1, \ldots, k_{n-1}\}. \]

The estimated parameter vector \( \hat{\theta}(k) \) remains unchanged over the interval \( [k_n, k_{n-1}] \), i.e., \( \hat{\theta}(k) = \hat{\theta}(k_n) \) for \( k \in [k_n, k_n + 1, k_{n-1}] \). In order to improve the accuracy of parameter estimation and speed up the convergence rate, we expand the term \( y(k) - \hat{\phi}^T(k, \hat{\alpha}) \hat{\theta}(k_n) \) in Eq. (15) to an \( L \)-dimensional innovation vector,
\[ E(L, k, \hat{\alpha}) = [y(k) - \hat{\phi}^T(k, \hat{\alpha}) \hat{\theta}(k_n) - \hat{\phi}^T(k, \hat{\alpha}) \hat{\theta}(k_n)] \]

Define the input-output (information) matrix \( \Phi(L, k, \hat{\alpha}) \) and the stacked output vector \( Y(L, k) \) as:
\[ \Phi(L, k, \hat{\alpha}) = [\hat{\phi}(k, \hat{\alpha}), \hat{\phi}(k_{n-1}, \hat{\alpha}), \ldots, \hat{\phi}(k_1, \hat{\alpha})] \in \mathbb{R}^{n \times L}, \]
\[ Y(L, k) = [y(k), y(k_{n-1}), \ldots, y(k_1)]^T \in \mathbb{R}^{n \times 1}. \]

The \( L \)-dimensional innovation vector \( E(L, k, \hat{\alpha}) \) can be also expressed as:
\[ E(L, k, \hat{\alpha}) = Y(L, k) - \hat{\Phi}^T(L, k, \hat{\alpha}) \hat{\theta}(k_n) \]

Since \( \hat{\Phi}(1, k, \hat{\alpha}) = \hat{\phi}(k, \hat{\alpha}), \hat{\Phi}(1, k, \hat{\alpha}) = y(k) - \hat{\phi}^T(k, \hat{\alpha}) \hat{\theta}(k_n) \)
Eq.(15) can be equivalently written as:
\[ \hat{\theta}(k) = \hat{\theta}(k_n) + \mu(k_n) \hat{\Phi}(1, k, \hat{\alpha}) E(L, k, \hat{\alpha}) \]
Replacing the 1's in \( \Phi(L, k, \hat{\alpha}) \) and \( E(L, k, \hat{\alpha}) \) with \( L \) and taking \( \mu(k_n) = 1/r(k) \) and \( r(k_n) = r(k)+\|\hat{\Phi}(L, k_n, \hat{\alpha})\|^2 \), the fractional order \( \hat{\alpha} \) is known and a multi-innovation gradient descent algorithm (MIGD) is shown as follows:
\[ \hat{\theta}(k) = \hat{\theta}(k_n) + \frac{\hat{\Phi}(L, k_n, \hat{\alpha})}{r(k_n)} E(L, k, \hat{\alpha}) \]

For the initial values of the identified parameters, we choose \( \hat{\theta}(0) \) as a real vector with small entries, e.g., \( \hat{\theta}(0) = I_n / p_0 \), \( I_n \) represents a dimensional column vector whose entries are all 1.

After the estimated parameter vector \( \hat{\theta} \) is obtained, the estimates of the vector \( \hat{a} \) elements are the first \( n_a \) values of \( \hat{\theta} \), \( \hat{b} \) can be acquired from the \( n_a + 1 \) to \( n_a + n_b \) elements of \( \hat{\theta} \). For the
estimation vector $\hat{\theta}$, we notice that for $c_i$, we have $n_n$ estimates $\hat{c}_i$. Therefore, we use the mean value to be calculated as its estimate,

$$\hat{c}_j = \frac{1}{n_n} \sum_{i=1}^{n_n} \hat{c}_{ni}, \quad j = 2,3,...,m . \quad (29)$$

For the performance analysis of the developed algorithm, we will give some discussions in the next section.

4 The fractional order estimation of FOHS

For the identification of FOHS, both the fractional order $\alpha$ and the parameters $\{a, b, c\}$ need to be identified. The corresponding identification algorithms can be divided into two stages: one for order update, another for parameter estimation. The different identification algorithms of the two stages are carried out alternately.

Define an error function as follows:

$$e(k) = y(k) - \phi^T(k) \hat{\theta} . \quad (30)$$

Then the identification problem can be converted into minimizing the following objective function:

$$J = \frac{1}{N} \sum_{k=1}^{N} e^2(k) , \quad (31)$$

where $N$ is the length of the total sampling data.

For the objective function in Eq. (31), the Levenberg-Marquardt (L-M) algorithm is adopted to confirm the fractional order $\alpha$ as follows:

$$\hat{\alpha}^{(m+1)} = \hat{\alpha}^{(m)} - \left[ J' + \lambda I \right]^{-1} J' \hat{\theta}^{(m)} . \quad (32)$$

The update rule is based on the calculation of the gradient and Hessian $J'$ and $J''$ with respect to fractional order $\hat{\alpha}$ and $\lambda$ is a tuning parameter for the convergence. The calculation of the gradient and Hessian with respect to the fractional order $\hat{\alpha}$ ($J'_\alpha$ and $J''_\alpha$) can be performed as follows:

$$J'_\alpha = -\frac{2}{N} \left[ \frac{\partial \phi^T(k, \hat{\alpha})}{\partial \hat{\alpha}} \right] y(k) - \phi^T(k, \hat{\alpha}) \hat{\theta}$$

$$= -\frac{2}{N} \left[ \frac{\partial \phi^T(k, \hat{\alpha})}{\partial \hat{\alpha}} \right] y(k) - \phi^T(k, \hat{\alpha}) \hat{\theta}$$

$$= -\frac{2}{N} \left[ \frac{\partial \phi^T(k, \hat{\alpha})}{\partial \hat{\alpha}} \right] y(k) - \phi^T(k, \hat{\alpha}) \hat{\theta}$$

where $\sigma_y(k) / \hat{\alpha} = \frac{\partial \phi^T(k, \hat{\alpha})}{\partial \hat{\alpha}}$ is the output sensitivity function with respect to $\hat{\alpha}$. The sensitivity function is shown as follows:

$$\sigma_y(k) / \hat{\alpha} = \frac{\partial \phi^T(k, \hat{\alpha} + \delta \hat{\alpha}) - \phi^T(k, \hat{\alpha})}{\delta \hat{\alpha}}$$

(33)

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

The Hessian $J''_\alpha$ can be calculated by:

$$J''_\alpha = \frac{2}{N} \left[ \frac{\partial \phi^T(k, \hat{\alpha})}{\partial \hat{\alpha}} \right] \left[ \frac{\partial \phi^T(k, \hat{\alpha})}{\partial \hat{\alpha}} \right]^T . \quad (34)$$

Hence, the gradient $J'_\alpha$ and $J''_\alpha$ are shown as follows:

$$J'_\alpha = -\frac{2}{N} \left[ \frac{\partial \phi^T(k, \hat{\alpha})}{\partial \hat{\alpha}} \right] y(k) - \phi^T(k, \hat{\alpha}) \hat{\theta}$$

$$J''_\alpha = \frac{2}{N} \left[ \frac{\partial \phi^T(k, \hat{\alpha})}{\partial \hat{\alpha}} \right]^T \left[ \frac{\partial \phi^T(k, \hat{\alpha})}{\partial \hat{\alpha}} \right] . \quad (35a)$$

For the single-innovation L-M algorithm in Eq.(32), it can be considered as the single-innovation estimation algorithm by using the error $y(k) - \phi^T(k, \hat{\alpha}) \hat{\theta}$. For the single-innovation L-M algorithm, it can suffer from low convergence speed and low modeling accuracy. For this reason, we propose a multi-innovation L-M identification algorithm (MILM).

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

$$\Xi(L,k,\hat{\alpha})= [\sigma_y(k) / \hat{\alpha}, \sigma_y(k_2) / \hat{\alpha}, \ldots, \sigma_y(k_{s+1}) / \hat{\alpha}] \in R^s$$

(36)

Using the definitions of $\phi(L,k,\hat{\alpha})$ and $Y(L,k)$, the multi-innovation L-M identification algorithm are arranged as:

$$\hat{\alpha}^{(m+1)} = \hat{\alpha}^{(m)} - \frac{2}{N} \left[ \frac{\partial \Xi^T(L,k,\hat{\alpha})}{\partial \hat{\alpha}} \Xi(L,k,\hat{\alpha}) + \lambda I \right]^{-1} \left[ \frac{\partial \Xi^T(L,k,\hat{\alpha})}{\partial \hat{\alpha}} \Xi(L,k,\hat{\alpha}) + \lambda I \right] \hat{\alpha}^{(m)}$$

(37)

In this paper, the MIGD method is proposed to deal with the parameters identification of FOHS and the MILM algorithm is derived to estimate the fractional order. According to the theory of interactive estimation and the principle of hierarchical identification, the two proposed algorithms are used to identify the parameters and estimate the fractional order alternately. In each iteration, the parameter estimation depends on the previous estimation of fractional order. In turn, the estimation of fractional order is based on the parameter estimation of the previous iteration, and both of them perform a complete hierarchical calculation process. The whole steps of the proposed method can be summarized as follows:

Step 1: Let $m = 1, s=1$, and set the initial values $\hat{\theta}^0, \hat{\alpha}^0, p$ and $\delta \hat{\alpha}$;

Step 2: Collect the input-output data $\{u(k), y(k)\}$, $k = 1,2,...,N$;

Step 3: Construct the information matrix $\phi(L,k,\hat{\alpha})$, the stacked output vector $Y(L,k)$ using Eqs. (17a) and (17b), respectively, and compute the output fractional order sensitivity function $\sigma_y(k) / \hat{\alpha}$ using Eq. (33);
Step 4: Construct the stacked sensitivity function vector \( \mathbf{Z}(L, k, \hat{\alpha}) \) using Eq. (36);

Step 5: Compute \( J_k \) and \( J_n \) by using Eq. (35a) and Eq. (35b), and acquire the fractional order \( \hat{\alpha} \) using Eq. (37);

Step 6: Update the parameter estimation \( \hat{\theta}^{(m)}(k) \) using Eqs. (20)-(28), and compute \( \hat{c}_j \) (\( j = 2, 3, \ldots, m \)) using Eq. (29);

Step 7: Let \( s = s + 1, \) if \( s \leq N \) return Step 6, else go to Step 8.

Step 8: Compute the objective function in Eq. (31);

Step 9: If \( \left| J(\hat{\theta}^{(m+1)}) - J(\hat{\theta}^{(m)}) \right| \leq \varepsilon \), then set \( \hat{\theta} = \hat{\theta}^{(m)}, \hat{\alpha} = \alpha^{(m)} \) and \( J(\hat{\theta}) = J(\hat{\theta}^{(m)}) \), otherwise \( m = m + 1, \) \( s = 1 \), go to Step 3.

5 Performance analysis

Let us introduce some notations. The symbol \( [X] \) denotes the determinant of the square matrix \( X \), i.e.,
\[ |X| = \text{det}[X] . \]
The symbols \( \lambda_{\max}(X) \) and \( \lambda_{\min}(X) \) represent the maximum and minimum eigenvalues of the matrix \( X \), respectively. The following lemmas are required to give the main convergence results.

Lemma 1. (Martingale convergence theorem)
Assume \( \{T_k\} \) and \( \{\xi_k\} \) to be time sequences of non-negative random variables, and they are the incremental \( \sigma \) algebra sequences, satisfying the following inequality,
\[ E(T_{k+1} - T_k) \leq \varepsilon \xi_k, \text{ a.s.} \]

If \( \sum_{k=1}^{\infty} \xi_k < \infty \), then \( T_k \) almost certainly converges to a finite random variable \( C_1 \), namely
\[ \lim_{k \to \infty} T_k = C_1 < \infty . \]

Lemma 2. For the MIGD algorithm in Eqs. (20)-(28), the estimates \( \hat{\alpha} \) of fractional order is obtained from Eq. (37) and the positive constant \( \varepsilon \) (\( \varepsilon > 1 \)) are given, and the following inequality holds:
\[ \sum_{i=1}^{n} \left[ \phi(L, k, \hat{\alpha}) \right]^2 \leq \frac{1}{\varepsilon - 1} < \infty , \quad \varepsilon > 1 . \]

Proof. Using Eq. (23), the following inequality holds,
\[ \sum_{i=1}^{n} \left| \phi(L, k, \hat{\alpha}) \right|^2 \leq \frac{1}{\varepsilon - 1} < \infty , \quad \varepsilon > 1 . \]
Proof. From Eqs. (13)-(14), the following equation can be achieved,
\[ y(k_i - i) - v(k_i - i) = y_0(k_i - i) = \phi^T(k_i - i, \alpha)\theta. \] (44)

Using Eq. (9), we have
\[ A(z) y_0(k_i - i) = B(z) v(k_i - i). \]

or
\[ A(z)[y(k_i - i) - v(k_i - i)] = B(z) v(k_i - i). \]

Using the above equation and Eq. (24), we have
\[ A(z)[y(k_i - i) - \hat{\phi}^T(k_i - i, \hat{\alpha})\hat{\theta}(k_i) - v(k_i - i)] \]
\[ = -A(z)\hat{\phi}_0^T(k_i - i, \hat{\alpha})\hat{\theta}(k_i) + A(z)[y(k_i - i) - v(k_i - i)] \]
\[ = -A(z)\hat{\phi}_0^T(k_i - i, \hat{\alpha}) + B(z) v(k_i - i) \]
\[ = -\hat{\gamma}_0(k_i - i) - a\Delta^s\hat{\phi}_0(k_i - i - 1) - a\Delta^s\hat{\phi}_0(k_i - i - 2) - ... \]
\[ + b\Delta^s\alpha u(k_i - i - n) + b\Delta^s\alpha u(k_i - i - 1) + b\Delta^s\alpha u(k_i - i - 2) + ... \]
\[ = -\hat{\gamma}_0(k_i - i) + \hat{\phi}^T(k_i - i, \hat{\alpha})\theta \]
\[ = -\hat{\phi}^T(k_i - i, \hat{\alpha})\theta(k_i) \] (45)

Taking \( i = 0, k_1, k_2, ..., k_{L - 1} \) and combining these equations in (41), (43), (45) give
\[ A(z)[E(L_k, \hat{\alpha}) - V(L_k)] \]
\[ = \hat{\phi}^T(L_k, \hat{\alpha})\theta(k_i) = \bar{Y}(L_k, \hat{\alpha}) \] (46)

This proves Lemma 4.

Define
\[ S(k_i) = \sum_{r = r(k_i)}^{2} \frac{1}{r(k_i)} \left[ \bar{Y}(L_k, \hat{\alpha}) - \frac{1}{2} E(L_k, \hat{\alpha}) - V(L_k) \right] \]
\[ \times \{E(L_k, \hat{\alpha}) - V(L_k)\} + C_2, C_2 > 0 \]

Theorem 1. For the FOHS in (14) and the MIGD algorithm in (20)-(28), the fractional order \( \hat{\alpha} \) is obtained from Eq. (37), assume that the noise signal \( \{v(k_i)\} \) is a white noise signal with zero mean and variance \( \sigma_v^2 \), and satisfies

(A1) \( \mathbb{E}[v(k_i)] = 0 \), \( \mathbb{E}[v(k_i)^2] = \sigma_v^2 \);

(A2) \( A(z) = \frac{1}{2} \) is strictly positive real.

Then
\[ \lim_{r \to \infty} \hat{\phi}(k_i) = \mathbb{E}(k_i) = C_2, \text{ m.s.} \]

Proof. Substituting Eq. (20) into Eq. (42) yields
\[ \hat{\phi}(k_i) = \hat{\phi}(k_i) + \frac{\phi(L_k, \hat{\alpha})}{r(k_i)} E(L_k, \hat{\alpha}) \] (47)

Taking the norm of both sides and using Eqs. (46), (47), \( r(k_i) \geq \|\phi(L_k, \hat{\alpha})\| \) and \( \|X\| \leq \|X\|\|Y\| \) have
\[ \|\hat{\phi}(k_i)\|^2 = \|\hat{\phi}(k_i)\|^2 + \frac{\phi(L_k, \hat{\alpha})}{r(k_i)} E(L_k, \hat{\alpha}) \]
\[ = \|\hat{\phi}(k_i)\|^2 + \frac{2}{r(k_i)} \|\phi(L_k, \hat{\alpha})\|^2 E(L_k, \hat{\alpha}) + ... \]
\[ + \frac{\phi(L_k, \hat{\alpha})}{r(k_i)} E(L_k, \hat{\alpha}) \]
\[ \leq \|\hat{\phi}(k_i)\|^2 + \frac{2}{r(k_i)} \|\phi(L_k, \hat{\alpha})\|^2 E(L_k, \hat{\alpha}) + ... \]
\[ + \frac{\phi(L_k, \hat{\alpha})}{r(k_i)} E(L_k, \hat{\alpha}) \]

Adding \( S(k_i) \) to both sides and letting
\[ T(s) = \|\hat{\phi}(k_i)\|^2 + S(k_i) \]
\[ T(s) = T(s - 1) + \frac{\phi(L_k, \hat{\alpha})}{r(k_i)} E(L_k, \hat{\alpha}) + ... \]
\[ - \frac{2}{r(k_i)} \|\phi(L_k, \hat{\alpha})\|^2 E(L_k, \hat{\alpha}) + ... \]

Since \( r(k_i) \), \( \phi(L_k, \hat{\alpha}) \), \( \bar{Y}(L_k, \hat{\alpha}) \), \( E(L_k, \hat{\alpha}) - V(L_k) \) are statistically independent of the white noise vector \( V(L_k) \), taking the
expectation have
\[
E[T(s)] \leq E[T(s-1)] + E \left[ \frac{\| \hat{\Phi}(L,k_s,\hat{\alpha})V(L,k_s) \|^2}{r^2(k_s)} \right] .
\] (49)

Using Lemma 2, the sum of the last term on the right-hand side from \( s = 1 \) to \( s = \infty \) is finite, i.e.,
\[
\sum_{s=1}^{\infty} E \left[ \frac{\| \hat{\Phi}(L,k_s,\hat{\alpha})V(L,k_s) \|^2}{r^2(k_s)} \right] \lesssim \sum_{s=1}^{\infty} \sigma^2 \lesssim \infty
\] (50)

Applying Lemma 1 to Eq. (49), we conclude that \( T(s) = \left[ \hat{\theta}(k_s) \right]^2 + S(k_s) \) converges m.s. to a constant \( C_2 \). This proves Theorem 1.

Theorem 1 gives the upper limit of parameter estimation error. That is, the upper limit of parameter estimation error is bounded in the mean square sense. Using Eq. (44), the following equation is obtained,
\[
\tilde{Y}(L,k_s,\hat{\alpha}) - \frac{1}{2} [E(L,k_s,\hat{\alpha}) - V(L,k_s)] = \left[ A(z) - \frac{1}{2} \right] \left[ E(L,k_s,\hat{\alpha}) - V(L,k_s) \right] .
\] (51)

Since \( A(z) - \frac{1}{2} \) is strictly positive real, according to the discrete positive real lemma from [23], there exists a positive constant \( C \) such that we have
\[
\sum_{s=1}^{N} \frac{C}{r(k_s)} \|E(L,k_s,\hat{\alpha}) - V(L,k_s)\|^2 \leq S(k_s) .
\] (52)

The estimation results of each iteration can affect the next parameter identification of fractional order system because the two proposed algorithms are alternately implemented in turn. If the fractional order estimate approximates the true value, the parameter estimates of fractional order system will also approach the true values. Thus, the upper limit of parameter estimation error is very small.

6 Example simulation
6.1 Simulation example 1

Let us consider the fractional nonlinear system of the commensurate order \( \alpha = 0.6 \) with \( n_s = n_v = 2 \), \( h = 1 \). The discrete-time transfer function is shown as

\[
y_0(k) = \frac{B(z)}{A(z)} u(k) = \frac{b_1 z^{-\gamma} + b_2 z^{-\eta}}{1 + a_1 z^{-\mu} + a_2 z^{-\nu}} u(k),
\] (53)

\[
= -0.4 z^{-0.6} - 0.2 z^{-1.2} - 0.4 z^{-0.2} - 0.2 z^{-0.6}, \quad \alpha = 0.6.
\]

The overall output of the fractional order system is shown as follows:
\[
y(k) = -a_1 \Delta^{0.6} y_0(k-1) - a_2 \Delta^{1.6} y_0(k-2) + ... + b_1 \Delta^{0.6} u(k-1) + b_2 \Delta^{0.6} u(k-2) + ... + c_1 \Delta^{1.6} u(k-1) + c_2 \Delta^{0.6} u(k-2) + v(k)
\] (56)

where the real parameter vector are:
\[
a = [a_1, a_2]^T = [0.1, 0.2]^T,
\]
\[
b = [b_1, b_2]^T = [-0.4, -0.2]^T,
\]
\[
c = [c_1, c_2]^T = [1, 0.5]^T.
\]

Thus, the parameter vector to be estimated is \( \theta = [a_1, a_2, b_1, b_2, c_1, c_2]^T = [0.1, 0.2, -0.4, -0.2, -0.2, -0.1]^T \).

The input \( u(k) \) is taken as a persistent excitation random signal with zero mean and unit variance \( \sigma^2 \). We collected the input-output data \{ \( u(k_s) \), \( y(k_s) \) \}, where \( k_s = k_{max} - k \), is a uniform random integer between 1 and a maximum value \( N=2 \). Taking \( L = 1, 3, 5 \) and applying the MIGD algorithm and the MILM algorithm, the parameter estimates and their relative errors \( \delta = (\hat{\theta} - \theta)^T (\hat{\theta} - \alpha) \) are shown in Table 1. 1, Table 1. 2 Table 1. 2 show the parameter estimates and relative errors under different noise signal ratios \( \delta_n = 14.13\%, 18.27\% \), different data lengths \( K = 5000, 10000, 20000 \) and different information lengths \( L = 1, 3, 5 \). As shown in Fig. 4, under the two types of noise-signal ratios, the lower the noise-signal ratio, the smaller the parameter estimation error, indicating that the proposed algorithm can work better for low signal-to-noise ratios.
### Table 1. The parameter estimation and the relative errors of the FOS with $L = 1, 3, 5$ ($\delta_{\omega} = 14.13\%$)

| $L$ | $K$  | $a_1$  | $a_2$  | $b_1$  | $b_2$  | $c_1 \times b_1$ | $c_2 \times b_2$ | $\alpha$ | $\delta$ |
|-----|------|--------|--------|--------|--------|-----------------|-----------------|---------|---------|
| 1   | 5000 | 0.1223 | 0.2462 | -0.3980 | -0.1907 | -0.1984 | -0.0897 | 0.5489 | 0.0054 |
| 10000 | 0.1244 | 0.2191 | -0.4103 | -0.2152 | -0.1885 | -0.0899 | 0.5777 | 0.002   |
| 20000 | 0.1298 | 0.2003 | -0.4085 | -0.2243 | -0.2012 | -0.1041 | 0.5833 | 0.0018 |
| 3   | 5000 | 0.0749 | 0.1975 | -0.4111 | -0.1936 | -0.1927 | -0.0844 | 0.5892 | 0.0012 |
| 10000 | 0.0839 | 0.1643 | -0.4006 | -0.2232 | -0.2043 | -0.1159 | 0.6742 | 0.0079 |
| 20000 | 0.1474 | 0.1986 | -0.3965 | -0.1996 | -0.1891 | -0.0922 | 0.5578 | 0.0042 |
| 5   | 5000 | 0.1210 | 0.2144 | -0.3900 | -0.2105 | 0.1889 | -0.1015 | 0.6202 | 0.0014 |
| 10000 | 0.1077 | 0.1496 | -0.3357 | -0.2165 | 0.1562 | -0.0982 | 0.8668 | 0.0804 |
| 20000 | 0.0680 | 0.2026 | -0.3054 | -0.1726 | 0.1652 | -0.0857 | 0.7496 | 0.0345 |
| True values | 0.1 | 0.2 | -0.4 | -0.2 | -0.1 | 0.6 |

### Table 2. The parameter estimation and the relative errors of the FOS with $L = 1, 3, 5$ ($\delta_{\omega} = 18.27\%$)

| $L$ | $K$  | $a_1$  | $a_2$  | $b_1$  | $b_2$  | $c_1 \times b_1$ | $c_2 \times b_2$ | $\alpha$ | $\delta$ |
|-----|------|--------|--------|--------|--------|-----------------|-----------------|---------|---------|
| 1   | 5000 | 0.1020 | 0.1548 | -0.3693 | -0.2108 | -0.1927 | -0.1068 | 0.6703 | 0.0081 |
| 10000 | 0.1114 | 0.1751 | -0.3608 | -0.1884 | -0.1834 | -0.0847 | 0.6192 | 0.0033 |
| 20000 | 0.0656 | 0.1765 | -0.3911 | -0.2196 | -0.1836 | -0.0921 | 0.6432 | 0.0079 |
| 3   | 5000 | 0.1318 | 0.2244 | -0.3729 | -0.1790 | -0.1651 | -0.0789 | 0.5808 | 0.0048 |
| 10000 | 0.1296 | 0.2583 | -0.4138 | -0.1494 | -0.1904 | -0.0584 | 0.4003 | 0.0487 |
| 20000 | 0.1228 | 0.1399 | -0.3791 | -0.2352 | -0.1949 | -0.1213 | 0.7204 | 0.0207 |
| 5   | 5000 | 0.0582 | 0.1388 | -0.2825 | -0.1160 | 0.1565 | -0.0478 | 0.5724 | 0.0317 |
| 10000 | 0.1236 | 0.1196 | -0.2827 | -0.1785 | 0.1351 | -0.0841 | 0.8706 | 0.0989 |
| 20000 | 0.1034 | 0.1488 | -0.3035 | -0.2061 | 0.1008 | -0.0475 | 0.9551 | 0.1507 |
| True values | 0.1 | 0.2 | -0.4 | -0.2 | -0.1 | 0.6 |

### Fig. 4 Mean square of parameter estimation error under different $K, L$ and $\delta_{\omega}$

Fig. 5 is a box plot of system parameter estimation error with different information length and data length under the noise signal ratio of $\delta_{\omega} = 14.13\%$. The narrower the cabinet and the closer to the 0 scale line, the smaller the parameter estimation error. It can be clearly seen that when $K = 5000$ and $L = 3$, the parameter estimation error is the smallest, and the parameter estimation curve is shown in the Fig. 6, different color curves correspond to corresponding
color coordinates. Therefore, combined with the method proposed in this paper, the estimated parameters when the number of information is 3, the data length is 5000, and the noise-to-signal ratio is 14.13% are used to model the system. The estimated output and actual output of the system obtained by the model are shown in Fig. 7, comparing the estimated output with the actual output and amplifying it partially, it can be clearly seen that the actual output data overlaps with the estimated output data and is a complete fit.

\[
\bar{y}(k) = c_1 u(k) + c_2 u^2(k) + c_3 u^3(k) + c_4 u^*(k) = u(k) + 0.4 u^2(k) + 0.35 u^3(k) + 0.2 u^*(k)
\]

\[
y(k) = y_0(k) + v(k),
\]

where

\[
A(z) = 1 + 0.1 z^{-0.8} + 0.2 z^{-1.6},
\]

\[
B(z) = -0.4 z^{-0.8} - 0.2 z^{-1.6}, \quad \alpha = 0.8.
\]

The overall output of the fractional order system is shown as follows:

\[
y(k) = -a_i \Delta^{0.8} y_0(k-1) - a_j \Delta^{0.8} y_0(k-2) + ...
\]

\[
+ \sum b_i \Delta^{0.8} u(k-i) + \sum c_i \Delta^{0.8} u^2(k-i) + 
\]

\[
+ \sum e_i \Delta^{0.8} u^3(k-i) + \sum f_i \Delta^{0.8} u^*(k-i) +
\]

\[
+ \sum g_i \Delta^{0.8} u(k-i) + \sum h_i \Delta^{0.8} u^2(k-i) + ...
\]

(58)

where the real parameter vector are:

\[
a = [a_1, a_2]^T = [0.1, 0.2]^T,
\]

\[
b = [b_1, b_2]^T = [-0.4, -0.2]^T,
\]

\[
c = [c_1, c_2, c_3, c_4]^T = [1, 0.4, 0.35, 0.2]^T.
\]

Thus, the parameter vector to be identified is:

\[
\theta = [a_1, a_2, c_1, c_2, c_3, c_4, b_1, b_2, d_1, d_2, e_1, e_2, f_1, f_2, g_1, g_2, h_1, h_2, i_1, i_2, j_1, j_2, k_1, k_2]^T
\]

\[
= [0.1, 0.2, -0.4, -0.2, -0.16, -0.08, -0.14, -0.07, -0.08, -0.04]^T
\]

For the FOS in (58), we take dual-rate sampling to deal with it. All the inputs \( \{u(k) : k = 0, 1, 2, ...\} \) are available, but only scarce outputs \( \{y(kq) : k = 0, 1, 2, ...\} \) are available \((q \geq 2\) is a positive integer). In such system, the input-output data available are \( \{u(k), y(kq) : k = 0, 1, 2, ...\} \). Thus, the inter-sample outputs (or missing outputs) \( y(kq + i) : i = 1, 2, ..., q-1 \) are not available. In simulation, the dual-rate coefficient \( q \) is selected as 2. It can be seen from example 1 that the method proposed in this paper has a better effect on de-noise signal ratio, so this example uses low noise signal ratio \( \delta = 14.13\% \) for simulation. Taking \( L = 1, 3, 5 \) and applying the proposed algorithms in our paper, the parameter estimates and their relative errors \( \delta = (\theta - \hat{\theta})^2 + (\alpha - \hat{\alpha})^2 \) are shown in Table. 3. The relative error diagram of the parameter estimation with different innovation lengths \( L=1, 3, 5 \) and data lengths \( K = 5000, 10000, 20000 \) are shown in Figs. 8, 9.

![Parameter estimation curve](image)

**Fig. 6** Parameter estimation curve \((L = 3, K = 5000, \delta_m = 14.13\%)

![Simulated and estimated outputs](image)

**Fig. 7** Simulated and estimated outputs

6.2 Simulation example 2

Consider a following discrete-time fractional order Hammerstein system with the commensurate order \( \alpha = 0.8 \).

\[
y_0(k) = \frac{B(z)}{A(z)} \bar{y}(k) = \frac{b_1 z^{-\gamma_1} + b_2 z^{-\gamma_2}}{1 + a_1 z^{-\gamma} + a_2 z^{-2\gamma}} \bar{y}(k)
\]

\[
= \frac{-0.4 z^{-0.8} - 0.2 z^{-1.6}}{1 + 0.1 z^{-0.8} + 0.2 z^{-1.6}} \bar{y}(k)
\]

\[
\bar{y}(k) = c_1 u(k) + c_2 u^2(k) + c_3 u^3(k) + c_4 u^*(k)
\]

\[
y(k) = y_0(k) + v(k),
\]
Table 3. The parameter estimation and the relative errors of the FOS with $L = 1, 3, 5$

| L  | $K$  | $a_1$ | $a_2$ | $b_1$ | $b_2$ | $c_2 \times b_1$ | $c_3 \times b_2$ | $c_2 \times b_1$ | $c_1 \times b_2$ | $c_0 \times b_3$ | $\alpha$ | $\delta$ |
|----|------|-------|-------|-------|-------|------------------|------------------|------------------|------------------|------------------|--------|--------|
| 1  | 5000 | 0.1332| 0.1861| -0.3522| -0.1702| -0.1146         | -0.0655          | -0.1242         | -0.0633          | -0.0849         | -0.0318 | 0.7247 | 0.0128 |
| 10000| 0.0654| 0.1604| -0.3502| -0.1618| -0.1677| -0.0758          | -0.1060          | -0.0498          | -0.0689          | -0.0343         | 0.8948 | 0.0175 |
| 20000| 0.0727| 0.2423| -0.4164| -0.1703| -0.1835| -0.0924          | -0.1156          | -0.0834          | -0.0651          | -0.0220         | 0.7618 | 0.0072 |
| 3  | 5000 | 0.0871| 0.1944| -0.3267| -0.1548| -0.1178         | -0.0399          | -0.1480          | -0.0410          | -0.0842         | -0.0414 | 0.7890 | 0.0121 |
| 10000| 0.1387| 0.1528| -0.4688| -0.1497| -0.1261| -0.0562          | -0.1672          | -0.1370          | -0.0954          | -0.0511         | 0.6969 | 0.0289 |
| 20000| 0.1140| 0.0966| -0.4386| -0.2157| -0.1508| -0.0431          | -0.1179          | -0.1272          | -0.0922          | -0.0806         | 0.9065 | 0.0310 |
| 5  | 5000 | 0.1266| 0.2040| -0.4201| -0.2011| -0.1309         | -0.0858          | -0.1334          | -0.0572          | -0.0849         | -0.0390 | 0.7600 | 0.0038 |
| 10000| 0.0117| -0.00052| -0.3664| -0.1838| -0.1667| -0.0520          | -0.1345          | 0.0431           | -0.0113         | -0.0095         | 0.0875 | 0.0742 |
| 20000| 0.0014| -0.00049| -0.4114| -0.1767| -0.1435| -0.0394          | -0.0988          | -0.0328          | -0.0309         | -0.0234         | 0.9992 | 0.0980 |

True values: $a_1 = 0.1$, $a_2 = 0.2$, $b_1 = -0.4$, $b_2 = -0.2$, $c_2 \times b_1 = -0.16$, $c_3 \times b_2 = -0.08$, $\alpha = -0.14$, $\delta = -0.07$.

Figs. 8, 9 show the relative error diagrams of parameter estimates with different innovation lengths $L = 1, 3, 5$ and data lengths $K = 5000$, 10000, 20000, combined with the error radar diagram of Fig. 8 and the error box plot of Fig. 9, the parameter estimation errors caused by different information lengths and different data lengths are different. It can be seen from the figures that the parameter estimation error is the smallest when $L = 5$ and $K = 5000$. The parameter estimation curve is shown in Fig. 10. Therefore, the estimated parameters at this time are combined with the method in this paper, and the estimated output of the system and the actual output obtained are shown in Fig. 11. The partial magnification shows that the output is basically fitted, which verifies the effectiveness of the method proposed in this paper.
7 Conclusion

In this paper, we have studied the identification problems for FOHS with scarce measurements. The MIGD algorithm and the MILM algorithm are proposed to deal with the identification of FOHS systems with scarce measurements and fraction order, and the convergence performance of MIGD algorithm is analyzed in the random statistical frame. Choose the appropriate number of information and data length, and use the method proposed in this paper to identify system parameters with better accuracy. For the multi-rate FOHS systems with multi-variables, they can be regarded as a class of fraction order systems with scarce measurement. In the future, we will study the multi-rate FOHS with multi-variables.

Declaration of conflicting interests

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

Funding Statement

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).

Reference

1. Zou S Z, Zhang L, Hu X S, Wang Z P etc., A review of fractional-order techniques applied to lithiumion batteries, lead-acid batteries and supercapacitors. Journal of Power Sources., 390(1), 286-296 (2018)
2. Chen S Z, Liu F W, Turner I, Hu X L. Numerical inversion of the fractional derivative index and surface thermal flux for an anomalous heat conduction model in a multi-layer medium. Applied Mathematical Modeling. 59(3), 514-526 (2018)
3. Victor S, Mali R, Garnier H, Oustaloup A. Parameter and differentiation order estimation in fractional models. Automatica, 49(4), 926-935 (2013)
4. Deng J. Higher-order stochastic averaging for a SDOF fractional viscoelastic system under bounded noise excitation. Journal of Frankin Institute. 354(17), 7917-7945 (2017)
5. Shalaby R, Mohammad E H, Belal A Z. Fractional order modeling and control for under-actuated inverted pendulum. Communication of Nonlinear Science Numerical Simulation., doi: 10.1016/j.cnsns.2019.02.023. (2019)
6. Abuaihsta T, Kertzscher J. Fractional-order modeling and parameter identification of electrical coils. Fract. Calc. Appl. Anal. 22(1), 193-216 (2019)
7. 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. International Journal of System Science. 47(1), 122-134 (2016)
8. Cui R Z, Wei Y H, Chen Y Q, Cheng S S, Wang Y. An innovative parameter estimation for fractional-order systems in the presence of outliers. Nonlinear Dynamics. 89(1), 453-463 (2017)
9. Liu D Y, Laleg-Kirati T M, Gibaru O, Perruquetti W. Identification of fractional order systems using modulating functions method. American Control Conference, IEEE Washington, USA, 1679-1684. (2013)
10. Dai Y, Wei Y H, Hu Y S, Wang Y. Modulating function-based identification for fractional order systems. Neurocomputing., 173, 1959-1966 (2016)
11. Victor S, Mali R. Parameter and differentiation order estimation in fractional models. Automatica. 49(4), 926-935 (2013)
12. Mali R, Aoun M, Oustaloup A. Synthesis of fractional kautz-like basis with two periodically repeating complex conjugate modes. Control, Communication and Signal Processing, First International Symposium. (2004)
13. Tang Y G, Liu H F, Wang W W, Lian Q S, Guan X P. Parameter identification of fractional order systems using block pulse functions. Signal Process. 107, 272-281
14. Gildas B, Alina V, Guillaume B, Mariana M, Pinheiro M. Order and parameter identification for a non-integer-order model of an EGG system. IFAC Papers on Line., 51(15), 772-777 (2018)

15. Marcin B, Piotr Duch, Tenreiro J A, Piotr Ostalczyk, Dominik S. Commensurate and non-commensurate fractional-order discrete models of an electric individual-wheel drive on an Autonomous platform. Entropy. 22, 300-313 (2020)

16. Ding F, Liu G L, Liu X P. Parameter estimation with scarce measurements. Automatica, 47(3), 1646-1655 (2011)

17. Ding F, Ding J. Least squares parameter estimation for systems with irregularly missing data. International Journal of Adaptive Control and Signal Processing, 24(7), 540–553 (2010)

18. Broersen, P. M. T., Waele, S., Bos, R. Autoregressive spectral analysis when observations are missing. 40(9), 1495–1504 (2004)

19. Jin Q B, Wang Z, Liu X P. Auxiliary model-based inter-varying multi-innovation least squares identification for multivariable OE-like systems with scarce measurements. Journal of Process Control. 35(1), 154-168 (2015)

20. Mohammad J M, Hamed M, Mohammad T. Recursive identification of multiple-input single-output fractional-order Hammerstein model with time delay. Applied Soft Computing., 70(3), 486-500 (2018)

21. Qi Z D, Sun Q, Ge W P, He Y K. Nonlinear modeling of PEMFC based on fractional order subspace identification. Asian Journal of Control. 22(5), 1892-1900 (2020)

22. Wang J C, Wei Y H, Liu T Y, Li A, Wang Y. Fully parametric identification for continuous time fractional order Hammerstein systems. Journal of the Franklin Institute. 357(1), 651-666 (2020)

23. Ding F. System identification: multi-innovation identification theory and method. Science Press. Beijing. (2016)