Iterative parameter identification for fractional-order block-oriented nonlinear systems with application to a battery model

Yan Ji*, Junwei Wang, Xiangxiang Meng

College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, PR China.

Abstract

This paper investigates parameter and order identification of a class of block-oriented nonlinear systems. By using the hierarchical identification principle, the system is divided into two subsystems, which are a linear block system and a nonlinear block system. For the purpose of solving the difficulty of estimating two sets of parameter vectors, the over-parameterization method and the key item separation technique are used, respectively. Therefore, a two-stage over-parameterization gradient-based iterative algorithm and a key term separation two-stage gradient-based iterative algorithm are derived. The simulation results indicate that the proposed algorithms are effective. Finally, the proposed method is evaluated through a battery model. The results show well agreement with the real system outputs.

Keywords: Parameter estimation, Key term separation, Over-parameterization, Hierarchical identification, Gradient search, Battery model

1. Introduction

Nonlinear phenomena widely exist in various systems, e.g., photovoltaic cell models [1], lithium-ion batteries [2], and stirred-tank reactors [3]. On the identification of nonlinear systems, Ghani et al. presented a new method for photovoltaic cells, their method is based on the examination of the power voltage data from which a system of five residual equations are derived and solved via the multi-variable Newton-Raphson method [4]. For a lithium-ion battery model, Li et al. combined the bias compensation recursive least squares algorithms and the extended Kalman filter to alleviate the impact of the noises [5]. Moreover, some methods have been proposed for the nonlinear systems identification, e.g., the maximum likelihood methods [6, 7, 8], the least squares methods [9, 10, 11], and the gradient search methods [12, 13, 14].

The block-oriented nonlinear systems are popular for their simplicity and ability to accurately describe a wide variety of nonlinear systems [15, 16, 17, 18]. The block structures, like Hammerstein systems, have the ability of flexible combination of various static nonlinear elements and various dynamic linear elements. Wang et al. considered switch detection and robust identification for slowly switched Hammerstein systems, and proposed a two-identifier-based switch detection scheme to improve the precision of the time-invariant parameter estimation [19]. For the multivariable Hammerstein systems, it is difficult to parameterize the system into an auto-regression form to which the standard least square method can be applied. Wang et al. maximized the logarithmic likelihood function about each parameter vector to get their estimates [20]. It is worth noting that some complex systems are difficult to be accurately described by the traditional integer-order systems. This motivates us to study the parameter identification of the fractional-order systems.

In many practical nonlinear processes, the fractional behavior cannot be ignored. Moreover, the existence of orders increase the difficulty of system identification, which makes many scholars study the fractional-order systems [21, 22, 23, 24]. Hu et al. established a improved second-order equivalent circuit model based on the

*Corresponding author
Email address: `yjichina@163.com` (Yan Ji*)

fractional calculus theory, and used the mixed-swarm-based cooperative particle swarm optimization algorithm to identify the parameters of the equivalent circuit model [25]. For fractional-order Hammerstein-Wiener systems, an output-error approach was developed by using the robust Levenberg-Marquardt algorithm, and was applied to the benchmark test experiment of the robot arm [26]. In the application of heating processes, Hammar et al. researched the parameterization of the Hammerstein system by transforming the fractional-order polynomial nonlinear state-space model [27]. Sersour et al. used the heuristic particle swarm optimization to identify unmeasurable internal variables combined with the key item separation technique [28].

Considering the nonlinearity of the channel, we use the hierarchical identification principle [29, 30, 31] to divide the original system into several subsystems or sub-identification models. For the purpose of solving the difficulty of estimating the two sets of parameter vectors, we use the over-parameterization method and the key term separation technique, respectively. The gradient search is basic for nonlinear optimization problems [32, 33, 34], a two-stage gradient-based iterative algorithm was developed for the fractional-order block-oriented nonlinear systems. The main contributions of this paper are as follows.

- By using the model decomposition technique, a fractional-order block-oriented nonlinear system is decomposed into two subsystems, which are a linear subsystem and a nonlinear subsystem.
- Based on the iterative search, we propose a two-stage over-parameterization gradient-based iterative (2S-OP-GI) algorithm, and a key term separation two-stage gradient-based iterative (KT-2S-GI) algorithm. In addition, the comparison of computation amount between the two algorithms is given.
- Two examples are given, in which the two-stage over-parameterization gradient-based iterative algorithm is applied to a battery model by using the proposed hierarchical identification method.

This paper is organized as follows. Section 2 describes the identification model for a class of fractional-order nonlinear systems. A 2S-OP-GI algorithm is presented in Section 3. Section 4 utilizes the key term separation technique to deduce a KT-2S-GI algorithm. Section 5 analyzes the computational efficiency of the 2S-OP-GI algorithm and the KT-2S-GI algorithm. The numerical examples illustrate the performance of the proposed methods in Section 6. Finally, some conclusions are given in Section 7.

2. System description and identification model

Let us introduce some symbols. $A =: X$ or $X := A$ represents that X is defined by A. The symbol $\boldsymbol{I}_n$ stands for an identity matrix of size $n \times n$. $\mathbf{1}_n$ represents an n-dimensional column vector whose elements all are 1, that is, $\mathbf{1}_n := [1, 1, \cdots, 1]^{\mathrm{T}} \in \mathbb{R}^n$. The superscript T denotes the matrix/vector transpose. Consider the following fractional-order block-oriented nonlinear system,

$$y(l) = \frac{B(z)}{A(z)} f(u(l)) + e(l), \tag{1}$$

where $y(l)$ is the output of the system, $u(l)$ is the input of the system, $e(l) := \frac{1}{C(z)} v(l)$ is an autoregressive noise and $v(l)$ is the random white noise with zero mean and variance σ^2, $f(\cdot)$ is the nonlinear element of the system, $A(z)$, $B(z)$ and $C(z)$ are three fractional-order polynomials in the unit backward shift operator $[z^{-1}y(l) = y(l-1),\ zy(l) = y(l+1)]$, define as

$$\begin{aligned} A(z) &:= a_{n_a} z^{-\alpha_{n_a}} + a_{n_a-1} z^{-\alpha_{n_a-1}} + \cdots + a_1 z^{-\alpha_1} + 1, \\ B(z) &:= b_{n_b} z^{-\alpha_{n_b}} + b_{n_b-1} z^{-\alpha_{n_b-1}} + \cdots + b_1 z^{-\alpha_1}, \\ C(z) &:= c_{n_c} z^{-\alpha_{n_c}} + c_{n_c-1} z^{-\alpha_{n_c-1}} + \cdots + c_1 z^{-\alpha_1} + 1, \end{aligned}$$

where α_{n_a}, α_{n_b} and α_{n_c} are fractional orders, and the nonlinear relation is given by

$$x(l) = f(u(l)) = \gamma_1 f_1(u(l)) + \gamma_2 f_2(u(l)) + \cdots + \gamma_m f_m(u(l)). \tag{2}$$

In this paper, the commensurate orders are the multiple of the same value α, such as $\alpha_j = j\alpha$. $e(l)$ can be obtained

$$\begin{aligned} e(l) &= [1 - C(z)]e(l) + v(l) \\ &= [c_1, c_2, \cdots, c_{n_c}] \begin{bmatrix} -\triangle^{\alpha} e(l-1) \\ -\triangle^{\alpha} e(l-2) \\ \vdots \\ -\triangle^{\alpha} e(l-n_c) \end{bmatrix} + v(l). \end{aligned} \tag{3}$$

According Equations (1)–(3), we get

$$\begin{aligned} y(l) &= [1 - A(z)]y(l) + B(z)f(u(l)) + e(l) \\ &= -\sum_{i=1}^{n_a} a_i \triangle^{\alpha} y(l-i) + \sum_{i=1}^{n_b} b_i \sum_{j=1}^{m} \gamma_j \triangle^{\alpha} f_j(u(l-i)) - \sum_{i=1}^{n_c} c_i \triangle^{\alpha} e(l-i) + v(l). \end{aligned} \tag{4}$$

Define the parameter vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ of the linear part and $\boldsymbol{\gamma}$ of the nonlinear part as

$$\boldsymbol{a} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n_a} \end{bmatrix} \in \mathbb{R}^{n_a}, \quad \boldsymbol{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n_b} \end{bmatrix} \in \mathbb{R}^{n_b}, \quad \boldsymbol{c} := \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{n_c} \end{bmatrix} \in \mathbb{R}^{n_c}, \quad \boldsymbol{\gamma} := \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_m \end{bmatrix} \in \mathbb{R}^{m}.$$

Define the information vectors/matrix:

$$\begin{aligned} \boldsymbol{\varphi}_a(l) &:= [-\triangle^{\alpha} y(l-1), -\triangle^{\alpha} y(l-2), \cdots, -\triangle^{\alpha} y(l-n_a)]^{\mathrm{T}} \in \mathbb{R}^{n_a}, \\ \boldsymbol{\varphi}_c(l) &:= [-\triangle^{\alpha} e(l-1), -\triangle^{\alpha} e(l-2), \cdots, -\triangle^{\alpha} e(l-n_c)]^{\mathrm{T}} \in \mathbb{R}^{n_c}, \\ \boldsymbol{F}(l) &:= \begin{bmatrix} \triangle^{\alpha} f_1(u(l-1)) & \triangle^{\alpha} f_1(u(l-2)) & \cdots & \triangle^{\alpha} f_1(u(l-n_b)) \\ \triangle^{\alpha} f_2(u(l-1)) & \triangle^{\alpha} f_2(u(l-2)) & \cdots & \triangle^{\alpha} f_m(u(l-n_b))] \\ \vdots & \vdots & & \vdots \\ \triangle^{\alpha} f_m(u(l-1)) & \triangle^{\alpha} f_m(u(l-2)) & \cdots & \triangle^{\alpha} f_m(u(l-n_b)) \end{bmatrix} \in \mathbb{R}^{m \times n_b}. \end{aligned}$$

Then Equation (4) can be rewritten as

$$\begin{aligned} y(l) &= -\sum_{i=1}^{n_a} a_i \triangle^{\alpha} y(l-i) + \sum_{i=1}^{n_b} b_i \sum_{j=1}^{m} \gamma_j \triangle^{\alpha} f_j(u(l-i)) - \sum_{i=1}^{n_c} c_i \triangle^{\alpha} e(l-i) + v(l) \\ &= \boldsymbol{\varphi}_a^{\mathrm{T}}(l)\boldsymbol{a} + \boldsymbol{\gamma}^{\mathrm{T}} \boldsymbol{F}(l)\boldsymbol{b} + \boldsymbol{\varphi}_c^{\mathrm{T}}(l)\boldsymbol{c} + v(l). \end{aligned} \tag{5}$$

Equation (5) is the identification model of the fractional-order block-oriented nonlinear system in (1). The objective of this paper is to propose the iterative identification algorithms to estimate the unknown parameters of the fractional-order block-oriented nonlinear systems by using the decomposition technique.

3. The two-stage over-parameterization gradient-based iterative algorithm

The section aims to apply the gradient search to derive an iterative estimation algorithm for the system. Let $\boldsymbol{F}_i(l) \in \mathbb{R}^{1 \times n_b}$ be the ith row of the information matrix $\boldsymbol{F}(l)$. Redefine the parameter vectors

$$\begin{aligned} \boldsymbol{\theta} &:= [\gamma_1 \boldsymbol{b}^{\mathrm{T}}, \gamma_2 \boldsymbol{b}^{\mathrm{T}}, \cdots, \gamma_m \boldsymbol{b}^{\mathrm{T}}] \in \mathbb{R}^{n_b m}, \\ \boldsymbol{\vartheta} &:= [\boldsymbol{a}, \boldsymbol{c}]^{\mathrm{T}} \in \mathbb{R}^{n_a + n_c}. \end{aligned}$$

The corresponding information vector is defined as

$$\begin{aligned} \boldsymbol{\varphi}_F(l) &:= [\boldsymbol{F}_1(l), \boldsymbol{F}_2(l), \cdots, \boldsymbol{F}_m(l)] \in \mathbb{R}^{n_b m}, \\ \boldsymbol{\varphi}(l) &:= [\boldsymbol{\varphi}_a^{\mathrm{T}}(l), \boldsymbol{\varphi}_c^{\mathrm{T}}(l)]^{\mathrm{T}} \in \mathbb{R}^{n_a + n_c}. \end{aligned}$$

The (5) can be written as

$$
\begin{aligned}
y(l) &= \boldsymbol{\varphi}_a^{\mathrm{T}}(l)\boldsymbol{a} + \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{F}(l)\boldsymbol{b} + \boldsymbol{\varphi}_c^{\mathrm{T}}(l)\boldsymbol{c} + v(l) \\
&= \boldsymbol{\varphi}_a^{\mathrm{T}}(l)\boldsymbol{a} + \boldsymbol{\varphi}_c^{\mathrm{T}}(l)\boldsymbol{c} + [\gamma_1, \gamma_2, \cdots, \gamma_m] \begin{bmatrix} \boldsymbol{F}_1(l)\boldsymbol{b} \\ \boldsymbol{F}_2(l)\boldsymbol{b} \\ \vdots \\ \boldsymbol{F}_m(l)\boldsymbol{b} \end{bmatrix} \\
&= \boldsymbol{\varphi}_a^{\mathrm{T}}(l)\boldsymbol{a} + \boldsymbol{\varphi}_c^{\mathrm{T}}(l)\boldsymbol{c} + [\boldsymbol{F}_1(l), \boldsymbol{F}_2(l), \cdots, \boldsymbol{F}_m(l)] \begin{bmatrix} \gamma_1\boldsymbol{b} \\ \gamma_2\boldsymbol{b} \\ \vdots \\ \gamma_m\boldsymbol{b} \end{bmatrix} + v(l) \\
&= \boldsymbol{\varphi}^{\mathrm{T}}(l)\boldsymbol{\vartheta} + \boldsymbol{\varphi}_F^{\mathrm{T}}(l)\boldsymbol{\theta} + v(l). \qquad (6)
\end{aligned}
$$

Introduce two transition variables:

$$
\begin{aligned}
y_1(l) &:= y(l) - \boldsymbol{\varphi}_F^{\mathrm{T}}(l)\boldsymbol{\theta}, \\
y_2(l) &:= y(l) - \boldsymbol{\varphi}^{\mathrm{T}}(l)\boldsymbol{\vartheta}.
\end{aligned}
$$

Thus, Equation (6) can be decomposed into two submodels:

$$
y_1(l) = \boldsymbol{\varphi}^{\mathrm{T}}(l)\boldsymbol{\vartheta} + v(l), \qquad (7)
$$

$$
y_2(l) = \boldsymbol{\varphi}_F^{\mathrm{T}}(l)\boldsymbol{\theta} + v(l). \qquad (8)
$$

Consider the input-output data $\{u(l), y(l) : l = 1, 2, \cdots, L\}$ and define the stacked output vectors and the stacked information matrices:

$$
\begin{aligned}
\boldsymbol{Y}(L) &:= [y(l), y(l+1), \cdots, y(l+L-1)]^{\mathrm{T}} \in \mathbb{R}^L, \\
\mathbf{Y}_1(L) &:= [y_1(l), y_1(l+1), \cdots, y_1(l+L-1)]^{\mathrm{T}} = \boldsymbol{Y}(L) - \boldsymbol{\Phi}_F(l)\boldsymbol{\theta} \in \mathbb{R}^L, \\
\mathbf{Y}_2(L) &:= [y_2(l), y_2(l+1), \cdots, y_2(l+L-1)]^{\mathrm{T}} = \boldsymbol{Y}(L) - \boldsymbol{\Phi}(l)\boldsymbol{\vartheta} \in \mathbb{R}^L, \\
\boldsymbol{\Phi}(L) &:= [\boldsymbol{\varphi}(l), \boldsymbol{\varphi}(l+1), \cdots, \boldsymbol{\varphi}(l+L-1)]^{\mathrm{T}} \in \mathbb{R}^{L\times(n_a+n_c)}, \\
\boldsymbol{\Phi}_F(L) &:= [\boldsymbol{\varphi}_F(l), \boldsymbol{\varphi}_F(l+1), \cdots, \boldsymbol{\varphi}_F(l+L-1)]^{\mathrm{T}} \in \mathbb{R}^{L\times(n_b m)}.
\end{aligned}
$$

For the subsystems in (7) and (8), we define two criterion functions:

$$
\begin{aligned}
J_1(\boldsymbol{\vartheta}) &:= \frac{1}{2}\|\mathbf{Y}_1(l) - \boldsymbol{\Phi}(l)\boldsymbol{\vartheta}\|^2, \\
J_2(\boldsymbol{\theta}) &:= \frac{1}{2}\|\mathbf{Y}_2(l) - \boldsymbol{\Phi}_F(l)\boldsymbol{\theta}\|^2.
\end{aligned}
$$

Minimizing $J_1(\boldsymbol{\vartheta})$ and $J_2(\boldsymbol{\theta})$, we can obtain the estimates of the parameter vectors $\boldsymbol{\vartheta}$, $\boldsymbol{\theta}$ and the order α, respectively. Let $\hat{\boldsymbol{\vartheta}}_k := \begin{bmatrix} \hat{\boldsymbol{a}}_k \\ \hat{\boldsymbol{c}}_k \end{bmatrix}$ be the kth iterative estimate of the parameter vector $\boldsymbol{\vartheta} = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{c} \end{bmatrix}$, and $\hat{\boldsymbol{\theta}}_k$ be the kth iterative estimate of the parameter vector $\boldsymbol{\theta}$, $\hat{\alpha}_k$ be the kth iterative estimate of the parameter α, and $\mu_{1,k}$, $\mu_{2,k}$ and $\mu_{3,k}$ be the kth iterative step sizes. Furthermore, we can obtain the gradient-based iterative relations:

$$
\hat{\boldsymbol{\vartheta}}_k = \hat{\boldsymbol{\vartheta}}_{k-1} + \mu_{1,k}\boldsymbol{\Phi}^{\mathrm{T}}(L)[\mathbf{Y}_1(L) - \boldsymbol{\Phi}(L)\hat{\boldsymbol{\vartheta}}_{k-1}], \qquad (9)
$$

$$
\mu_{1,k} = \underset{\mu_{1,k}\geq 0}{\operatorname{argmin}}\, J_1\big(\hat{\boldsymbol{\vartheta}}_{k-1} - \mu_{1,k}\frac{\partial J_1(\hat{\boldsymbol{\vartheta}}_{k-1}, \hat{\alpha}_{k-1})}{\partial \boldsymbol{\vartheta}}\big), \qquad (10)
$$

$$
\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_{k-1} + \mu_{2,k}\boldsymbol{\Phi}_F^{\mathrm{T}}(L)[\mathbf{Y}_2(L) - \boldsymbol{\Phi}_F(L)\hat{\boldsymbol{\theta}}_{k-1}], \qquad (11)
$$

$$
\mu_{2,k} = \underset{\mu_{2,k}\geq 0}{\operatorname{argmin}}\, J_2\big(\hat{\boldsymbol{\theta}}_{k-1} - \mu_{2,k}\frac{\partial J_1(\hat{\boldsymbol{\theta}}_{k-1}, \hat{\alpha}_{k-1})}{\partial \boldsymbol{\theta}}\big), \qquad (12)
$$

$$
\hat{\alpha}_k = \hat{\alpha}_{k-1} + \mu_{3,k}\left[\frac{\partial(\boldsymbol{\Phi}(L)\hat{\boldsymbol{\vartheta}}_k + \boldsymbol{\Phi}_F(L)\hat{\boldsymbol{\theta}}_{k-1})}{\partial \alpha}\right]^{\mathrm{T}}[\boldsymbol{Y}(L) - \boldsymbol{\Phi}(L)\hat{\boldsymbol{\vartheta}}_k - \boldsymbol{\Phi}_F(L)\hat{\boldsymbol{\theta}}_{k-1}], \qquad (13)
$$

$$\mu_{3,k} = \underset{\mu_{3,k}\geq 0}{\text{argmin}}\, J_2\big(\hat{\alpha}_{k-1} - \mu_{3,k}\frac{\partial J_2(\hat{\boldsymbol{\theta}}_{k-1}, \hat{\alpha}_{k-1})}{\partial \alpha}\big). \tag{14}$$

However, it is not difficult to find that the above iterative relations is incapable of calculating the parameter estimates. First, the information vector $\boldsymbol{\Phi}(l)$ contains the unknown intermediate variable $\triangle^{\alpha}e(l)$. Second, the order of the information vectors $\boldsymbol{\Phi}(l)$ and $\boldsymbol{\Phi}_F(l)$ are non-integer, and the order has to be identified. Replacing the unknown terms $\triangle^{\alpha}e(l)$ and α with their iterative estimates $\triangle^{\hat{\alpha}_{k-1}}\hat{e}_{k-1}(l)$ and $\hat{\alpha}_{k-1}$, combining (9)–(14), we can obtain the two-stage over-parameterized gradient-based iterative (2S-OP-GI) algorithm for estimating the parameter vectors $\boldsymbol{\vartheta}$, $\boldsymbol{\theta}$ and the order α:

$$\hat{\boldsymbol{\vartheta}}_k = \hat{\boldsymbol{\vartheta}}_{k-1} + \mu_{1,k}\hat{\boldsymbol{\Phi}}_k^{\mathrm{T}}(L)[\boldsymbol{Y}(L) - \hat{\boldsymbol{\Phi}}_k(l)\hat{\boldsymbol{\vartheta}}_{k-1} - \hat{\boldsymbol{\Phi}}_{F,k}(l)\hat{\boldsymbol{\theta}}_{k-1}], \tag{15}$$

$$\mu_{1,k} = \underset{\mu_{1,k}\geq 0}{\text{argmin}}\, J_1\big(\hat{\boldsymbol{\vartheta}}_{k-1} - \mu_{1,k}\frac{\partial J_1(\hat{\boldsymbol{\vartheta}}_{k-1}, \hat{\alpha}_{k-1})}{\partial \boldsymbol{\vartheta}}\big), \tag{16}$$

$$\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_{k-1} + \mu_{2,k}\hat{\boldsymbol{\Phi}}_{F,k}^{\mathrm{T}}(L)[\boldsymbol{Y}(L) - \hat{\boldsymbol{\Phi}}_k(l)\hat{\boldsymbol{\vartheta}}_{k-1} - \hat{\boldsymbol{\Phi}}_{F,k}(l)\hat{\boldsymbol{\theta}}_{k-1}], \tag{17}$$

$$\mu_{2,k} = \underset{\mu_{2,k}\geq 0}{\text{argmin}}\, J_2\big(\hat{\boldsymbol{\theta}}_{k-1} - \mu_{2,k}\frac{\partial J_1(\hat{\boldsymbol{\theta}}_{k-1}, \hat{\alpha}_{k-1})}{\partial \boldsymbol{\theta}}\big), \tag{18}$$

$$\hat{\alpha}_k = \hat{\alpha}_{k-1} + \mu_{3,k}\left[\frac{\partial(\hat{\boldsymbol{\Phi}}_k(L)\hat{\boldsymbol{\vartheta}}_{k-1} + \hat{\boldsymbol{\Phi}}_{F,k}(L)\hat{\boldsymbol{\theta}}_{k-1})}{\partial \alpha}\right]^{\mathrm{T}}[\boldsymbol{Y}(L) - \hat{\boldsymbol{\Phi}}_k(l)\hat{\boldsymbol{\vartheta}}_{k-1} - \hat{\boldsymbol{\Phi}}_{F,k}(l)\hat{\boldsymbol{\theta}}_{k-1}], \tag{19}$$

$$\mu_{3,k} = \underset{\mu_{3,k}\geq 0}{\text{argmin}}\, J_2\big(\hat{\alpha}_{k-1} - \mu_{3,k}\frac{\partial J_2(\hat{\boldsymbol{\theta}}_{k-1}, \hat{\alpha}_{k-1})}{\partial \alpha}\big), \tag{20}$$

$$\hat{\boldsymbol{\Phi}}_k(l) = [\hat{\boldsymbol{\varphi}}_k(l), \hat{\boldsymbol{\varphi}}_k(l+1), \cdots, \hat{\boldsymbol{\varphi}}_k(l+L-1)]^{\mathrm{T}}, \tag{21}$$

$$\hat{\boldsymbol{\Phi}}_{F,k}(l) = [\hat{\boldsymbol{\varphi}}_{F,k}(l), \hat{\boldsymbol{\varphi}}_{F,k}(l+1), \cdots, \hat{\boldsymbol{\varphi}}_{F,k}(l+L-1)]^{\mathrm{T}}, \tag{22}$$

$$\hat{\boldsymbol{\varphi}}_k(l) = [\hat{\boldsymbol{\varphi}}_{a,k}^{\mathrm{T}}(l), \hat{\boldsymbol{\varphi}}_{c,k}^{\mathrm{T}}(l)]^{\mathrm{T}}, \tag{23}$$

$$\hat{\boldsymbol{\varphi}}_{a,k}(l) = [-\triangle^{\hat{\alpha}_{k-1}}y(l-1), -\triangle^{\hat{\alpha}_{k-1}}y(l-2), \cdots, -\triangle^{\hat{\alpha}_{k-1}}y(l-n_a)]^{\mathrm{T}}, \tag{24}$$

$$\hat{\boldsymbol{\varphi}}_{c,k}(l) = [-\triangle^{\hat{\alpha}_{k-1}}e(l-1), -\triangle^{\hat{\alpha}_{k-1}}e(l-2), \cdots, -\triangle^{\hat{\alpha}_{k-1}}e(l-n_a)]^{\mathrm{T}}, \tag{25}$$

$$\hat{\boldsymbol{\varphi}}_{F,k}(l) = [\hat{\boldsymbol{F}}_{1,k-1}(l), \hat{\boldsymbol{F}}_{2,k-1}(l), \cdots, \hat{\boldsymbol{F}}_{m,k-1}(l)]^{\mathrm{T}}, \tag{26}$$

$$\hat{\boldsymbol{F}}_{i,k}(l) = [\triangle^{\hat{\alpha}_{k-1}}f_i(u(l-1)), \triangle^{\hat{\alpha}_{k-1}}f_i(u(l-2)), \cdots, \triangle^{\hat{\alpha}_{k-1}}f_i(u(l-n_b))], \tag{27}$$

$$\hat{e}_k(l) = \hat{\boldsymbol{\varphi}}_{c,k}^{\mathrm{T}}(l)\hat{\boldsymbol{c}}_k + v(l), \tag{28}$$

$$\hat{\boldsymbol{\vartheta}}_k = [\hat{a}_{1,k}, \hat{a}_{2,k}, \cdots, \hat{a}_{n_a,k}, \hat{c}_{2,k}, \hat{c}_{3,k}, \cdots, \hat{c}_{n,k}]^{\mathrm{T}}, \tag{29}$$

$$\hat{\boldsymbol{\theta}}_{i,k} = [\widehat{b_i\gamma_1}_k^{\mathrm{T}}, \widehat{b_i\gamma_2}_k^{\mathrm{T}}, \cdots, \widehat{b_i\gamma_m}_k^{\mathrm{T}}]^{\mathrm{T}}, \quad i = 1, 2, \cdots, n_b. \tag{30}$$

After the estimate $\hat{\boldsymbol{\theta}}_k$ of $\boldsymbol{\theta}$ is obtained by using the 2S-OP-GI algorithm, the parameter estimation vector $\hat{\boldsymbol{\theta}}_k$ contains the product of $\boldsymbol{\gamma}$ and $\boldsymbol{b}$. The original parameters $b_1, b_2, \cdots, b_{n_b}, \gamma_1, \gamma_2, \cdots, \gamma_m$ need to be extracted from $\boldsymbol{\theta}$. Then, we let $\hat{\boldsymbol{\theta}}_{1,k} = \widehat{b_1\gamma_j}_k, j = 1, 2, \cdots, m$. Let $\widehat{b_1\gamma_j}_k = \hat{b}_{1,k}\hat{\gamma}_{j,k}$ and $\hat{\boldsymbol{b}}_k$ be the estimate of $\boldsymbol{b}$, and we have

$$\begin{aligned}\sum_{j=1}^{m}\boldsymbol{\theta}_{1,k}^2 &= (\widehat{b_1\gamma_1}_k)^2 + (\widehat{b_1\gamma_2}_k)^2 + \cdots + (\widehat{b_1\gamma_m}_k)^2 \\ &= (\hat{b}_{1,k}\hat{\gamma}_{1,k})^2 + (\hat{b}_{1,k}\hat{\gamma}_{2,k})^2 + \cdots + (\hat{b}_{1,k}\hat{\gamma}_{m,k})^2 \\ &= \hat{b}_{1,k}^2\|\hat{\gamma}_k^2\|.\end{aligned}$$

The iterative estimate $\hat{b}_{1,k}$ can be computed by

$$\hat{b}_{1,k} = \text{sgn}[\hat{\boldsymbol{\theta}}_{1,k}]\text{sqrt}\big(\sum_{\mathrm{i}=1}^{\mathrm{n}}\hat{\boldsymbol{\theta}}_{\mathrm{i,k}}^2\big). \tag{31}$$

Then, we can compute the estimate $\hat{\gamma}_{j,k}$ of γ_j, and the estimate $\hat{b}_{i,k}$ of b_i by

$$\hat{\gamma}_{j,k} = \hat{\boldsymbol{\theta}}_{j,k}/\hat{b}_{1,k}, \quad j = 1, 2, \cdots, m. \tag{32}$$

$$\hat{b}_{i,k} = \frac{1}{n}\sum_{j=1}^{m}\frac{\widehat{b_i\gamma_j}_k}{\hat{\gamma}_{j,k}} = \frac{1}{n}\sum_{j=1}^{m}\frac{\hat{\boldsymbol{\theta}}_k(im-m+j)}{\hat{\gamma}_{j,k}}, \quad i=2,3,\cdots,n_b. \tag{33}$$

Equations (15)–(33) make up the 2S-OP-GI algorithm. The steps of computing the parameter estimates $\hat{\boldsymbol{\theta}}_k$, $\hat{\boldsymbol{\vartheta}}_k$ nad $\hat{\alpha}_k$ are listed as follows.

1. For $k \leqslant 0$, all the variables are set to zero. Let $k=1$, and set the initial values: $\hat{\boldsymbol{\theta}}_0 = \mathbf{1}_{n_b m}/p_0$, $\hat{\boldsymbol{\vartheta}}_0 = \mathbf{1}_{n_a+n_c}/p_0$, $p_0 = 10^6$ and $\hat{\alpha}_0$ to be a random number, the parameter estimation accuracy ϵ.
2. Collect the input and output data $u(l)$ and $y(l)$ and set the data length L.
3. Calculate $\hat{e}_k(l)$ and $\hat{\boldsymbol{F}}_{i,k}(l)$ according to (28) and (27), form the information vectors $\hat{\boldsymbol{\varphi}}_{a,k}(l)$, $\hat{\boldsymbol{\varphi}}_{c,k}(l)$, $\hat{\boldsymbol{\varphi}}_{F,k}(l)$ and $\hat{\boldsymbol{\varphi}}_k(l)$ using (24), (25), (26) and (23).
4. Construct the stacked information matrices $\hat{\boldsymbol{\Phi}}_k(l)$ and $\hat{\boldsymbol{\Phi}}_{F,k}(l)$ by (21) and (22), select a large $\mu_{1,k}(l)$, $\mu_{2,k}(l)$ and $\mu_{3,k}(l)$ according to (16), (18) and (20).
5. Update the parameter estimation vectors $\hat{\boldsymbol{\vartheta}}_k$ and $\hat{\boldsymbol{\theta}}_k$ by using (15) and (17), and update the order estimate $\hat{\alpha}_k$ using (19).
6. If $\|\hat{\boldsymbol{\vartheta}}_k - \hat{\boldsymbol{\vartheta}}_{k-1}\| + \|\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_{k-1}\| + \|\hat{\alpha}_k - \hat{\alpha}_{k-1}\| > \varepsilon$, increase k by 1 and go to Step 3; otherwise, obtain the parameter estimation vectors $\hat{\boldsymbol{\vartheta}}_k$ and $\hat{\boldsymbol{\theta}}_k$, and the order $\hat{\alpha}_k$ and terminate the process.

4. The key term separation two-stage gradient-based iterative algorithm

In this section, a key term separation identification model is established based on the key term separation technique. The goal is to derive a key term separation two-stage gradient-based iterative algorithm for estimating the parameter vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{\gamma}$ and α from available observation data.

4.1. The key term separation identification model

To obtain the unique parameter estimates, one has to let $b_1 = 1$. Choose $x(k)$ as the key term to parameterize the fractional-order block-oriented nonlinear system. Then the Equation (5) output can be expressed as

$$\begin{aligned} y(l) &= -\sum_{i=1}^{n_a} a_i \triangle^{\alpha} y(l-i) + \sum_{j=1}^{m}\gamma_j \triangle^{\alpha} f_j(u(l-1)) + \sum_{i=2}^{n_b} b_i \sum_{j=1}^{m}\gamma_j \triangle^{\alpha} f_j(u(l-i)) - \sum_{i=1}^{n_c} c_i \triangle^{\alpha} e(l-i) + v(l) \\ &= \boldsymbol{\phi}_a^{\mathrm{T}}(l)\boldsymbol{a} + \boldsymbol{\phi}_f^{\mathrm{T}}(l)\boldsymbol{\gamma} + \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{F}(l)\boldsymbol{b}' + \boldsymbol{\phi}_c^{\mathrm{T}}(l)\boldsymbol{c} + v(l) \\ &= \boldsymbol{\phi}^{\mathrm{T}}(l)\boldsymbol{\vartheta} + \boldsymbol{\phi}_F(l)\boldsymbol{b}' + v(l), \end{aligned} \tag{34}$$

where the corresponding information and parameter vectors are defined as

$$\begin{aligned} \boldsymbol{\phi}_a(l) &:= [-\triangle^{\alpha} y(l-1), -\triangle^{\alpha} y(l-2), \cdots, -\triangle^{\alpha} y(l-n_a)]^{\mathrm{T}} \in \mathbb{R}^{n_a}, \\ \boldsymbol{\phi}_f(l) &:= [\triangle^{\alpha} f_1(u(l-1)), \triangle^{\alpha} f_2(u(l-1)), \cdots, \triangle^{\alpha} f_m(u(l-1))]^{\mathrm{T}} \in \mathbb{R}^{m}, \\ \boldsymbol{\phi}_c(l) &:= [-\triangle^{\alpha} e(l-1), -\triangle^{\alpha} e(l-2), \cdots, -\triangle^{\alpha} e(l-n_c)]^{\mathrm{T}} \in \mathbb{R}^{n_c}, \\ \boldsymbol{\phi}(l) &:= [\boldsymbol{\phi}_a^{\mathrm{T}}(l), \boldsymbol{\phi}_f^{\mathrm{T}}(l), \boldsymbol{\phi}_c^{\mathrm{T}}(l)]^{\mathrm{T}} \in \mathbb{R}^{n_a+n_c+m}, \\ \boldsymbol{\phi}_F(l) &:= \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{F}(l) \in \mathbb{R}^{n_b-1}, \\ \boldsymbol{F}(l) &:= \begin{bmatrix} \triangle^{\alpha} f_1(u(l-2)) & \triangle^{\alpha} f_1(u(l-3)) & \cdots & \triangle^{\alpha} f_1(u(l-n_b)) \\ \triangle^{\alpha} f_2(u(l-2)) & \triangle^{\alpha} f_2(u(l-3)) & \cdots & \triangle^{\alpha} f_m(u(l-n_b))] \\ \vdots & \vdots & & \vdots \\ \triangle^{\alpha} f_m(u(l-2)) & \triangle^{\alpha} f_m(u(l-3)) & \cdots & \triangle^{\alpha} f_m(u(l-n_b)) \end{bmatrix} \in \mathbb{R}^{m\times(n_b-1)}, \\ \boldsymbol{a} &:= [a_1, a_2, \cdots, a_{n_a}]^{\mathrm{T}} \in \mathbb{R}^{n_a}, \\ \boldsymbol{b}' &:= [b_2, b_3, \cdots, b_{n_b}]^{\mathrm{T}} \in \mathbb{R}^{n_b-1}, \\ \boldsymbol{c} &:= [c_1, c_2, \cdots, c_{n_c}]^{\mathrm{T}} \in \mathbb{R}^{n_c}, \end{aligned}$$

$$\boldsymbol{\gamma} := [\gamma_1, \gamma_2, \cdots, \gamma_m]^{\mathrm{T}} \in \mathbb{R}^m,$$
$$\boldsymbol{\vartheta} := [\boldsymbol{a}^{\mathrm{T}}, \boldsymbol{c}^{\mathrm{T}}, \boldsymbol{\gamma}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{n_a+n_c+m}.$$

Based on the model in (34), the decomposed two submodels are given by

$$y_3(l) = \boldsymbol{\phi}^{\mathrm{T}}(l)\boldsymbol{\vartheta} + v(l), \tag{35}$$
$$y_4(l) = \boldsymbol{\phi}_F(l)\boldsymbol{b}' + v(l). \tag{36}$$

Therefore, we can get the identification model and derive the key term separation two-stage gradient-based iterative algorithm.

4.2. The key term separation two-stage gradient-based iterative algorithm

According to the key term separation identification model in (35) and (36), construct the information matrices $\boldsymbol{\Phi}(L)$ and $\boldsymbol{\Phi}_F(L)$ and the system outputs $\boldsymbol{Y}(L)$, $\boldsymbol{Y}_1(L)$ and $\boldsymbol{Y}_2(L)$ as

$$\boldsymbol{Y}(L) := [y(l), y(l+1), \cdots, y(l+L-1)]^{\mathrm{T}} \in \mathbb{R}^L,$$
$$\boldsymbol{Y}_1(L) := [y_1(l), y_1(l+1), \cdots, y_1(l+L-1)]^{\mathrm{T}} = \boldsymbol{Y}(L) - \boldsymbol{\Phi}_F(l)\boldsymbol{b}' \in \mathbb{R}^L,$$
$$\boldsymbol{Y}_2(L) := [y_2(l), y_2(l+1), \cdots, y_2(l+L-1)]^{\mathrm{T}} = \boldsymbol{Y}(L) - \boldsymbol{\Phi}(l)\boldsymbol{\vartheta} \in \mathbb{R}^L,$$
$$\boldsymbol{\Phi}(L) := [\boldsymbol{\phi}(l), \boldsymbol{\phi}(l+1), \cdots, \boldsymbol{\phi}(l+L-1)]^{\mathrm{T}} \in \mathbb{R}^{L\times(n_a+n_c+m)},$$
$$\boldsymbol{\Phi}_F(L) := [\boldsymbol{\phi}_F(l), \boldsymbol{\phi}_F(l+1), \cdots, \boldsymbol{\phi}_F(l+L-1)]^{\mathrm{T}} \in \mathbb{R}^{L\times(n_b-1)}.$$

Consider the input-output data $\{u(l), y(l) : l = 1, 2, \cdots, L\}$, define two criterion functions:

$$J_3(\boldsymbol{\vartheta}, \alpha) := \frac{1}{2}\sum_{j=1}^{L}[y_3(l) - \boldsymbol{\phi}^{\mathrm{T}}(l)\boldsymbol{\vartheta}]^2,$$
$$J_4(\boldsymbol{\gamma}, \alpha) := \frac{1}{2}\sum_{j=1}^{L}[y_4(l) - \boldsymbol{\phi}_F(l)\boldsymbol{b}']^2.$$

Let $k = 1, 2, 3, \cdots$ be an iterative variable, $\hat{\boldsymbol{\vartheta}}_k \in \mathbb{R}^{n_a+n_c+m}$, $\hat{\boldsymbol{b}}'_k \in \mathbb{R}^{n_b-1}$ and $\hat{\alpha}_k$ be the estimates of the parameter vectors $\boldsymbol{\vartheta}$ and $\boldsymbol{b}'$ and order α at iteration k, and $\rho_{1,k}$, $\rho_{2,k}$ and $\rho_{3,k}$ be the iterative step sizes. Using the negative gradient search and minimizing $J_3(\boldsymbol{\vartheta}, \alpha)$ and $J_4(\boldsymbol{\gamma}, \alpha)$ lead to the following gradient-based iterative relations for computing $\hat{\boldsymbol{\vartheta}}_k$, $\hat{\boldsymbol{b}}'_k$ and $\hat{\alpha}_k$:

$$\hat{\boldsymbol{\vartheta}}_k = \hat{\boldsymbol{\vartheta}}_{k-1} + \rho_{1,k}\boldsymbol{\Phi}^{\mathrm{T}}(L)[\boldsymbol{Y}_1(L) - \boldsymbol{\Phi}(L)\hat{\boldsymbol{\vartheta}}_{k-1}], \tag{37}$$
$$\hat{\boldsymbol{b}}'_k = \hat{\boldsymbol{b}}'_{k-1} + \rho_{2,k}\boldsymbol{\Phi}_F^{\mathrm{T}}(L)[\boldsymbol{Y}_2(L) - \boldsymbol{\Phi}_F(L)\hat{\boldsymbol{b}}'_{k-1}], \tag{38}$$
$$\hat{\alpha}_k = \hat{\alpha}_{k-1} + \rho_{3,k}\left[\frac{\partial(\boldsymbol{\Phi}(L)\hat{\boldsymbol{\vartheta}}_{k-1} + \boldsymbol{\Phi}_F(L)\hat{\boldsymbol{b}}'_{k-1})}{\partial\alpha}\right]^{\mathrm{T}}[\boldsymbol{Y}(L) - \boldsymbol{\Phi}(L)\hat{\boldsymbol{\vartheta}}_{k-1} - \boldsymbol{\Phi}_F(L)\hat{\boldsymbol{b}}'_{k-1}]. \tag{39}$$

Replacing $\boldsymbol{\Phi}(L)$ and $\boldsymbol{\Phi}_F(L)$ in (37) – (39) with their estimates $\hat{\boldsymbol{\Phi}}_k(L)$ and $\hat{\boldsymbol{\Phi}}_F(L)$ yields the following gradient-based iterative algorithm for estimating $\boldsymbol{\vartheta}$, $\boldsymbol{b}'$ and α:

$$\hat{\boldsymbol{\vartheta}}_k = \hat{\boldsymbol{\vartheta}}_{k-1} + \rho_{1,k}\hat{\boldsymbol{\Phi}}_k^{\mathrm{T}}(L)[\boldsymbol{Y}(L) - \hat{\boldsymbol{\Phi}}_k(l)\hat{\boldsymbol{\vartheta}}_{k-1} - \hat{\boldsymbol{\Phi}}_{F,k}(l)\hat{\boldsymbol{\theta}}_{k-1}], \tag{40}$$
$$\rho_{1,k} = \underset{\rho_{1,k}\geq 0}{\operatorname{argmin}}\, J_1\Big(\hat{\boldsymbol{\vartheta}}_{k-1} - \rho_{1,k}\frac{\partial J_1(\hat{\boldsymbol{\vartheta}}_{k-1}, \hat{\alpha}_{k-1})}{\partial\boldsymbol{\vartheta}}\Big), \tag{41}$$
$$\hat{\boldsymbol{b}}'_k = \hat{\boldsymbol{b}}'_{k-1} + \rho_{2,k}\hat{\boldsymbol{\Phi}}_{F,k}^{\mathrm{T}}(L)[\boldsymbol{Y}(L) - \hat{\boldsymbol{\Phi}}_k(l)\hat{\boldsymbol{\vartheta}}_{k-1} - \hat{\boldsymbol{\Phi}}_{F,k}(l)\hat{\boldsymbol{b}}'_{k-1}], \tag{42}$$
$$\rho_{2,k} = \underset{\rho_{2,k}\geq 0}{\operatorname{argmin}}\, J_2\Big(\hat{\boldsymbol{b}}'_{k-1} - \rho_{2,k}\frac{\partial J_1(\hat{\boldsymbol{b}}'_{k-1}, \hat{\alpha}_{k-1})}{\partial\boldsymbol{b}'}\Big), \tag{43}$$
$$\hat{\alpha}_k = \hat{\alpha}_{k-1} + \rho_{3,k}\left[\frac{\partial(\hat{\boldsymbol{\Phi}}_k(L)\hat{\boldsymbol{\vartheta}}_{k-1} + \hat{\boldsymbol{\Phi}}_{F,k}(L)\hat{\boldsymbol{b}}'_{k-1})}{\partial\alpha}\right]^{\mathrm{T}}[\boldsymbol{Y}(L) - \hat{\boldsymbol{\Phi}}_k(l)\hat{\boldsymbol{\vartheta}}_{k-1} - \hat{\boldsymbol{\Phi}}_{F,k}(l)\hat{\boldsymbol{b}}'_{k-1}], \tag{44}$$

$$\rho_{3,k} = \mathop{\mathrm{argmin}}_{\rho_{3,k}\geq 0} J_2\big(\hat{\alpha}_{k-1} - \rho_{3,k}\frac{\partial J_2(\hat{\boldsymbol{b}'}_{k-1}, \hat{\alpha}_{k-1})}{\partial \alpha}\big), \tag{45}$$

$$\hat{\boldsymbol{\Phi}}_k(l) = [\hat{\boldsymbol{\phi}}_k(l), \hat{\boldsymbol{\phi}}_k(l+1), \cdots, \hat{\boldsymbol{\phi}}_k(l+L-1)]^{\mathrm{T}}, \tag{46}$$

$$\hat{\boldsymbol{\Phi}}_{F,k}(l) = [\hat{\boldsymbol{\phi}}_{F,k}(l), \hat{\boldsymbol{\phi}}_{F,k}(l+1), \cdots, \hat{\boldsymbol{\phi}}_{F,k}(l+L-1)]^{\mathrm{T}}, \tag{47}$$

$$\hat{\boldsymbol{\phi}}_k(l) = [\hat{\boldsymbol{\phi}}^{\mathrm{T}}_{a,k}(l), \hat{\boldsymbol{\phi}}^{\mathrm{T}}_{f,k}(l), \hat{\boldsymbol{\phi}}^{\mathrm{T}}_{c,k}(l)]^{\mathrm{T}}, \tag{48}$$

$$\hat{\boldsymbol{\phi}}_{a,k}(l) = [-\triangle^{\hat{\alpha}_{k-1}} y(l-1), -\triangle^{\hat{\alpha}_{k-1}} y(l-2), \cdots, -\triangle^{\hat{\alpha}_{k-1}} y(l-n_a)]^{\mathrm{T}}, \tag{49}$$

$$\hat{\boldsymbol{\phi}}_{f,k}(l) = [\triangle^{\hat{\alpha}_{k-1}} f_1(u(l-1)), \triangle^{\hat{\alpha}_{k-1}} f_2(u(l-1)), \cdots, \triangle^{\hat{\alpha}_{k-1}} f_m(u(l-1))]^{\mathrm{T}}, \tag{50}$$

$$\hat{\boldsymbol{\phi}}_{c,k}(l) = [-\triangle^{\hat{\alpha}_{k-1}} e(l-1), -\triangle^{\hat{\alpha}_{k-1}} e(l-2), \cdots, -\triangle^{\hat{\alpha}_{k-1}} e(l-n_a)]^{\mathrm{T}}, \tag{51}$$

$$\hat{F}_k(l) := \begin{bmatrix} \triangle^{\hat{\alpha}_{k-1}} f_1(u(l-2)) & \triangle^{\hat{\alpha}_{k-1}} f_1(u(l-3)) & \cdots & \triangle^{\hat{\alpha}_{k-1}} f_1(u(l-n_b)) \\ \triangle^{\hat{\alpha}_{k-1}} f_2(u(l-2)) & \triangle^{\hat{\alpha}_{k-1}} f_2(u(l-3)) & \cdots & \triangle^{\hat{\alpha}_{k-1}} f_m(u(l-n_b))] \\ \vdots & \vdots & & \vdots \\ \triangle^{\hat{\alpha}_{k-1}} f_m(u(l-2)) & \triangle^{\hat{\alpha}_{k-1}} f_m(u(l-3)) & \cdots & \triangle^{\hat{\alpha}_{k-1}} f_m(u(l-n_b)) \end{bmatrix}, \tag{52}$$

$$\hat{e}_k(l) = \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{c,k}(l)\hat{\boldsymbol{c}}_k + v(l), \tag{53}$$

$$\hat{\boldsymbol{\vartheta}}_k = [\hat{a}_{1,k}, \hat{a}_{2,k}, \cdots, \hat{a}_{n_a,k}, \hat{\gamma}_{1,k}, \hat{\gamma}_{2,k}, \cdots, \hat{\gamma}_{m,k}, \hat{c}_{2,k}, \hat{c}_{3,k}, \cdots, \hat{c}_{n_c,k}]^{\mathrm{T}}, \tag{54}$$

$$\hat{\boldsymbol{b}'}_k = [\hat{b}_{1,k}, \hat{b}_{2,k}, \cdots, \hat{b}_{n_b,k}]^{\mathrm{T}}. \tag{55}$$

Equations (40)–(55) make up the KT-2S-GI algorithm. The steps of computing the parameter estimates $\hat{\boldsymbol{\vartheta}}_k$, $\hat{\boldsymbol{b}'}_k$ and $\hat{\alpha}_k$ are listed as follows:

1. For $k \leqslant 0$, all the variables are set to zero. Let $k = 1$, and set the initial values: $\hat{\boldsymbol{\vartheta}}_0 = \mathbf{1}_{n_a+n_c+m}/p_0$, $\hat{\boldsymbol{b}'}_0 = \mathbf{1}_{n_b-1}/p_0$, $p_0 = 10^6$ and $\hat{\alpha}_0$ to be a random number, the parameter estimation accuracy ϵ.
2. Collect the input and output data $u(l)$ and $y(l)$, set the data length L.
3. Calculate $\hat{e}_k(l)$ according to (53), form the information vectors $\hat{\boldsymbol{\phi}}_{a,k}(l)$, $\hat{\boldsymbol{\phi}}_{f,k}(l)$, $\hat{\boldsymbol{\varphi}}_{c,k}(l)$ and $\hat{\boldsymbol{\phi}}_k(l)$ using (49), (50), (51) and (48).
4. Construct the stacked information matrices $\hat{\boldsymbol{\Phi}}_k(l)$, $\hat{\boldsymbol{\Phi}}_{F,k}(l)$ and $\hat{F}_k(l)$ by (46), (47) and (52), select a large $\rho_{1,k}(l)$, $\rho_{2,k}(l)$ and $\rho_{3,k}(l)$ according to (41), (43) and (45).
5. Update the parameter estimation vectors $\hat{\boldsymbol{\vartheta}}_k$ and $\hat{\boldsymbol{b}'}_k$ using (40) and (42), update the order estimate $\hat{\alpha}_k$ using (44).
6. If $\|\hat{\boldsymbol{\vartheta}}_k - \hat{\boldsymbol{\vartheta}}_{k-1}\| + \|\hat{\boldsymbol{b}'}_k - \hat{\boldsymbol{b}'}_{k-1}\| + \|\hat{\alpha}_k - \hat{\alpha}_{k-1}\| > \varepsilon$, increase k by 1 and go to Step 3; otherwise, obtain the parameter estimation vectors $\hat{\boldsymbol{\vartheta}}_k$ and $\hat{\boldsymbol{b}'}_k$, and the order $\hat{\alpha}_k$ and terminate the process.

5. The comparison of the computational efficiency

The following discusses the computational efficiency of the 2S-OP-GI algorithm and the KT-2S-GI algorithm. The flop (floating point operation) counting is a simple approach to the measuring of program efficiency since it ignores subscripting, memory traffic, and the countless other overheads associated with program execution, the flop counting is just a quick accounting method that captures only one of the several dimensions of the efficiency issue although multiplication (division) and addition (subtraction) with different lengths are different. The computational efficiency of the 2S-OP-GI algorithm and the KT-2S-GI algorithm is shown in Tables 1 and 2, where $n_1 := n_a + n_c + n_b m$ and $n_2 := n_a + n_c + m + n_b - 1$. By comparing the calculation amount of the two algorithms, we have

$$\begin{aligned} N_1 - N_2 &= (4k+2)(n_1 - n_2)L + (n_b m - n_b + 1)2L - (n_1 - n_2)k \\ &= (4k+2)(n_b m - m - n_b + 1)L + (n_b m - n_b + 1)2L - (n_b m - m - n_b + 1)k > 0. \end{aligned}$$

Table 1: The computational efficiency of the 2S-OP-GI algorithm

Variables	Expressions	Multiplications	Additions
$\hat{\boldsymbol{\vartheta}}_k$	$\hat{\boldsymbol{\vartheta}}_k = \hat{\boldsymbol{\vartheta}}_{k-1} + \mu_{1,k}\hat{\boldsymbol{\Phi}}_k^{\mathrm{T}}(L)\hat{\boldsymbol{E}}_k$	$(L+1)(n_a+n_c)k$	$(n_a+n_c)Lk$
$\mu_{1,k}$	$\mu_{1,k} = \underset{\mu_{1,k}\geq 0}{\operatorname{argmin}} J_1\big(\hat{\boldsymbol{\vartheta}}_{k-1} - \mu_{1,k}\frac{\partial J_1(\hat{\boldsymbol{\vartheta}}_{k-1},\hat{\alpha}_{k-1})}{\partial \boldsymbol{\vartheta}}\big)$	$(n_a+n_c)L+1$	$(n_a+n_c)L-1$
$\hat{\boldsymbol{\theta}}_k$	$\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_{k-1} + \mu_{2,k}\hat{\boldsymbol{\Phi}}_{F,k}^{\mathrm{T}}(L)\hat{\boldsymbol{E}}_k$	$(L+1)n_bmk$	n_bmLk
$\mu_{2,k}$	$\mu_{2,k} = \underset{\mu_{2,k}\geq 0}{\operatorname{argmin}} J_2\big(\hat{\boldsymbol{\theta}}_{k-1} - \mu_{2,k}\frac{\partial J_1(\hat{\boldsymbol{\theta}}_{k-1},\hat{\alpha}_{k-1})}{\partial \boldsymbol{\theta}}\big)$	n_bmL+1	n_bmL-1
$\hat{\boldsymbol{E}}_k$	$\hat{\boldsymbol{E}}_k = \boldsymbol{Y}(L) - \hat{\boldsymbol{\Phi}}_k(l)\hat{\boldsymbol{\vartheta}}_{k-1} - \hat{\boldsymbol{\Phi}}_{F,k}(l)\hat{\boldsymbol{\theta}}_{k-1}$	n_1Lk	n_1kL
$\hat{\alpha}_k$	$\hat{\alpha}_k = \hat{\alpha}_{k-1} + \mu_{3,k}\left[\frac{\partial(\hat{\boldsymbol{\Phi}}_k(L)\hat{\boldsymbol{\vartheta}}_{k-1}+\hat{\boldsymbol{\Phi}}_{F,k}(L)\hat{\boldsymbol{\theta}}_{k-1})}{\partial \alpha}\right]^{\mathrm{T}}\hat{\boldsymbol{E}}_k$	$(L+1)k$	Lk
$\mu_{3,k}$	$\mu_{3,k} = \underset{\mu_{3,k}\geq 0}{\operatorname{argmin}} J_2\big(\hat{\alpha}_{k-1} - \mu_{3,k}\frac{\partial J_2(\hat{\boldsymbol{\theta}}_{k-1},\hat{\alpha}_{k-1})}{\partial \alpha}\big)$	n_bmL+1	n_bmL-1
Sum		$(2L+1)n_1k+(n_1+n_bm)L$	$(2n_1+1)kL$
		$+(L+1)k+3$	$+(n_1+n_bm)L-3$
Total flops		$N_1 := (4k+2)n_1L+2L(n_bm+k)+n_1k+k$	

Table 2: The computational efficiency of the KT-2S-GI algorithm

Variables	Expressions	Multiplications	Additions
$\hat{\boldsymbol{\vartheta}}_k$	$\hat{\boldsymbol{\vartheta}}_k = \hat{\boldsymbol{\vartheta}}_{k-1} + \rho_{1,k}\hat{\boldsymbol{\Phi}}_k^{\mathrm{T}}(L)\hat{\boldsymbol{E}}_k$	$(L+1)(n_a+n_c+m)k$	$(n_a+n_c+m)Lk$
$\rho_{1,k}$	$\rho_{1,k} = \underset{\rho_{1,k}\geq 0}{\operatorname{argmin}} J_1\big(\hat{\boldsymbol{\vartheta}}_{k-1} - \rho_{1,k}\frac{\partial J_1(\hat{\boldsymbol{\vartheta}}_{k-1},\hat{\alpha}_{k-1})}{\partial \boldsymbol{\vartheta}}\big)$	$(n_a+n_c+m)L+1$	$(n_a+n_c+m)L-1$
$\hat{\boldsymbol{b}'}_k$	$\hat{\boldsymbol{b}'}_k = \hat{\boldsymbol{b}'}_{k-1} + \rho_{2,k}\hat{\boldsymbol{\Phi}}_{F,k}^{\mathrm{T}}(L)\hat{\boldsymbol{E}}_k$	$(L+1)(n_b-1)k$	$(n_b-1)Lk$
$\rho_{2,k}$	$\rho_{2,k} = \underset{\rho_{2,k}\geq 0}{\operatorname{argmin}} J_2\big(\hat{\boldsymbol{b}'}_{k-1} - \rho_{2,k}\frac{\partial J_1(\hat{\boldsymbol{b}'}_{k-1},\hat{\alpha}_{k-1})}{\partial \boldsymbol{b}'}\big)$	$(n_b-1)L+1$	$(n_b-1)L-1$
$\hat{\boldsymbol{E}}_k$	$\hat{\boldsymbol{E}}_k = \boldsymbol{Y}(L) - \hat{\boldsymbol{\Phi}}_k(l)\hat{\boldsymbol{\vartheta}}_{k-1} - \hat{\boldsymbol{\Phi}}_{F,k}(l)\hat{\boldsymbol{\theta}}_{k-1}$	n_2Lk	n_2kL
$\hat{\alpha}_k$	$\hat{\alpha}_k = \hat{\alpha}_{k-1} + \rho_{3,k}\left[\frac{\partial(\hat{\boldsymbol{\Phi}}_k(L)\hat{\boldsymbol{\vartheta}}_{k-1}+\hat{\boldsymbol{\Phi}}_{F,k}(L)\hat{\boldsymbol{b}'}_{k-1})}{\partial \alpha}\right]^{\mathrm{T}}\hat{\boldsymbol{E}}_k$	$(L+1)k$	Lk
$\rho_{3,k}$	$\rho_{3,k} = \underset{\rho_{3,k}\geq 0}{\operatorname{argmin}} J_2\big(\hat{\alpha}_{k-1} - \rho_{3,k}\frac{\partial J_2(\hat{\boldsymbol{b}'}_{k-1},\hat{\alpha}_{k-1})}{\partial \alpha}\big)$	$(n_b-1)L+1$	$(n_b-1)L-1$
Sum		$(2L+1)n_2k+(n_2+n_b-1)L$	$(2n_2+1)kL$
		$+(L+1)k+3$	$+(n_2+n_b-1)L-3$
Total flops		$N_2 := (4k+2)n_2L+2L(n_b-1+k)+n_2k+k$	

We can clearly see that the calculation amount of the KT-2S-GI algorithm is less than that of the 2S-OP-GI algorithm. For example, when $n_a = 10$, $n_b = 10$, $n_c = 10$, $m = 10$, $L = 1000$ and $k = 30$, we have $N_1 - N_2 > 1.0 \times 10^7$.

6. Examples

6.1. A fractional-order block-oriented nonlinear system

Consider the following nonlinear system,

$$y(l) = \frac{B(z)}{A(z)}x(l) + \frac{1}{C(z)}v(l),$$
$$x(l) = \gamma_1 f(u(l)) + \gamma_2 f^2(u(l)),$$

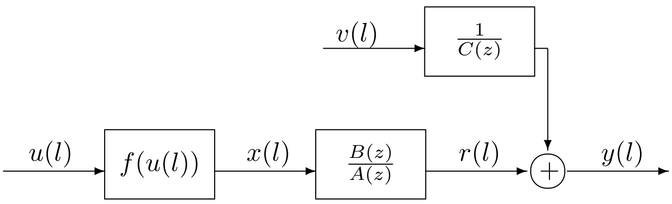

Figure 1: The fractional-order block-oriented nonlinear system

$$
\begin{aligned}
B(z) &= b_1 z^{-\alpha} + b_2 z^{-2\alpha} + b_3 z^{-3\alpha}, \\
A(z) &= a_1 z^{-\alpha} + a_2 z^{-2\alpha}, \\
C(z) &= c_1 z^{-\alpha} + c_2 z^{-2\alpha}.
\end{aligned}
$$

In simulation, the input $\{u(l)\}$ is taken as a persistent excitation signal sequence with zero mean and unit variance, and $\{v(l)\}$ is taken as a white noise sequence with zero mean and variance σ^2. Taking the data length $L = 1000$, we collect the data $u(l)$ and $y(l)$, and use the 2S-OP-GI algorithm to estimate the parameters of the example system, the parameter estimates and their errors with variances $\sigma^2 = 0.50^2$ and $\sigma^2 = 1.00^2$ are shown in Tables 3–4. The relative parameter estimation errors $\delta := \|\hat{\boldsymbol{\vartheta}}_k - \boldsymbol{\vartheta}\|/\|\boldsymbol{\vartheta}\|$ are shown in Figures 2–3 with different variances.

Table 3: The 2S-OP-GI estimates and their errors with $\sigma^2 = 0.50^2$

k	a_1	a_2	c_1	c_2	b_1	b_2	b_3	γ_1	γ_2	α	δ
1	-0.03188	-0.01594	-0.00329	0.01433	0.02488	-0.02612	0.05862	0.00367	0.01513	0.45000	0.73129
10	-0.03692	-0.01478	-0.00652	0.01819	0.03182	-0.08995	-0.06937	0.00853	0.02198	0.48998	0.53138
20	-0.03733	-0.01450	-0.00667	0.01891	0.07537	-0.15564	-0.22203	0.00988	0.02290	0.48565	0.29156
30	-0.03741	-0.01442	-0.00672	0.01928	0.11750	-0.17444	-0.30365	0.01127	0.02319	0.48870	0.15689
40	-0.03745	-0.01438	-0.00675	0.01951	0.14423	-0.17938	-0.34389	0.01236	0.02336	0.49099	0.08679
50	-0.03747	-0.01436	-0.00678	0.01965	0.16042	-0.18077	-0.36530	0.01312	0.02347	0.49283	0.04830
60	-0.03749	-0.01434	-0.00679	0.01975	0.17039	-0.18115	-0.37752	0.01363	0.02353	0.49438	0.02647
70	-0.03750	-0.01433	-0.00680	0.01981	0.17668	-0.18125	-0.38486	0.01397	0.02357	0.49573	0.01485
80	-0.03751	-0.01433	-0.00681	0.01988	0.18419	-0.18130	-0.39317	0.01438	0.02361	0.49848	0.01274
90	-0.03751	-0.01433	-0.00681	0.01988	0.18454	-0.18130	-0.39354	0.01440	0.02361	0.49869	0.01315
100	-0.03751	-0.01433	-0.00681	0.01989	0.18518	-0.18131	-0.39420	0.01444	0.02361	0.49910	0.01397
True values	-0.03800	-0.01400	-0.00800	0.02000	0.18000	-0.17500	-0.39000	0.01650	0.02400	0.50000	

Table 4: The 2S-OP-GI estimates and their errors with $\sigma^2 = 1.00^2$

k	a_1	a_2	c_1	c_2	b_1	b_2	b_3	γ_1	γ_2	α	δ
1	-0.03207	-0.01602	-0.00341	0.01471	0.04442	-0.03795	0.10556	0.00420	0.01600	0.45000	0.78106
10	-0.03709	-0.01445	-0.00649	0.01838	0.07044	-0.13368	-0.08875	0.01114	0.02444	0.48314	0.47314
20	-0.03739	-0.01418	-0.00662	0.01907	0.11046	-0.18239	-0.24993	0.01351	0.02526	0.48231	0.23035
30	-0.03746	-0.01409	-0.00667	0.01942	0.13976	-0.18651	-0.32964	0.01568	0.02554	0.48298	0.11026
40	-0.03749	-0.01405	-0.00670	0.01964	0.15701	-0.18381	-0.36794	0.01729	0.02571	0.48372	0.05398
50	-0.03752	-0.01402	-0.00672	0.01979	0.16714	-0.18109	-0.38812	0.01838	0.02581	0.48447	0.03121
60	-0.03753	-0.01400	-0.00673	0.01988	0.17328	-0.17913	-0.39956	0.01910	0.02587	0.48523	0.02864
70	-0.03754	-0.01399	-0.00674	0.01993	0.17710	-0.17784	-0.40637	0.01957	0.02591	0.48600	0.03253
80	-0.03754	-0.01398	-0.00674	0.01997	0.17955	-0.17700	-0.41056	0.01988	0.02593	0.48677	0.03637
90	-0.03754	-0.01398	-0.00675	0.01999	0.18115	-0.17646	-0.41321	0.02008	0.02594	0.48754	0.03911
100	-0.03754	-0.01398	-0.00675	0.02001	0.18224	-0.17612	-0.41493	0.02021	0.02594	0.48831	0.04092
True values	-0.03800	-0.01400	-0.00800	0.02000	0.18000	-0.17500	-0.39000	0.01650	0.02400	0.50000	

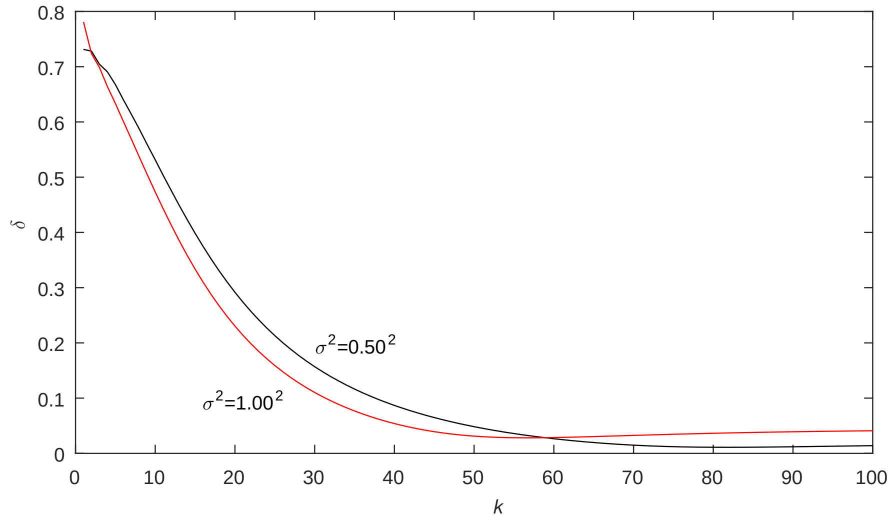

Figure 2: The 2S-OP-GI parameter estimation errors δ versus k with variances $\sigma^2 = 0.50^2$ and $\sigma^2 = 1.00^2$

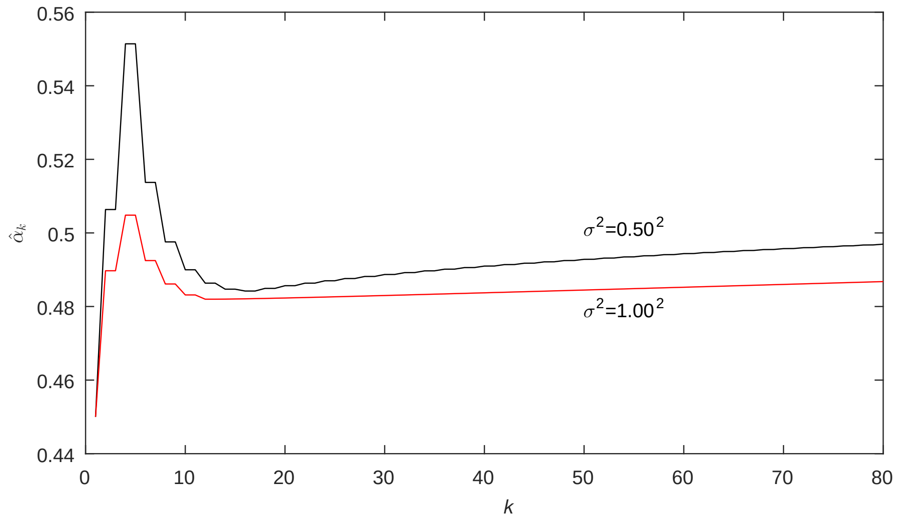

Figure 3: The 2S-OP-GI order estimate $\hat{\alpha}_k$ versus k with variances $\sigma^2 = 0.50^2$ and $\sigma^2 = 1.00^2$

Applying the KT-2S-GI algorithm to estimate parameters of the example system, the parameter estimates and their errors with variance $\sigma^2 = 0.50^2$ and $\sigma^2 = 1.00^2$ are shown in Tables 5–6. The relative parameter estimation errors $\delta := \|\hat{\boldsymbol{\vartheta}}_k - \boldsymbol{\vartheta}\|/\|\boldsymbol{\vartheta}\|$ are shown in Figures 4–5 with different variances.

Table 5: The KT-2S-GI estimates and their errors with $\sigma^2 = 0.50^2$

k	a_1	a_2	c_1	c_2	b_2	b_3	γ_1	γ_2	α	δ
1	0.25876	-0.23522	0.09220	-0.33735	0.00296	-0.05155	6.37678	-0.52996	0.30010	4.34842
10	-0.01683	-0.00092	0.33804	-0.61127	-0.01335	-0.04716	-1.75176	-0.03598	0.30014	0.04750
20	-0.02004	-0.00602	0.34128	-0.60673	-0.01045	-0.01517	-1.75761	-0.03841	0.30014	0.03744
30	-0.02029	-0.00621	0.34109	-0.60661	-0.00917	0.00233	-1.75842	-0.03859	0.30014	0.03364
40	-0.02036	-0.00629	0.34101	-0.60659	-0.00866	0.01187	-1.75861	-0.03869	0.30014	0.03242
50	-0.02039	-0.00633	0.34096	-0.60659	-0.00847	0.01706	-1.75870	-0.03874	0.30014	0.03207
60	-0.02041	-0.00635	0.34093	-0.60659	-0.00842	0.01989	-1.75875	-0.03877	0.30014	0.03199
70	-0.02042	-0.00636	0.34092	-0.60658	-0.00842	0.02144	-1.75878	-0.03879	0.30014	0.03197
80	-0.02042	-0.00637	0.34091	-0.60658	-0.00844	0.02228	-1.75880	-0.03879	0.30014	0.03197
90	-0.02043	-0.00637	0.34090	-0.60658	-0.00846	0.02274	-1.75881	-0.03880	0.30014	0.03198
100	-0.02043	-0.00638	0.34090	-0.60658	-0.00847	0.02299	-1.75881	-0.03880	0.30014	0.03198
True values	-0.02100	-0.00700	0.35000	-0.61000	-0.00800	0.02300	-1.70000	-0.04000	0.30000	0.00000

Table 6: The KT-2S-GI estimates and their errors with $\sigma^2 = 1.00^2$

k	a_1	a_2	c_1	c_2	b_2	b_3	γ_1	γ_2	α	δ
1	0.19260	-0.22320	0.08830	-0.32386	0.00379	-0.05052	7.23389	-0.57125	0.30009	4.80684
10	-0.02236	-0.00452	0.33273	-0.60215	-0.01058	-0.02240	-1.78762	-0.04101	0.30012	0.05393
20	-0.02266	-0.00832	0.33721	-0.59947	-0.00878	0.00912	-1.78298	-0.04324	0.30012	0.04603
30	-0.02272	-0.00844	0.33710	-0.59946	-0.00840	0.01886	-1.78318	-0.04339	0.30012	0.04560
40	-0.02275	-0.00847	0.33705	-0.59945	-0.00835	0.02187	-1.78329	-0.04343	0.30012	0.04561
50	-0.02276	-0.00848	0.33704	-0.59945	-0.00837	0.02281	-1.78333	-0.04344	0.30012	0.04563
60	-0.02276	-0.00849	0.33703	-0.59945	-0.00838	0.02310	-1.78334	-0.04345	0.30012	0.04564
70	-0.02276	-0.00849	0.33703	-0.59945	-0.00838	0.02319	-1.78334	-0.04345	0.30012	0.04564
80	-0.02276	-0.00849	0.33703	-0.59945	-0.00839	0.02322	-1.78334	-0.04345	0.30012	0.04564
90	-0.02276	-0.00849	0.33703	-0.59945	-0.00839	0.02322	-1.78334	-0.04345	0.30012	0.04564
100	-0.02276	-0.00849	0.33703	-0.59945	-0.00839	0.02323	-1.78334	-0.04345	0.30012	0.04564
True values	-0.02100	-0.00700	0.35000	-0.61000	-0.00800	0.02300	-1.70000	-0.04000	0.30000	0.00000

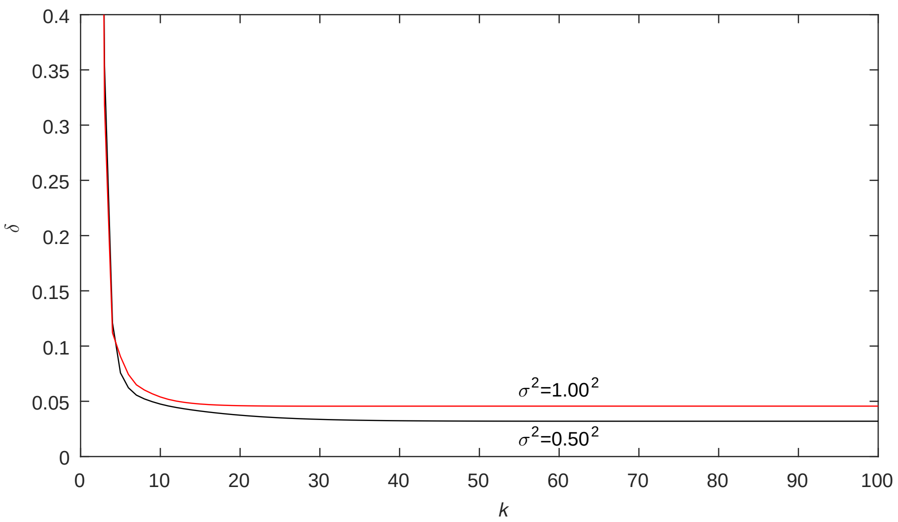

Figure 4: The KT-2S-GI parameter estimation errors δ versus k with variances $\sigma^2 = 0.50^2$ and $\sigma^2 = 1.00^2$

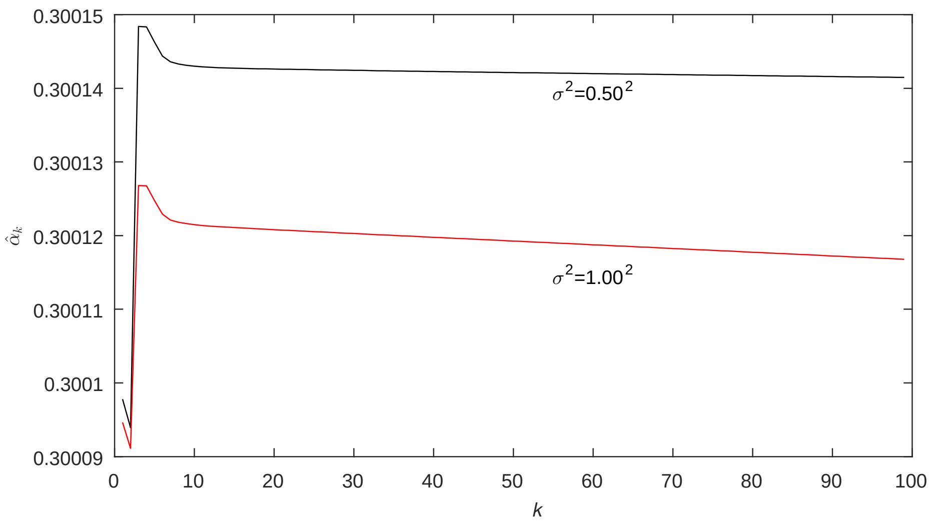

Figure 5: The KT-2S-GI order estimation errors $\hat{\alpha}_k$ versus k with variances $\sigma^2 = 0.50^2$ and $\sigma^2 = 1.00^2$

From Tables 3–6 and Figures 2–5, we can draw the following conclusions.

- Both algorithms can produce mire accurate parameter estimates under lower noise variances.

- With the iterative variable k increasing, the parameter estimation errors given by the 2S-OP-GI and KT-2S-GI algorithms become smaller as shown in Tables 3–6 and Figures 2 and 4.

6.2. A battery model

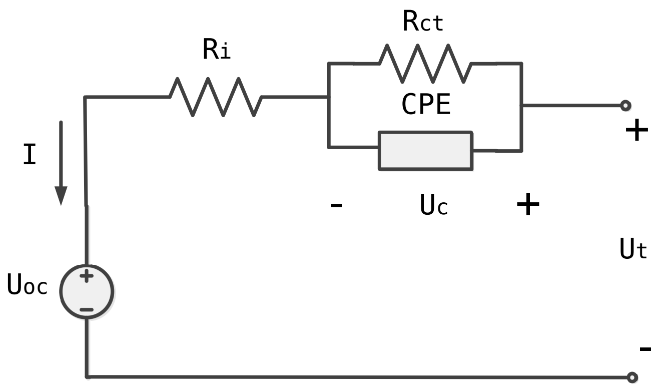

Figure 6: The battery model

Take a battery model as an example. Tian et al. researched a fractional-order battery model [35], as follows:

$$U_t(l) = \frac{(R_i + R_{ct})I(l) + R_i\tau\triangle^{\alpha}I(l) - \tau\triangle^{(\alpha)}U_t(l) + (1+\tau\triangle^{\alpha})U_{oc}(l)}{1+\tau},$$

where $\triangle^{(\alpha)} = \triangle^{\alpha}U_t(l) - U_t(l)$, $\tau = R_{ct}Q$. Obviously, the battery model has strong nonlinearity. The corresponding information vector and parameter vector of this example are given by

$$\begin{aligned}
&\boldsymbol{\theta} := [\frac{R_i + R_{ct}}{1+\tau}, \frac{R_i\tau}{1+\tau}, \frac{\tau}{1+\tau}, \frac{(1+\tau\triangle^{\alpha})U_{oc}(l)}{1+\tau}],\\
&\boldsymbol{\phi} := [I(l), \triangle^{\alpha}I(l), -\triangle^{(\alpha)}U_t(l), 1],\\
&R_i = 0.001(\Omega), R_{ct} = 0.002(\Omega), \alpha = 0.7, Q = 320, U_{oc} = 4(\text{V}).
\end{aligned}$$

The parameters in the battery model may be influenced by many factors, such as temperature, battery state of charge (SOC), capacity degradation, and others. In the process of the experiment, we added a white noise sequence with zero mean and variance σ^2. Figure 7 shows the current and voltage, and use the 2S-OP-GI algorithm to estimate parameters of the example system. From Table 7 and Figures 8–11, we can draw the following conclusions.

- Figure 8 shows a better the voltage fitting effect, and the voltage fitting error is shown in Figure 9.
- It can be seen from Figure 10 that the power of the battery has a better fitting effect and the corresponding error is small.
- The parameter estimates and their errors with variance $\sigma^2 = 0.10^2$ are shown in Table 7 and Figure 11. With the iterative variable k increasing, the parameter estimation errors become smaller.

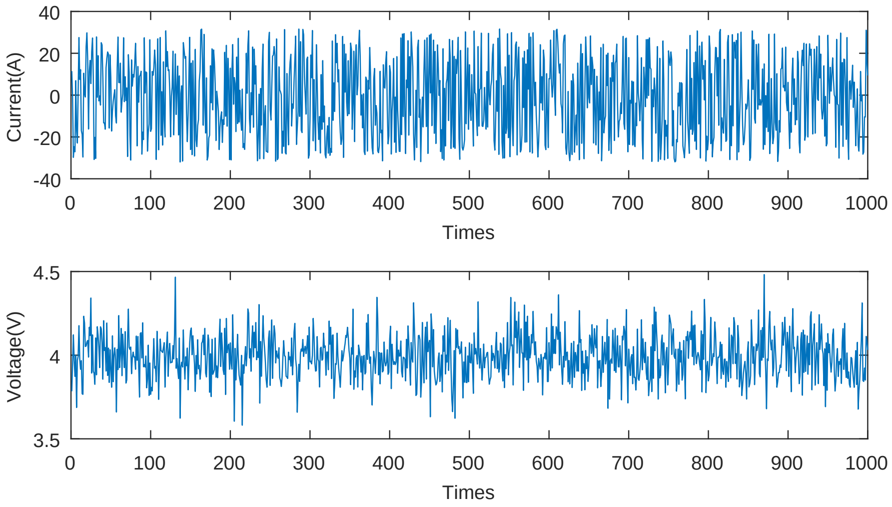

Figure 7: The current and voltage profiles used for simulation

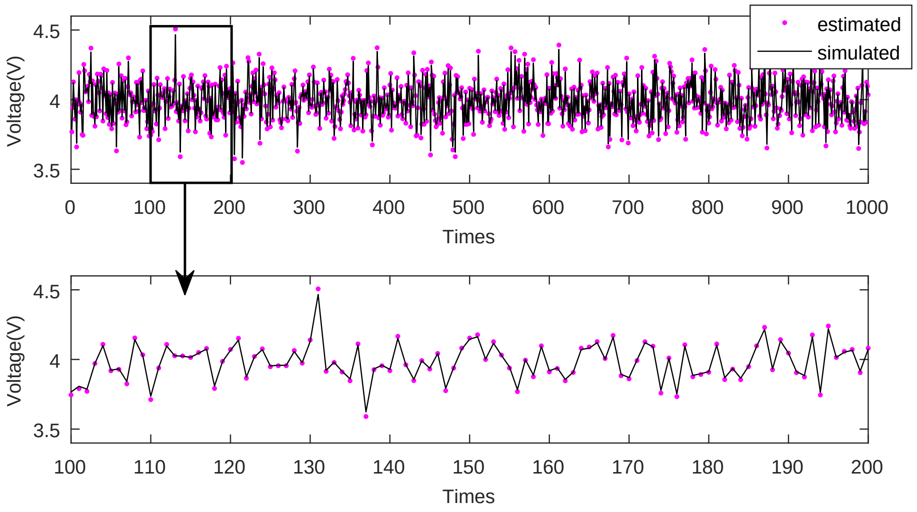

Figure 8: The estimated voltage and simulated voltage profiles

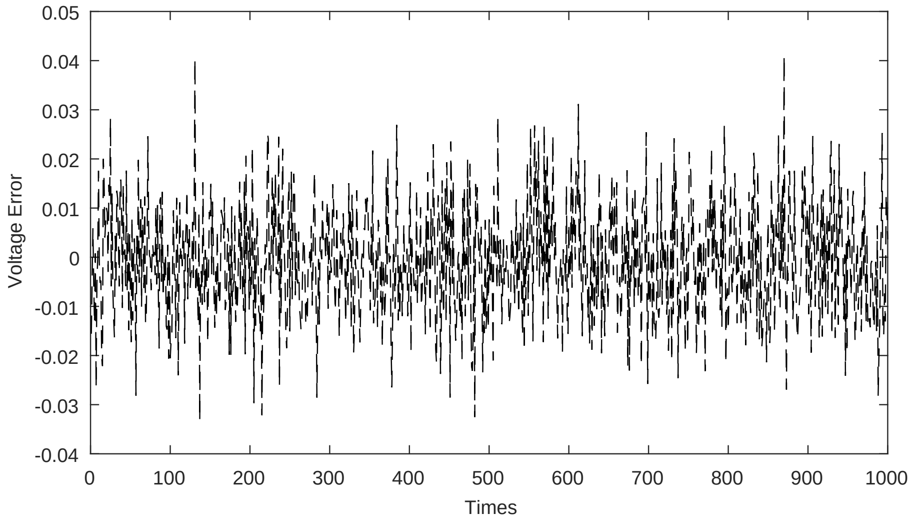

Figure 9: The voltage error profiles

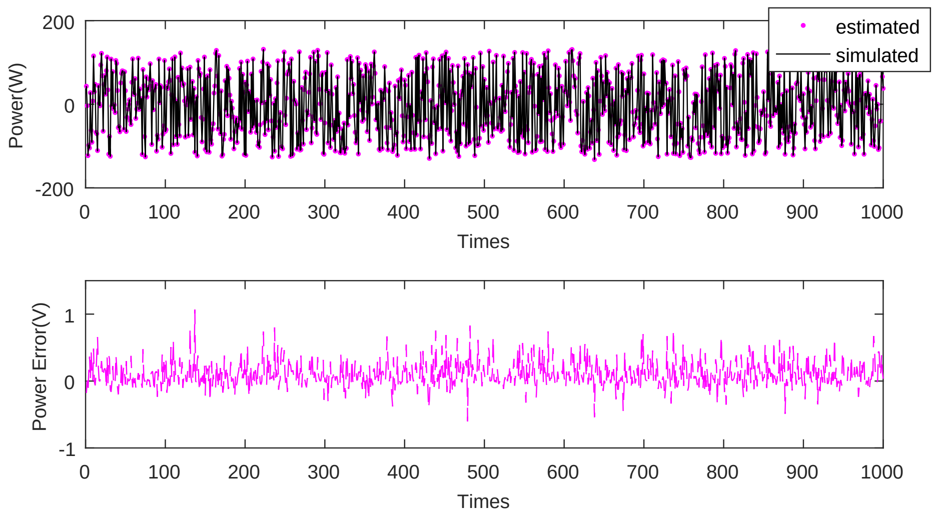

Figure 10: The estimated power and simulated power profiles and power error profiles

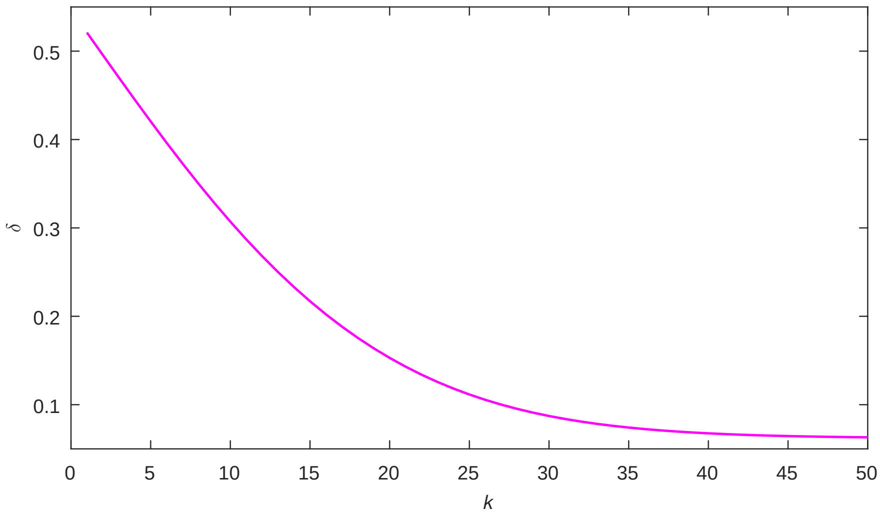

Figure 11: The estimation errors δ versus k with variances $\sigma^2 = 0.10^2$

Table 7: The estimates and their errors with $\sigma^2 = 0.10^2$

k	R_i	R_{ct}	Q	U_{oc}	δ
1	0.00264	0.00522	153.20650	4.09963	0.51205
5	0.00200	0.00430	185.32437	4.09812	0.40989
10	0.00150	0.00359	221.68104	4.09624	0.29475
15	0.00120	0.00316	250.90263	4.09440	0.20349
20	0.00102	0.00290	272.36504	4.09258	0.13887
25	0.00091	0.00274	287.21689	4.09079	0.09794
30	0.00084	0.00265	297.18491	4.08903	0.07529
35	0.00080	0.00258	303.84385	4.08731	0.06480
40	0.00077	0.00254	308.37487	4.08562	0.06100
45	0.00075	0.00251	311.57917	4.08396	0.06029
50	0.00074	0.00248	313.97059	4.08233	0.06085
True values	0.00100	0.00200	320.00000	4.00000	0.00000

7. Conclusions

This paper studies two algorithms, namely a 2S-OP-GI algorithm and a KT-2S-GI algorithm, for identifying the fractional-order block-oriented nonlinear systems based on the hierarchical identification principle. The results of example 1 show that both algorithms are effective. In terms of the computational analysis, the dimension of the matrix involved in the KT-2S-GI algorithm is smaller than that of the 2S-OP-GI algorithm, so the computational efficiency of the KT-2S-GI algorithm is higher. In example 2, a good model fitting is obtained by applying the battery, which further verified the effectiveness of the proposed method. Further research will focus on exploring new identification methods for fractional-order nonlinear systems in combination with other techniques and strategies.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 61472195, No. 61773356) and the Natural Science Foundation of Shandong Province (ZR2020MF160).

Conflict of Interest

The authors declare that they have no conflict of interest.

Data Availability Statement

All data generated or analyzed during this study are included in this article.

References

[1] A. Al-Subhi, Parameters estimation of photovoltaic cells using simple and efficient mathematical models, Solar Energy 209 (2020) 245-257.

[2] Y. Wang, M. Li, Z. Chen, Experimental study of fractional-order models for lithium-ion battery and ultra-capacitor: Modeling, system identification, and validation, Applied Energy 278(15) (2020) 115736.

[3] M. El-Koujok, M. Benammar, N. Meskin, M. Al-Naemi, R. Langari, Multiple sensor fault diagnosis by evolving data-driven approach, Information Sciences 259(221) (2014) 346-358.

[4] F. Ghani, E.F. Fernandez, F. Almonacid, et al., The numerical computation of lumped parameter values using the multi-dimensional Newton-Raphson method for the characterisation of a multi-junction CPV module using the five-parameter approach, Solar Energy 149 (2017) 302-313.

[5] Y.G. Li, J.Q. Chen, F.C. Lan, Enhanced online model identification and state of charge estimation for lithium-ion battery under noise corrupted measurements by bias compensation recursive least squares, Journal of Power Sources 456 (2020) 227984.

[6] J.L. Sun, X.G. Liu, A novel APSO-aided maximum likelihood identification method for Hammerstein systems, Nonlinear Dynamics 73(1-2) (2013) 449-462.

[7] X.M. Liu, Y.M. Fan, Maximum likelihood extended gradient-based estimation algorithms for the input nonlinear controlled autoregressive moving average system with variable-gain nonlinearity, International Journal of Robust and Nonlinear Control 31 (2021) doi:10.1002/rnc.5450.

[8] V. Kulikova, Maximum likelihood estimation of linear stochastic systems in the class of sequential square-root orthogonal filtering methods, Automation and Remote Control 72(4) (2011) 766-786.

[9] X. Zhang, F. Ding, Recursive parameter estimation and its convergence for bilinear systems, IET Control Theory and Applications 14(5) (2020) 677-687.

[10] Y. Ji, X.K. Jiang, L.J. Wan, Hierarchical least squares parameter estimation algorithm for two-input Hammerstein finite impulse response systems, Journal of the Franklin Institute 357(8) (2020) 5019-5032.

[11] M.H. Li, X.M. Liu, Maximum likelihood least squares based iterative estimation for a class of bilinear systems using the data filtering technique, International Journal of Control Automation and Systems 18(6) (2020) 1581-1592.

[12] F. Han, Q. Liu, A diversity-guided hybrid particle swarm optimization based on gradient search, Neurocomputing 137(5) (2014) 234-240.

[13] Y.L. Chu, Y. Fan, Y.Y. He, W.Q. Wu, An improved multi-node Newton iteration search method based on PSO, Optik 232 (2021) 166404.

[14] Z.B. Chen, X.S. Chen, A note on "solving the general Sylvester discrete-time periodic matrix equations via the gradient based iterative method", Applied Mathematics Letters 119 (2021) 107149.

[15] J.H. Li, J.L. Zhang, Maximum likelihood identification of dual-rate Hammerstein output-error moving average system, IET Control Theory and Applications 14(8) (2020) 1089-1101.

[16] X.H. Wang, F. Zhu, F. Ding, The modified extended Kalman filter based recursive estimation for Wiener nonlinear systems with process noise and measurement noise, International Journal of Adaptive Control and Signal Processing 34(10) (2020) 1321-1340.

[17] F. Li, L. Jia, Parameter estimation of Hammerstein-Wiener nonlinear system with noise using special test signals, Neurocomputing 344(7) (2019) 37-48.

[18] L.Sersour, T.Djamah, M.Bettayeb, Fractional Wiener system identification using heuristic optimization technique based on key term principle, IFAC-PapersOnLine 52(13) (2019) 189-193.

[19] Z. Wang, H.R. An, X.L. Luo, Switch detection and robust parameter estimation for slowly switched Hammerstein systems, Nonlinear Analysis: Hybrid Systems 32 (2019) 202-213.

[20] D.Q. Wang, Q.H. Fan, Y. Ma, An interactive maximum likelihood estimation method for multivariable Hammerstein systems, Journal of the Franklin Institute 357(17) (2020) 12986-13005.

[21] Z. Zhang, J. Zhang, Asymptotic stabilization of general nonlinear fractional-order systems with multiple time delays, Nonlinear Dynamics 102 (2020) 605-619.

[22] C. Yin, Y.H. Cheng, Y.Q. Chen, B.D. Stark, S.M. Zhong, Adaptive fractional-order switching-type control method design for 3D fractional-order nonlinear systems, Nonlinear Dynamics 82 (2015) 39 – 52.

[23] Y.X. Kang, S.H. Mao, Y.H. Zhang, Variable order fractional grey model and its application, Applied Mathematical Modelling 97 (2021) 619-635.

[24] Z.R. Lu, G. Liu, J.K. Liu, Y.M. Chen, L. Wang, Parameter identification of nonlinear fractional-order systems by enhanced response sensitivity approach, Nonlinear Dynamics 95 (2019) 1495-1512.

[25] M.H. Hu, Y.X. Li, S.X. Li, C.Y. Fu, D.T. Qin, Lithium-Ion Battery Modeling and Parameter Identification Based on Fractional Theory, Energy 165(15) (2018) 153-163.

[26] K. Hammar, T. Djamah, M. Bettayeb, Nonlinear system identification using fractional Hammerstein-Wiener models, Nonlinear Dynamics 98(3) (2019) 2327-2338.

[27] K. Hammar, T. Djamah, M. Bettayeb, Identification of fractional Hammerstein system with application to a heating process, Nonlinear Dynamics 96 (2019) 2613-2626.

[28] L. Sersour, T. Djamah, M. Bettayeb, Fractional Wiener system identification using heuristic optimization technique based on key term principle, IFAC-PapersOnLine 52(13) (2019) 189-193.

[29] S.J. Dong, L. Yu, W.A. Zhang, B. Chen, Robust hierarchical identification of Wiener systems in the presence of dynamic disturbances, Journal of the Franklin Institute 357(6) (2020) 3809-3834.

[30] B. Niu, Z.D. Bao, D.F. Yu, et al., Hierarchical modeling method based on multilevel architecture surface restriction and its application in point-bar internal architecture of a complex meandering river, Journal of Petroleum Science and Engineering 205 (2021) 108808.

[31] Q.B. Jin, Z. Wang, R.G. Yang, J. Wang, An effective direct closed loop identification method for linear multivariable systems with colored noise, Journal of Process Control 24(5) (2014) 485-492.

[32] A. Golbabai, M. Javidi, Newton-like iterative methods for solving system of non-linear equations, Applied Mathematics and Computation 192(2) (2007) 546-551.

[33] Y.H. Wei, Y. Kang, W.D. Yin, Y. Wang, Generalization of the gradient method with fractional order gradient direction, Journal of the Franklin Institute 357 (2020) 2514-2532.

[34] M.H. Li, X.M. Liu, Filtering-based maximum likelihood gradient iterative estimation algorithm for bilinear systems with autoregressive moving average noise, Circuits Systems and Signal Processing 37(11) (2018) 5023-5048.

[35] J.P. Tian, R. Xiong, W.X. Shen, J. Wang, R.X. Yang, Online simultaneous identification of parameters and order of a fractional order battery model, Journal of Cleaner Production 247(20) (2020) 119147.