



# Hierarchical Principle-Based Iterative Parameter Estimation Algorithm for Dual-Frequency Signals

Siyu Liu<sup>1</sup> · Feng Ding<sup>1,2,3</sup> · Ling Xu<sup>1</sup> · Tasawar Hayat<sup>3</sup>

Received: 30 January 2018 / Revised: 17 December 2018 / Accepted: 17 December 2018 /  
Published online: 2 January 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

In this paper, we consider the parameter estimation problem of dual-frequency signals disturbed by stochastic noise. The signal model is a highly nonlinear function with respect to the frequencies and phases, and the gradient method cannot obtain the accurate parameter estimates. Based on the Newton search, we derive an iterative algorithm for estimating all parameters, including the unknown amplitudes, frequencies, and phases. Furthermore, by using the parameter decomposition, a hierarchical least squares and gradient-based iterative algorithm is proposed for improving the computational efficiency. A gradient-based iterative algorithm is given for comparisons. The numerical examples are provided to demonstrate the validity of the proposed algorithms.

**Keywords** Iterative algorithm · Signal modeling · Hierarchical identification · Least squares · Newton search · Gradient search

---

✉ Feng Ding  
fding@jiangnan.edu.cn

Siyu Liu  
sylvu12@126.com

Ling Xu  
lingxu0848@163.com

Tasawar Hayat  
fmgpak@gmail.com

<sup>1</sup> Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, People's Republic of China

<sup>2</sup> School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan 430068, People's Republic of China

<sup>3</sup> Department of Mathematics, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia

## 1 Introduction

Signal models are often encountered in scientific disciplines and engineering applications. The parameter estimation of signal models is very important, and then many estimation approaches have been presented [32,41,42]. As an original way, Fourier analysis plays a key role in signal processing [48,50]. However, there are certain limitations because the frequencies of the Fourier transform are just a collection of the individual frequencies of the periodic signals. Another classical method is the phase-locked loop technique, which automatically adjusts the frequency by locking the phase between the reference and the internal sine waves. Additionally, a self-correlation phase method [4] has been developed to estimate the frequencies. The parameter estimation methods can be applied to many areas [2,3,5,51].

Many methods can estimate the parameters of the signal models. For example, Stojica et al. [37] used the weighted least squares method to estimate the amplitudes of the sinusoidal signals. The amplitude and frequency of a sine signal were estimated with the property of global convergence in [20]. The amplitude and frequency estimation problem of multiple sine signals was investigated based on the adaptive identifier [21]. Recently, Hou [22] has presented a solution for estimating parameters of a biased multiple sinusoids using a combination of a frequency identifier and gradient algorithms. This paper is devoted to identify the parameters for dual-frequency signals, including the unknown amplitudes, frequencies, and phases.

The iterative methods play an important role not only in finding the solutions of nonlinear matrix equations, but also in deriving parameter estimation algorithms for signal models. Combining with some optimization tools, the iterative methods can be applied to estimate the parameters of the nonlinear systems [9,34]. For example, Xu [40] derived a Newton iterative algorithm for the parameter estimation of dynamical systems. Ding et al. [12] studied the iterative parameter identification for pseudo-linear systems with ARMA noise using the filtering technique.

As a classical optimization tool, the Newton method is a root-finding algorithm that utilizes the first few terms of the Taylor series of a function in the neighborhood of a suspected root. For a long time, the Newton method has received a great deal of attention, and there have been lots of applications in many fields, such as the minimization and maximization problems, the transcendental equations, the numerical verification for the solutions of nonlinear equations [30,35,49].

The previous work in [29,43] considered the Newton iterative (NI) algorithm for the nonlinear parameter estimation problems. The NI algorithm possesses high parameter estimation accuracy, while the computational cost is large for the reason that the inverse of the Hessian matrix needs to be computed at each iteration, and its components are second-order derivatives. In order to reduce the computational cost, this paper presents a hierarchical least squares and gradient-based iterative (HI) algorithm using the parameter decomposition. The key is to decompose the parameter vector into three parts and to estimate the parameters of each part, so the dimension of the parameter vector is reduced and the computational efficiency and the parameter estimation accuracy are improved.

The rest of this paper is organized as follows. Sections 2 and 3 derive two iterative algorithms based on the gradient search and the Newton search, respectively. Sec-

tion 4 provides a HI algorithm. Section 5 gives two numerical examples to illustrate the effectiveness of the proposed algorithms. Finally, some conclusions are drawn in Sect. 6.

## 2 The Gradient-Based Iterative Algorithm

The sine and cosine signals are typical periodic signals, whose waveforms are the sine and cosine curves in mathematics. Many complex signals can be decomposed into numerous sinusoidal signals with different frequencies and amplitudes by the Fourier series [1]. The sine signal differs from the cosine signal by  $\pi/2$  in the initial phase. This paper derives the parameter estimation algorithms for the dual-frequency cosine signals, which can be applied to the sine signals.

Consider the following dual-frequency cosine signal model:

$$y(t) = a \cos(\omega_1 t + \phi) + b \cos(\omega_2 t + \psi) + v(t), \quad (1)$$

where  $a > 0$  and  $b > 0$  are the amplitudes,  $\omega_1 > 0$  and  $\omega_2 > 0$  are the angular frequencies,  $\phi$  and  $\psi$  are the initial phases,  $t$  is a continuous-time variable,  $y(t)$  is the observations, and  $v(t)$  is a stochastic disturbance with zero mean and variance  $\sigma^2$ .

The actual signals are always interfered with various stochastic noises. In practical engineering, we can only get the discrete observed data. Suppose that the sampled data are  $y(t_i)$ ,  $i = 1, 2, 3, \dots, L$ , where  $t_i$  is the sampled time and  $L$  denotes the data length.

Use the observed data  $y(t_i)$  and the model output to construct the criterion function

$$J_1(\boldsymbol{\theta}) := \frac{1}{2} \sum_{i=1}^L [y(t_i) - a \cos(\omega_1 t_i + \phi) - b \cos(\omega_2 t_i + \psi)]^2.$$

The criterion function represents the square sum of the errors between the model outputs and the observed data. It is desired that this criterion function is as small as possible, which is equivalent to minimizing  $J_1(\boldsymbol{\theta})$  and obtaining the estimates of the parameter vector  $\boldsymbol{\theta} := [a, b, \omega_1, \omega_2, \phi, \psi]^T \in \mathbb{R}^6$ .

Letting the partial derivative of  $J_1(\boldsymbol{\theta})$  with respect to  $\boldsymbol{\theta}$  gives

$$\begin{aligned} \text{grad}[J_1(\boldsymbol{\theta})] &:= \frac{\partial J_1(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\ &= \left[ \frac{\partial J_1(\boldsymbol{\theta})}{\partial a}, \frac{\partial J_1(\boldsymbol{\theta})}{\partial b}, \frac{\partial J_1(\boldsymbol{\theta})}{\partial \omega_1}, \frac{\partial J_1(\boldsymbol{\theta})}{\partial \omega_2}, \frac{\partial J_1(\boldsymbol{\theta})}{\partial \phi}, \frac{\partial J_1(\boldsymbol{\theta})}{\partial \psi} \right]^T \in \mathbb{R}^6. \end{aligned}$$

Define the stacked observed data  $\mathbf{Y}(L)$  and the stacked information matrix  $\boldsymbol{\Phi}(\boldsymbol{\theta}, L)$  as

$$\begin{aligned} \mathbf{Y}(L) &:= [y(t_1), y(t_2), \dots, y(t_L)]^T \in \mathbb{R}^L, \\ \boldsymbol{\Phi}(\boldsymbol{\theta}, L) &:= [\boldsymbol{\varphi}(\boldsymbol{\theta}, t_1), \boldsymbol{\varphi}(\boldsymbol{\theta}, t_2), \dots, \boldsymbol{\varphi}(\boldsymbol{\theta}, t_L)]^T \in \mathbb{R}^{L \times 6}. \end{aligned}$$

Define the information vector

$$\boldsymbol{\varphi}(\boldsymbol{\theta}, t_i) := [-\cos(\omega_1 t_i + \phi), -\cos(\omega_2 t_i + \psi), at_i \sin(\omega_1 t_i + \phi), bt_i \sin(\omega_2 t_i + \psi), a \sin(\omega_1 t_i + \phi), b \sin(\omega_2 t_i + \psi)]^T \in \mathbb{R}^6.$$

Let  $f(t_i) := a \cos(\omega_1 t_i + \phi) + b \cos(\omega_2 t_i + \psi) \in \mathbb{R}$ , and the stacked model output vector

$$\mathbf{F}(\boldsymbol{\theta}, L) := \begin{bmatrix} f(\boldsymbol{\theta}, t_1) \\ f(\boldsymbol{\theta}, t_2) \\ \vdots \\ f(\boldsymbol{\theta}, t_L) \end{bmatrix} \in \mathbb{R}^L.$$

Then, the gradient vector can be expressed as

$$\text{grad}[J_1(\boldsymbol{\theta})] := \boldsymbol{\Phi}^T(\boldsymbol{\theta}, L)[\mathbf{Y}(L) - \mathbf{F}(\boldsymbol{\theta}, L)].$$

Let  $k = 1, 2, 3, \dots$  be an iterative variable and  $\hat{\boldsymbol{\theta}}_k$  be the estimate of  $\boldsymbol{\theta}$  at iteration  $k$ . The partial derivatives at  $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{k-1}$  are given by

$$\frac{\partial J_1(\hat{\boldsymbol{\theta}}_{k-1})}{\partial a} = - \sum_{i=1}^L \cos(\hat{\omega}_{1,k-1} t_i + \hat{\phi}_{k-1}) w(\hat{\boldsymbol{\theta}}_{k-1}, t_i),$$

$$\frac{\partial J_1(\hat{\boldsymbol{\theta}}_{k-1})}{\partial b} = - \sum_{i=1}^L \cos(\hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1}) w(\hat{\boldsymbol{\theta}}_{k-1}, t_i),$$

$$\frac{\partial J_1(\hat{\boldsymbol{\theta}}_{k-1})}{\partial \omega_1} = \sum_{i=1}^L \hat{a}_{k-1} t_i \sin(\hat{\omega}_{1,k-1} t_i + \hat{\phi}_{k-1}) w(\hat{\boldsymbol{\theta}}_{k-1}, t_i),$$

$$\frac{\partial J_1(\hat{\boldsymbol{\theta}}_{k-1})}{\partial \omega_2} = \sum_{i=1}^L \hat{b}_{k-1} t_i \sin(\hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1}) w(\hat{\boldsymbol{\theta}}_{k-1}, t_i),$$

$$\frac{\partial J_1(\hat{\boldsymbol{\theta}}_{k-1})}{\partial \phi} = \sum_{i=1}^L \hat{a}_{k-1} \sin(\hat{\omega}_{1,k-1} t_i + \hat{\phi}_{k-1}) w(\hat{\boldsymbol{\theta}}_{k-1}, t_i),$$

$$\frac{\partial J_1(\hat{\boldsymbol{\theta}}_{k-1})}{\partial \psi} = \sum_{i=1}^L \hat{b}_{k-1} \sin(\hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1}) w(\hat{\boldsymbol{\theta}}_{k-1}, t_i),$$

$$w(\hat{\boldsymbol{\theta}}_{k-1}, t_i) := y(t_i) - \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1} t_i + \hat{\phi}_{k-1}) - \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1}).$$

Minimizing  $J_1(\boldsymbol{\theta})$  and selecting an iterative step size  $\mu_k \geq 0$ , we can get the gradient-based iterative (GI) algorithm for dual-frequency signal models:

$$\begin{aligned} \hat{\theta}_k &= \hat{\theta}_{k-1} - \mu_k \text{grad} \left[ J_1 \left( \hat{\theta}_{k-1} \right) \right] \\ &= \hat{\theta}_{k-1} - \mu_k \Phi^T \left( \hat{\theta}_{k-1}, L \right) \left[ Y(L) - F \left( \hat{\theta}_{k-1}, L \right) \right], \end{aligned} \tag{2}$$

$$Y(L) = [y(t_1), y(t_2), \dots, y(t_L)]^T, \tag{3}$$

$$\Phi \left( \hat{\theta}_{k-1}, L \right) = \left[ \varphi \left( \hat{\theta}_{k-1}, t_1 \right), \varphi \left( \hat{\theta}_{k-1}, t_2 \right), \dots, \varphi \left( \hat{\theta}_{k-1}, t_L \right) \right]^T, \tag{4}$$

$$F \left( \hat{\theta}_{k-1}, L \right) = \left[ f \left( \hat{\theta}_{k-1}, t_1 \right), f \left( \hat{\theta}_{k-1}, t_2 \right), \dots, f \left( \hat{\theta}_{k-1}, t_L \right) \right]^T, \tag{5}$$

$$\begin{aligned} \varphi \left( \hat{\theta}_{k-1}, t_i \right) &= \left[ -\cos \left( \hat{\omega}_{1,k-1} t_i + \hat{\phi}_{k-1} \right), -\cos \left( \hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1} \right), \right. \\ &\quad \hat{a}_{k-1} t_i \sin \left( \hat{\omega}_{1,k-1} t_i + \hat{\phi}_{k-1} \right), \\ &\quad \hat{b}_{k-1} t_i \sin \left( \hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1} \right), \hat{a}_{k-1} \sin \left( \hat{\omega}_{1,k-1} t_i + \hat{\phi}_{k-1} \right), \\ &\quad \left. \hat{b}_{k-1} \sin \left( \hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1} \right) \right]^T, \end{aligned} \tag{6}$$

$$f \left( \hat{\theta}_{k-1}, t_i \right) = \hat{a}_{k-1} \cos \left( \hat{\omega}_{1,k-1} t_i + \hat{\phi}_{k-1} \right) + \hat{b}_{k-1} \cos \left( \hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1} \right), \tag{7}$$

$$\mu_k = \arg \min J_1 \left( \hat{\theta}_{k-1} - \mu \text{grad} \left[ J_1 \left( \hat{\theta}_{k-1} \right) \right] \right), \tag{8}$$

$$\hat{\theta}_k = \left[ \hat{a}_k, \hat{b}_k, \hat{\omega}_{1,k}, \hat{\omega}_{2,k}, \hat{\phi}_k, \hat{\psi}_k \right]^T. \tag{9}$$

$\mu_k \leq \lambda_{\max}^{-1} [\Phi^T(\hat{\theta}_{k-1}, L)\Phi(\hat{\theta}_{k-1}, L)]$ , where the  $\lambda_{\max}[\Phi^T(\hat{\theta}_{k-1}, L)\Phi(\hat{\theta}_{k-1}, L)]$  denotes the maximum eigenvalue of the matrix  $[\Phi^T(\hat{\theta}_{k-1}, L)\Phi(\hat{\theta}_{k-1}, L)]$ . The procedure of the GI algorithm in (2)–(9) for dual-frequency signals is as follows.

1. Let  $k = 1$ , give a small number  $\varepsilon$  and an integer  $L$ , set the initial value  $\hat{\theta}_0 = \mathbf{1}_6/p_0$ , where  $p_0$  is generally taken as a large positive number, for example,  $p_0 = 10^6$ .
2. Collect the observed data  $y(t_i)$  and  $u(t_i)$ ,  $i = 1, 2, \dots, L$ , form  $Y(L)$  by (3).
3. Form the information vector  $\varphi(\hat{\theta}_{k-1}, t_i)$  by (6) and the stacked information matrix  $\Phi(\hat{\theta}_{k-1}, L)$  by (4).
4. Form  $f(\hat{\theta}_{k-1}, t_i)$  by (7) and the stacked model output vector  $F(\hat{\theta}_{k-1}, L)$  by (5).
5. Select a largest  $\mu_k$  satisfying (8), and update the estimate  $\hat{\theta}_k$  by (9).
6. Compare  $\hat{\theta}_k$  with  $\hat{\theta}_{k-1}$ , if  $\|\hat{\theta}_k - \hat{\theta}_{k-1}\| \geq \varepsilon$ , then increase  $k$  by 1 and go to Step 3; otherwise, obtain the iterative time  $k$  and the estimate  $\hat{\theta}_k$  and terminate this procedure.

**Remark 1** In fact, the one-dimensional search for finding the best step size  $\mu_k$  needs to solve a complicated nonlinear equation at each iteration, so we cannot obtain the algebraic solution for computing  $\mu_k$  by Eq. (8). To figure out this problem, we can use the cut-and-try method or the Newton iterative method to acquire the approximation. In what follows, we derive a Newton iterative algorithm for estimating the parameters.

### 3 The Newton Iterative Algorithm

The gradient method requires the first derivative, but the Newton method is quadratically convergent. In order to improve the parameter estimation accuracy, calculating the second partial derivative of the criterion function  $J_1(\boldsymbol{\theta})$  with respect to  $\boldsymbol{\theta}$  gives the Hessian matrix

$$\mathbf{H}(\boldsymbol{\theta}) := \frac{\partial^2 J_1(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} = \frac{\partial \text{grad}[J_1(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}^T} = \begin{bmatrix} h_{11}(\boldsymbol{\theta}) & h_{12}(\boldsymbol{\theta}) & \cdots & h_{16}(\boldsymbol{\theta}) \\ h_{21}(\boldsymbol{\theta}) & h_{22}(\boldsymbol{\theta}) & \cdots & h_{26}(\boldsymbol{\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ h_{61}(\boldsymbol{\theta}) & h_{62}(\boldsymbol{\theta}) & \cdots & h_{66}(\boldsymbol{\theta}) \end{bmatrix} \in \mathbb{R}^{6 \times 6},$$

$$h_{11}(\boldsymbol{\theta}) = \frac{\partial^2 J_1(\boldsymbol{\theta})}{\partial a^2} = \sum_{i=1}^L \cos^2(\omega_1 t_i + \phi),$$

$$h_{12}(\boldsymbol{\theta}) = \frac{\partial^2 J_1(\boldsymbol{\theta})}{\partial a \partial b} = \sum_{i=1}^L \cos(\omega_1 t_i + \phi) \cos(\omega_2 t_i + \psi),$$

$$h_{13}(\boldsymbol{\theta}) = \frac{\partial^2 J_1(\boldsymbol{\theta})}{\partial a \partial \omega_1} = \sum_{i=1}^L t_i \sin(\omega_1 t_i + \phi) [y(t_i) - b \cos(\omega_2 t_i + \psi)] - a t_i \sin(2\omega_1 t_i + 2\phi),$$

$$h_{14}(\boldsymbol{\theta}) = \frac{\partial^2 J_1(\boldsymbol{\theta})}{\partial a \partial \omega_2} = - \sum_{i=1}^L b t_i \cos(\omega_1 t_i + \phi) \sin(\omega_2 t_i + \psi),$$

$$h_{15}(\boldsymbol{\theta}) = \frac{\partial^2 J_1(\boldsymbol{\theta})}{\partial a \partial \phi} = \sum_{i=1}^L \sin(\omega_1 t_i + \phi) [y(t_i) - b \cos(\omega_2 t_i + \psi)] - \sin(2\omega_1 t_i + 2\phi),$$

$$h_{16}(\boldsymbol{\theta}) = \frac{\partial^2 J_1(\boldsymbol{\theta})}{\partial a \partial \psi} = - \sum_{i=1}^L b \cos(\omega_1 t_i + \phi) \sin(\omega_2 t_i + \psi).$$

For the sake of brevity, the derivation results of all  $h_{mn}$ ,  $1 = n < m \leq 6$  are not listed here, and the rest are presented in (10)–(39). Let  $\hat{\boldsymbol{\theta}}_k := [\hat{a}_k, \hat{b}_k, \hat{\omega}_{1,k}, \hat{\omega}_{2,k}, \hat{\phi}_k, \hat{\psi}_k]^T$  be the estimate of  $\boldsymbol{\theta}$  at iteration  $k$ . Based on the Newton search, we can derive the Newton iterative (NI) algorithm for dual-frequency signal models:

$$\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_{k-1} - \mathbf{H}^{-1}(\hat{\boldsymbol{\theta}}_{k-1}) \text{grad} [J_1(\hat{\boldsymbol{\theta}}_{k-1})], \tag{10}$$

$$\mathbf{Y}(L) = [y(t_1), y(t_2), \dots, y(t_L)]^T, \tag{11}$$

$$\hat{\boldsymbol{\Phi}}_k = \boldsymbol{\Phi}(\hat{\boldsymbol{\theta}}_{k-1}, L) = [\boldsymbol{\varphi}(\hat{\boldsymbol{\theta}}_{k-1}, t_1), \boldsymbol{\varphi}(\hat{\boldsymbol{\theta}}_{k-1}, t_2), \dots, \boldsymbol{\varphi}(\hat{\boldsymbol{\theta}}_{k-1}, t_L)]^T, \tag{12}$$

$$\hat{\mathbf{F}}_k = \mathbf{F}(\hat{\boldsymbol{\theta}}_{k-1}, L) = [f(\hat{\boldsymbol{\theta}}_{k-1}, t_1), f(\hat{\boldsymbol{\theta}}_{k-1}, t_2), \dots, f(\hat{\boldsymbol{\theta}}_{k-1}, t_L)]^T, \tag{13}$$

$$\begin{aligned} \hat{\boldsymbol{\varphi}}_k(t_i) &= \boldsymbol{\varphi}(\hat{\boldsymbol{\theta}}_{k-1}, t_i) \\ &= [-\cos(\hat{\omega}_{1,k-1} t_i + \hat{\phi}_{k-1}), -\cos(\hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1})], \end{aligned}$$

$$\begin{aligned} & \hat{a}_{k-1}t_i \sin(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}), \\ & \hat{b}_{k-1}t_i \sin(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}), \hat{a}_{k-1} \sin(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}), \\ & \hat{b}_{k-1} \sin(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) \Big]^\top, \end{aligned} \tag{14}$$

$$\begin{aligned} \hat{f}_k(t_i) &= f(\hat{\theta}_{k-1}, t_i) \\ &= \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) + \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}), \end{aligned} \tag{15}$$

$$\mathbf{H}(\hat{\theta}_{k-1}) = \begin{bmatrix} h_{11}(\hat{\theta}_{k-1}) & h_{12}(\hat{\theta}_{k-1}) & \cdots & h_{16}(\hat{\theta}_{k-1}) \\ h_{21}(\hat{\theta}_{k-1}) & h_{22}(\hat{\theta}_{k-1}) & \cdots & h_{26}(\hat{\theta}_{k-1}) \\ \vdots & \vdots & \ddots & \vdots \\ h_{61}(\hat{\theta}_{k-1}) & h_{62}(\hat{\theta}_{k-1}) & \cdots & h_{66}(\hat{\theta}_{k-1}) \end{bmatrix}, \tag{16}$$

$$h_{11}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \cos^2(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}), \tag{17}$$

$$h_{12}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}), \tag{18}$$

$$\begin{aligned} h_{13}(\hat{\theta}_{k-1}) &= \sum_{i=1}^L t_i \sin(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \left[ y(t_i) - \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) \right] \\ &\quad - \hat{a}_{k-1}t_i \sin(2\hat{\omega}_{1,k-1}t_i + 2\hat{\phi}_{k-1}), \end{aligned} \tag{19}$$

$$h_{14}(\hat{\theta}_{k-1}) = - \sum_{i=1}^L \hat{b}_{k-1}t_i \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \sin(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}), \tag{20}$$

$$\begin{aligned} h_{15}(\hat{\theta}_{k-1}) &= \sum_{i=1}^L \sin(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \left[ y(t_i) - \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) \right] \\ &\quad - \sin(2\hat{\omega}_{1,k-1}t_i + 2\hat{\phi}_{k-1}), \end{aligned} \tag{21}$$

$$h_{16}(\hat{\theta}_{k-1}) = - \sum_{i=1}^L \hat{b}_{k-1} \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \sin(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}), \tag{22}$$

$$h_{22}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \cos^2(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}), \tag{23}$$

$$h_{23}(\hat{\theta}_{k-1}) = - \sum_{i=1}^L \hat{a}_{k-1}t_i \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) \sin(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}), \tag{24}$$

$$\begin{aligned} h_{24}(\hat{\theta}_{k-1}) &= \sum_{i=1}^L t_i \sin(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) \left[ y(t_i) - \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \right] \\ &\quad - \hat{b}_{k-1}t_i \sin(2\hat{\omega}_{2,k-1}t_i + 2\hat{\psi}_{k-1}), \end{aligned} \tag{25}$$

$$h_{25}(\hat{\theta}_{k-1}) = -\sum_{i=1}^L \hat{a}_{k-1} \sin(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}), \quad (26)$$

$$h_{26}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \sin(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) \left[ y(t_i) - \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \right] - \sin(2\hat{\omega}_{2,k-1}t_i + 2\hat{\psi}_{k-1}), \quad (27)$$

$$h_{33}(\hat{\theta}_{k-1}) = \sum_{i=1}^L -\hat{a}_{k-1} \hat{b}_{k-1} t_i^2 \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) + \hat{a}_{k-1} t_i^2 y(t_i) \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) - \hat{a}_{k-1}^2 t_i^2 \cos(2\hat{\omega}_{1,k-1}t_i + 2\hat{\phi}_{k-1}), \quad (28)$$

$$h_{34}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \hat{a}_{k-1} \hat{b}_{k-1} t_i^2 \sin(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \sin(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}), \quad (29)$$

$$h_{35}(\hat{\theta}_{k-1}) = \sum_{i=1}^L -\hat{a}_{k-1} \hat{b}_{k-1} t_i \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) + \hat{a}_{k-1} t_i \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) y(t_i) - \hat{a}_{k-1}^2 t_i \cos(2\hat{\omega}_{1,k-1}t_i + 2\hat{\phi}_{k-1}), \quad (30)$$

$$h_{36}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \hat{a}_{k-1} \hat{b}_{k-1} t_i \sin(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \sin(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}), \quad (31)$$

$$h_{44}(\hat{\theta}_{k-1}) = \sum_{i=1}^L -\hat{a}_{k-1} \hat{b}_{k-1} t_i^2 \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) - \hat{b}_{k-1}^2 t_i^2 \cos(2\hat{\omega}_{2,k-1}t_i + 2\hat{\psi}_{k-1}) + \hat{b}_{k-1} t_i^2 \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) y(t_i), \quad (32)$$

$$h_{45}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \hat{a}_{k-1} \hat{b}_{k-1} t_i \sin(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) \sin(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}), \quad (33)$$

$$h_{46}(\hat{\theta}_{k-1}) = \sum_{i=1}^L -\hat{a}_{k-1} \hat{b}_{k-1} t_i \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) + \hat{b}_{k-1} t_i \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) y(t_i) - \hat{b}_{k-1}^2 t_i \cos(2\hat{\omega}_{2,k-1}t_i + 2\hat{\psi}_{k-1}), \quad (34)$$

$$h_{55}(\hat{\theta}_{k-1}) = \sum_{i=1}^L -\hat{a}_{k-1} \hat{b}_{k-1} \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) \cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) + \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}) y(t_i) - \hat{a}_{k-1}^2 \cos(2\hat{\omega}_{1,k-1}t_i + 2\hat{\phi}_{k-1}), \quad (35)$$



$$h_{56}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \hat{a}_{k-1} \hat{b}_{k-1} \sin(\hat{\omega}_{1,k-1} t_i + \hat{\phi}_{k-1}) \sin(\hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1}), \tag{36}$$

$$h_{66}(\hat{\theta}_{k-1}) = \sum_{i=1}^L -\hat{a}_{k-1} \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1}) \cos(\hat{\omega}_{1,k-1} t_i + \hat{\phi}_{k-1}) + \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1} t_i + \hat{\psi}_{k-1}) y(t_i) - \hat{b}_{k-1}^2 \cos(2\hat{\omega}_{2,k-1} t_i + 2\hat{\psi}_{k-1}), \tag{37}$$

$$h_{mn}(\hat{\theta}_{k-1}) = h_{nm}(\hat{\theta}_{k-1}), \quad 1 = n < m \leq 6, \tag{38}$$

$$\hat{\theta}_k = [\hat{a}_k, \hat{b}_k, \hat{\omega}_{1,k}, \hat{\omega}_{2,k}, \hat{\phi}_k, \hat{\psi}_k]^T. \tag{39}$$

The steps of the NI algorithm for computing  $\hat{\theta}_k$  are listed as follows.

1. To initialize, let  $k = 1$ ,  $\hat{\theta}_0 = \mathbf{1}_6/p_0$ , and give a small number  $\varepsilon$ .
2. Collect the observed data  $y(t_i)$  and  $u(t_i)$ ,  $i = 1, 2, \dots, L$ , form  $\mathbf{Y}(L)$  by (11).
3. Form the information vector  $\hat{\phi}_k(t_i)$  using (14), the stacked information matrix  $\hat{\Phi}_k$  using (12).
4. Form  $\hat{f}_k(t_i)$  using (15) and the stacked model output vector  $\hat{F}_k$  using (13).
5. Compute  $h_{mn}(\hat{\theta}_{k-1})$  by (17)–(38),  $1 = n < m \leq 6$ , form the Hessian matrix  $\mathbf{H}(\hat{\theta}_{k-1})$  by (16).
6. Update the estimate  $\hat{\theta}_k$  by (39).
7. Compare  $\hat{\theta}_k$  with  $\hat{\theta}_{k-1}$ , if  $\|\hat{\theta}_k - \hat{\theta}_{k-1}\| > \varepsilon$ , then increase  $k$  by 1, and go to Step 3; otherwise, obtain the estimate  $\hat{\theta}_k$  and terminate this procedure.

**Remark 2** The NI algorithm is an effective parameter estimation algorithm; however, it must calculate the inverse of the Hessian matrix at each iteration, which requires much computational cost. The following will derive a hierarchical least squares and gradient-based iterative algorithm, which not only ameliorates the estimation accuracy but also improves the computational efficiency.

### 4 The Hierarchical Iterative Algorithm

Inspired initially by the hierarchical control for large-scale systems based on the decomposition–coordination principle, we derive a hierarchical identification algorithm for the parameter estimation problem in this paper. According to the parameter decomposition, the original optimization problem is transformed into a combination of the nonlinear optimization and the linear optimization. In this way, the computational complexity is reduced greatly. The hierarchical method can be devoted to signal processing [44,47], nonlinear systems [26,31,52–54], and other fields [10,11,25].

Define three parameter vectors  $\alpha := [a, b]^T \in \mathbb{R}^2$ ,  $\omega := [\omega_1, \omega_2]^T \in \mathbb{R}^2$  and  $\beta := [\phi, \psi]^T \in \mathbb{R}^2$ . The information vector with respect to  $\alpha$  can be written as

$$\varphi_\alpha(t_i) = [-\cos(\omega_1 t_i + \phi), -\cos(\omega_2 t_i + \psi)]^T \in \mathbb{R}^2.$$

Define the information matrix with respect to  $\alpha$  as

$$\Phi_\alpha(L) := [\varphi_\alpha(t_1), \varphi_\alpha(t_2), \dots, \varphi_\alpha(t_L)]^T \in \mathbb{R}^{2 \times L}.$$

Define three criterion functions:

$$J_2(\alpha) := \frac{1}{2} \|Y(L) - \Phi_\alpha(L)\alpha\|^2,$$

$$J_3(\omega) := \frac{1}{2} \sum_{i=1}^L e^2(\omega, t_i),$$

$$J_4(\beta) := \frac{1}{2} \sum_{i=1}^L e^2(\beta, t_i),$$

where  $e(\omega, t_i) = e(\beta, t_i) = y(t_i) - a \cos(\omega_1 t_i + \phi) - b \cos(\omega_2 t_i + \psi)$ . Letting the first-order derivative of  $J_2(\alpha)$  with respect to  $\alpha$  be zero, we have

$$\frac{\partial J_2(\alpha)}{\partial \alpha} = -\Phi_\alpha^T(L)[Y(L) - \Phi_\alpha(L)\alpha] = \mathbf{0}.$$

The least squares (LS) solution is given by

$$\alpha = [\Phi_\alpha^T(L)\Phi_\alpha(L)]^{-1}\Phi_\alpha(L)Y(L). \quad (40)$$

Unfortunately, there are the unknown terms  $\omega_1$ ,  $\omega_2$ ,  $\phi$ , and  $\psi$  in  $\Phi_\alpha(L)$ . The parameter estimates cannot be obtained.

Taking the first-order derivative of  $J_3(\omega)$  with respect to  $\omega$  gives

$$\text{grad}[J_3(\omega)] = \frac{\partial J_3(\omega)}{\partial \omega} = \left[ \frac{\partial J_3(\omega)}{\partial \omega_1}, \frac{\partial J_3(\omega)}{\partial \omega_2} \right]^T \in \mathbb{R}^2.$$

Taking the first-order derivative of  $J_4(\beta)$  with respect to  $\beta$  gives

$$\text{grad}[J_4(\beta)] = \frac{\partial J_4(\beta)}{\partial \beta} = \left[ \frac{\partial J_4(\beta)}{\partial \phi}, \frac{\partial J_4(\beta)}{\partial \psi} \right]^T \in \mathbb{R}^2.$$

Let  $\hat{\omega}_k := [\hat{\omega}_{1,k}, \hat{\omega}_{2,k}]^T$  be the estimate of  $\omega$  at iteration  $k$  and  $\hat{\beta}_k := [\hat{\phi}_k, \hat{\psi}_k]^T \in \mathbb{R}^2$  be the estimate of  $\beta$  at iteration  $k$ . The gradient-based iterative (GI) algorithm for estimating the parameter vector  $\omega$  is given as follows:

$$\begin{aligned} \hat{\omega}_k &= \hat{\omega}_{k-1} - \nu_k \text{grad}[J_3(\hat{\omega}_{k-1})], \\ \text{grad}[J_3(\hat{\omega}_{k-1})] &= \left[ \frac{\partial J_3(\hat{\omega}_{k-1})}{\partial \omega_1}, \frac{\partial J_3(\hat{\omega}_{k-1})}{\partial \omega_2} \right]^T \in \mathbb{R}^2. \end{aligned} \quad (41)$$

and the GI parameter estimation algorithm for estimating the parameter vector  $\beta$  is given as follows:

$$\hat{\beta}_k = \hat{\beta}_{k-1} - \lambda_k \text{grad}[J_4(\hat{\beta}_{k-1})],$$

$$\text{grad}[J_4(\hat{\beta}_{k-1})] = \left[ \frac{\partial J_4(\hat{\beta}_{k-1})}{\partial \phi}, \frac{\partial J_4(\hat{\beta}_{k-1})}{\partial \psi} \right]^T \in \mathbb{R}^2, \tag{42}$$

where two step sizes  $\nu_k$  and  $\lambda_k$  can be determined by using the one-dimensional search. Here, we replace the unknown parameter vectors  $\alpha$ ,  $\omega$ , and  $\beta$  in (40), (41), and (42) with their previous estimates and obtain the hierarchical least squares and gradient-based iterative (HI) algorithm for the dual-frequency signals:

$$\hat{\alpha}_k = \left[ \hat{\Phi}_{\alpha,k-1}^T(L) \hat{\Phi}_{\alpha,k-1}(L) \right]^{-1} \hat{\Phi}_{\alpha,k-1}(L) Y, \tag{43}$$

$$\hat{\omega}_k = \hat{\omega}_{k-1} - \nu_k \text{grad}[J_3(\hat{\omega}_{k-1})], \tag{44}$$

$$\hat{\beta}_k = \hat{\beta}_{k-1} - \lambda_k \text{grad}[J_4(\hat{\beta}_{k-1})], \tag{45}$$

$$Y(L) = [y(t_1), y(t_2), \dots, y(t_L)]^T, \tag{46}$$

$$\hat{\Phi}_{\alpha,k-1}(L) = [\hat{\phi}_{\alpha,k-1}(t_1), \hat{\phi}_{\alpha,k-1}(t_2), \dots, \hat{\phi}_{\alpha,k-1}(t_L)]^T, \tag{47}$$

$$\hat{\phi}_{\alpha,k-1}(t_i) = \left[ -\cos(\hat{\omega}_{1,k-1}t_i + \hat{\phi}_{k-1}), -\cos(\hat{\omega}_{2,k-1}t_i + \hat{\psi}_{k-1}) \right]^T, \tag{48}$$

$$\text{grad}[J_3(\hat{\omega}_{k-1})] = \left[ \frac{\partial J_3(\hat{\omega}_{k-1})}{\partial \omega_1}, \frac{\partial J_3(\hat{\omega}_{k-1})}{\partial \omega_2} \right]^T, \tag{49}$$

$$\nu_k = \arg \min J_3(\hat{\omega}_{k-1} - \nu \text{grad}[J_3(\hat{\omega}_{k-1})]), \tag{50}$$

$$\text{grad}[J_4(\hat{\beta}_{k-1})] = \left[ \frac{\partial J_4(\hat{\beta}_{k-1})}{\partial \phi}, \frac{\partial J_4(\hat{\beta}_{k-1})}{\partial \psi} \right]^T, \tag{51}$$

$$\lambda_k = \arg \min J_4(\hat{\beta}_{k-1} - \lambda \text{grad}[J_4(\hat{\beta}_{k-1})]), \tag{52}$$

$$\hat{\theta}_k = [\hat{\alpha}_k^T, \hat{\omega}_k^T, \hat{\beta}_k^T]^T. \tag{53}$$

The steps of the HI algorithm in (43)–(53) for dual-frequency signals are as follows.

1. To initialize, let  $k = 1$ ,  $\hat{\alpha}_0 = \mathbf{1}_2/p_0$ , let  $\hat{\omega}_0$  and  $\hat{\beta}_0$  be small real number vectors, give a small number  $\varepsilon$ .
2. Collect the observed data  $y(t_i)$  and  $u(t_i)$ ,  $i = 1, 2, \dots, L$ , and form  $Y(L)$  by (46).
3. Form the information vector  $\hat{\phi}_{\alpha,k-1}(t_i)$  using (48) and the stacked information matrix  $\hat{\Phi}_{\alpha,k-1}(L)$  using (47).
4. Compute  $\frac{\partial J_3(\hat{\omega}_{k-1})}{\partial \omega_i}$ ,  $i = 1, 2$ , using (49) and form  $\text{grad}[J_3(\hat{\omega}_{k-1})]$ .
5. Compute  $\frac{\partial J_4(\hat{\beta}_{k-1})}{\partial \phi}$ ,  $\frac{\partial J_4(\hat{\beta}_{k-1})}{\partial \psi}$  using (51) and form  $\text{grad}[J_4(\hat{\beta}_{k-1})]$ .
6. Select the step sizes  $\nu_k$  by (50) and  $\lambda_k$  by (52).

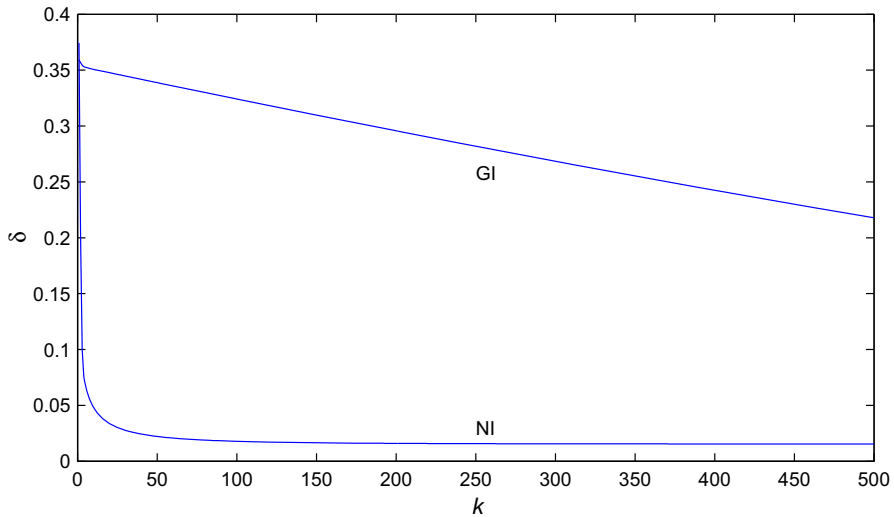


Fig. 1 NI and GI estimation errors  $\delta$  versus  $k$

7. Update the estimate  $\hat{\alpha}_k$  by (43), update the estimate  $\hat{\omega}_k$  by (44), and update the estimate  $\hat{\beta}_k$  using (45). Form  $\hat{\theta}_k$  using (53).
8. If  $\|\hat{\theta}_k - \hat{\theta}_{k-1}\| > \varepsilon$ , then increase  $k$  by 1, and go to Step 3; otherwise, obtain the estimate  $\hat{\theta}_k$  and terminate this procedure.

## 5 Examples

**Example 1** Consider the following dual-frequency cosine signal model:

$$y(t) = 1.48 \cos(0.2t + 2\pi/15) + 2.25 \cos(0.55t + \pi/4) + v(t).$$

The parameter vector to be estimated is

$$\theta = [a, b, \omega_1, \omega_2, \phi, \psi]^T = [1.48, 2.25, 0.20, 0.55, 2\pi/15, \pi/4]^T.$$

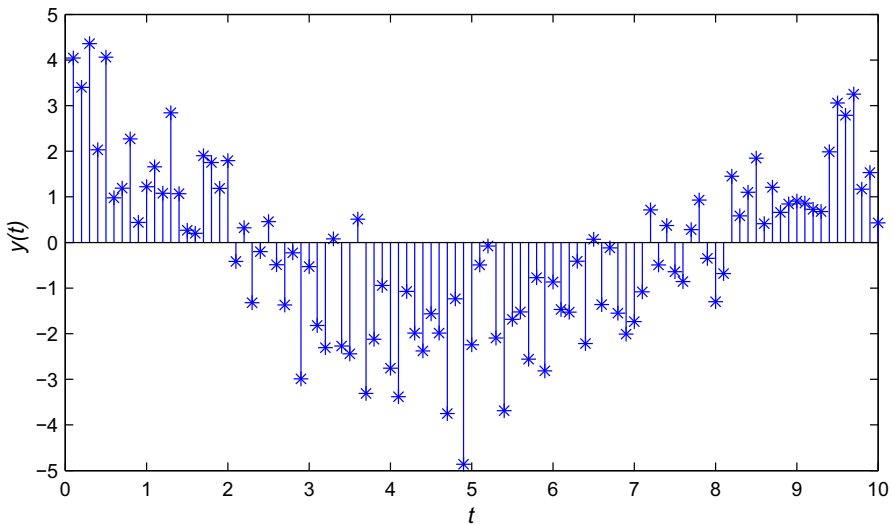
In simulation, the white noise signal  $v(t)$  with variance  $\sigma^2 = 1.00^2$  is added to the cosine combination signal. Set the sampling period  $T = 0.01$  s,  $t = nT$ ,  $n = 1, 2, 3, \dots$ , let the iterative time  $k = 500$ , the data length  $L = 2000$ . Apply the GI algorithm and the NI algorithm to estimate the parameter vector  $\theta$ . The parameter estimation errors  $\delta := \|\hat{\theta}_k - \theta\| / \|\theta\| \times 100\%$  versus  $k$  are illustrated in Fig. 1, and the parameter estimates and their estimation errors are exhibited in Table 1.

From the comparison of the GI algorithm and the NI algorithm, we can draw the following conclusions:

- The parameter estimation errors obtained by the GI algorithm and the NI algorithm gradually decrease as iteration  $k$  increases.

**Table 1** GI and the NI estimates and errors

Algorithms	$k$	$a$	$b$	$\omega_1$	$\omega_2$	$\phi$	$\psi$	$\delta$ (%)
GI	1	1.92278	3.14706	0.22793	0.72470	0.52167	-0.98430	35.91961
	2	1.92207	3.14427	0.19645	0.69311	0.51999	-0.98757	35.66197
	5	1.92055	3.14007	0.18202	0.58728	0.51977	-0.99580	35.25432
	100	1.88120	3.06929	0.21321	0.56784	0.51635	-0.98358	32.41767
	200	1.84200	2.99814	0.21216	0.56649	0.50975	-0.96985	29.57100
	500	1.73574	2.80346	0.20927	0.56275	0.48835	-0.93174	21.78905
NI	1	1.98755	2.27276	0.15553	0.61568	1.17722	-1.36504	37.43558
	2	1.04531	2.42861	0.23960	0.57526	0.13061	-0.95151	19.94760
	5	1.40530	2.33337	0.20962	0.57020	0.43107	-0.94600	6.81644
	100	1.46069	2.24935	0.20336	0.55343	0.39295	-0.82523	1.77920
	200	1.46554	2.24529	0.20326	0.55228	0.38788	-0.81589	1.59634
	500	1.46979	2.24272	0.20332	0.55143	0.38299	-0.80867	1.54416
True values		1.48000	2.25000	0.20000	0.55000	0.41888	0.78540	



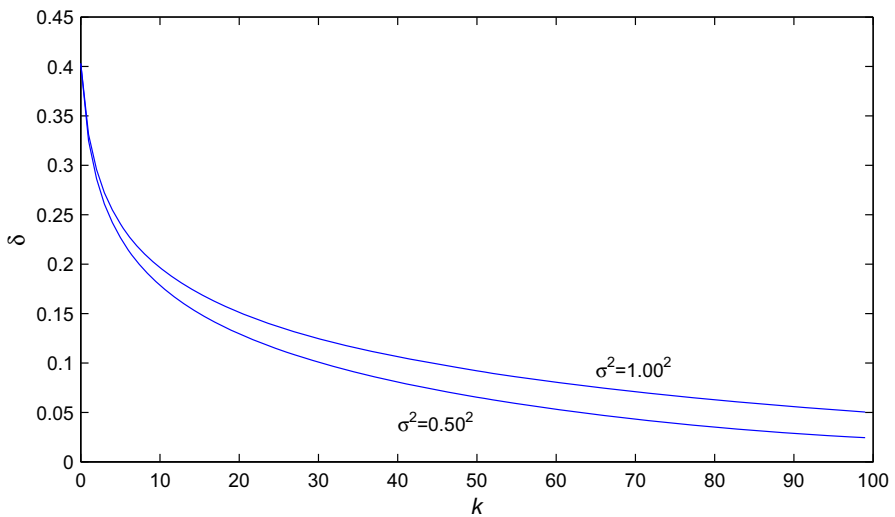
**Fig. 2** Measured data  $y(t_i)$  ( $\sigma^2 = 1.00^2$ )

- The NI parameter estimates approach their true values only in the dozens of iterations.
- Table 1 and Fig. 1 show that the NI algorithm has more accurate parameter estimates than those of the GI algorithm under the same condition.

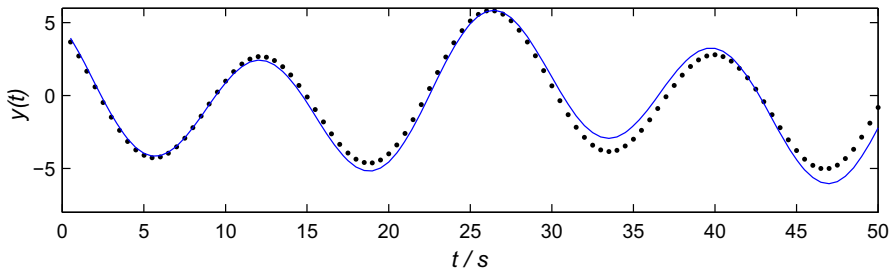
**Example 2** Consider the signal model in Example 1. In this simulation, the sampling period is 0.1s and the iterative time  $k = 100$ ;  $v(t)$  is the noise with zero mean and variances  $\sigma^2 = 0.50^2$  and  $\sigma^2 = 1.00^2$ , respectively. The distribution of the measured data is shown in Fig. 2.

**Table 2** HI estimates and errors

$\sigma^2$	$k$	$a$	$b$	$\omega_1$	$\omega_2$	$\phi$	$\psi$	$\delta$ (%)
1.00 <sup>2</sup>	1	1.71090	3.17709	0.45289	0.71831	0.22708	0.52635	33.11422
	2	1.67880	3.07949	0.44895	0.72404	0.22562	0.52828	29.56938
	5	1.62914	2.92850	0.44278	0.73301	0.22333	0.53130	24.09133
	10	1.58879	2.80580	0.43770	0.74040	0.22145	0.53379	19.64958
	20	1.54753	2.68037	0.43243	0.74806	0.21949	0.53636	15.12653
	50	1.49278	2.51392	0.42533	0.75839	0.21686	0.53984	9.19887
	100	1.45165	2.38886	0.41991	0.76627	0.21485	0.54249	4.97134
0.50 <sup>2</sup>	1	1.70814	3.15935	0.45286	0.71912	0.22712	0.52655	32.49730
	2	1.67468	3.05302	0.44891	0.72525	0.22568	0.52858	28.64766
	5	1.62290	2.88850	0.44272	0.73486	0.22342	0.53175	22.69714
	10	1.58083	2.75481	0.43761	0.74278	0.22156	0.53436	17.87076
	20	1.53782	2.61813	0.43232	0.75098	0.21963	0.53707	12.95545
	50	1.48075	2.43677	0.42520	0.76204	0.21703	0.54072	6.53725
	100	1.43787	2.30050	0.41975	0.77048	0.21504	0.54351	2.39653
True values		1.48000	2.25000	0.20000	0.55000	0.41888	0.78540	

**Fig. 3** HI estimation errors  $\delta$  versus  $k$ 

Apply the HI algorithm in (43)–(53) to estimate the parameter vectors  $\alpha$ ,  $\omega$ , and  $\beta$  of the dual-frequency signal model. The parameter estimates and errors  $\delta$  are shown in Table 2 and Fig. 3. In order to test the effectiveness of the estimated signal, the actual signal and its estimates are drawn in Fig. 4, where the solid line denotes the actual signal and the dot line denotes its estimates.



**Fig. 4** Actual signal and its estimates ( $\sigma^2 = 1.00^2$ )

From Fig. 3, we can see that the parameter estimation errors  $\delta$  obtained by the HI algorithm become small with  $k$  increasing. As the noise levels decrease, the HI algorithm can give more accurate parameter estimates; see the parameter estimation errors in the last column in Table 2 and the error curves in Fig. 3.

## 6 Conclusions

This paper presents a hierarchical least squares and gradient-based iterative algorithm to estimate all parameters of dual-frequency signal models from the observed data. The simulation results indicate that the proposed algorithm can improve the parameter estimation accuracy and reduce the computational burden. Similarly, the idea can be extended to estimate the parameters of multi-frequency signals and applied to signal detection and reconstruction, as well as in signal processing. The proposed algorithms in this paper can combine other methods such as the neural network methods [8,27] and the kernel methods [16,24] to study parameter identification of different systems [6,7,13–15,18,19,23,28,38,45,58] and can be applied to other fields [17,33,36,39,46,55–57].

**Acknowledgements** This work was supported by the National Natural Science Foundation of China (No. 61873111) and the 111 Project (B12018).

## References

1. B. Boashash, Estimating and interpreting the instantaneous frequency of a signal—part 1: fundamentals. *Proc. IEEE* **80**(4), 520–538 (1992)
2. Y. Cao, P. Li, Y. Zhang, Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing. *Future Gener. Comput. Syst.* **88**, 279–283 (2018)
3. Y. Cao, L.C. Ma, S. Xiao et al., Standard analysis for transfer delay in CTCS-3. *Chin. J. Electron.* **26**(5), 1057–1063 (2017)
4. Y. Cao, G. Wei, F.J. Chen, An exact analysis of modified covariance frequency estimation algorithm based on correlation of single-tone. *Signal Process.* **92**(11), 2785–2790 (2012)
5. Y. Cao, Y. Wen, X. Meng, W. Xu, Performance evaluation with improved receiver design for asynchronous coordinated multipoint transmissions. *Chin. J. Electron.* **25**(2), 372–378 (2016)

6. G.Y. Chen, M. Gan, C.L.P. Chen et al., A regularized variable projection algorithm for separable nonlinear least squares problems. *IEEE Trans. Autom. Control.* **64**(2), 1 (2019). <https://doi.org/10.1109/TAC.2018.2838045>
7. G.Y. Chen, M. Gan, F. Ding et al., Modified Gram–Schmidt method-based variable projection algorithm for separable nonlinear models. *IEEE Trans. Neural Netw. Learn. Syst.* **1**, 1 (2018). <https://doi.org/10.1109/TNNLS.2018.2884909>
8. M.Z. Chen, D.Q. Zhu, A workload balanced algorithm for task assignment and path planning of inhomogeneous autonomous underwater vehicle system. *IEEE Trans. Cogn. Dev. Syst.* **PP**, 1 (2018). <https://doi.org/10.1109/TCDS.2018.2866984>
9. J.L. Ding, Recursive and iterative least squares parameter estimation algorithms for multiple-input–output-error systems with autoregressive noise. *Circuits Syst. Signal Process.* **37**(5), 1884–1906 (2018)
10. J.L. Ding, The hierarchical iterative identification algorithm for multi-input–output-error systems with autoregressive noise. *Complexity* (2017) <https://doi.org/10.1155/2017/5292894> (Article ID 5292894)
11. F. Ding, H.B. Chen, L. Xu et al., A hierarchical least squares identification algorithm for Hammerstein nonlinear systems using the key term separation. *J. Frankl. Inst.* **355**(8), 3737–3752 (2018)
12. F. Ding, L. Xu, F.E. Alsaadi, T. Hayat, Iterative parameter identification for pseudo-linear systems with ARMA noise using the filtering technique. *IET Control Theory Appl.* **12**(7), 892–899 (2018)
13. M. Gan, C.L.P. Chen, G.Y. Chen et al., On some separated algorithms for separable nonlinear squares problems. *IEEE Trans. Cybern.* **48**(10), 2866–2874 (2018)
14. M. Gan, H.X. Li, H. Peng, A variable projection approach for efficient estimation of RBF-ARX model. *IEEE Trans. Cybern.* **45**(3), 462–471 (2015)
15. Z.W. Ge, F. Ding, L. Xu, A. Alsaedi, T. Hayat, Gradient-based iterative identification method for multivariate equation-error autoregressive moving average systems using the decomposition technique. *J. Frankl. Inst.* **356**, 1–19 (2019). <https://doi.org/10.1016/j.jfranklin.2018.12.002>
16. F.Z. Geng, S.P. Qian, An optimal reproducing kernel method for linear nonlocal boundary value problems. *Appl. Math. Lett.* **77**, 49–56 (2018)
17. P.C. Gong, W.Q. Wang, F.C. Li et al., Sparsity-aware transmit beamspace design for FDA-MIMO radar. *Signal Process.* **144**, 99–103 (2018)
18. Y. Gu, F. Ding, J.H. Li, States based iterative parameter estimation for a state space model with multi-state delays using decomposition. *Signal Process.* **106**, 294–300 (2015)
19. Y. Gu, J. Liu, X. Li, State space model identification of multirate processes with time-delay using the expectation maximization. *J. Frankl. Inst.* **356**, 1–17 (2019). <https://doi.org/10.1016/j.jfranklin.2018.08.030>
20. M. Hou, Amplitude and frequency estimator of a sinusoid. *IEEE Trans. Autom. Control* **50**(6), 855–858 (2005)
21. M. Hou, Estimation of sinusoidal frequencies and amplitudes using adaptive identifier and observer. *IEEE Trans. Autom. Control* **52**(3), 493–499 (2007)
22. M. Hou, Parameter identification of sinusoids. *IEEE Trans. Autom. Control* **57**(2), 467–472 (2012)
23. Y.B. Hu, Q. Zhou, H. Yu, Z. Zhou, F. Ding, Two-stage generalized projection identification algorithms for controlled autoregressive systems. *Circuits Syst. Signal Process.* **38**, 1–17 (2019). <https://doi.org/10.1007/s00034-018-0996-0>
24. X.Y. Li, B.Y. Wu, A new reproducing kernel collocation method for nonlocal fractional boundary value problems with non-smooth solutions. *Appl. Math. Lett.* **86**, 194–199 (2018)
25. F.B. Li, L.G. Wu, P. Shi, C.C. Lim, State estimation and sliding mode control for semi-Markovian jump systems with mismatched uncertainties. *Automatica* **51**, 385–393 (2015)
26. J.H. Li, W. Zheng, J.P. Gu, L. Hua, A recursive identification algorithm for Wiener nonlinear systems with linear state-space subsystem. *Circuits Syst. Signal Process.* **37**(6), 2374–2393 (2018)
27. X. Li, D.Q. Zhu, An adaptive SOM neural network method for distributed formation control of a group of AUVs. *IEEE Trans. Ind. Electron.* **65**(10), 8260–8270 (2018)
28. Q.Y. Liu, F. Ding, Auxiliary model-based recursive generalized least squares algorithm for multivariate output-error autoregressive systems using the data filtering. *Circuits Syst. Signal Process.* **38**, 1–21 (2019). <https://doi.org/10.1007/s00034-018-0871-z>
29. S.Y. Liu, L. Xu, F. Ding, Iterative parameter estimation algorithms for dual-frequency signal models. *Algorithms* **10**(4), 1–13 (2017). <https://doi.org/10.3390/a10040118>
30. P.P. Liu, S.G. Zhang, Newton’s method for solving a class of nonlinear matrix equations. *J. Comput. Appl. Math.* **255**, 254–267 (2014)



31. P. Ma, F. Ding, Q.M. Zhu, Decomposition-based recursive least squares identification methods for multivariate pseudolinear systems using the multi-innovation. *Int. J. Syst. Sci.* **49**(5), 920–928 (2018)
32. M. Nafar, G.B. Gharehpetian, T. Niknam, Using modified fuzzy particle swarm optimization algorithm for parameter estimation of surge arresters model. *Int. J. Innov. Comput. Inf. Control* **8**(1B), 567–581 (2012)
33. J. Pan, W. Li, H.P. Zhang, Control algorithms of magnetic suspension systems based on the improved double exponential reaching law of sliding mode control. *Int. J. Control Autom. Syst.* **16**(6), 2878–2887 (2018)
34. J. Pan, H. Ma, X. Jiang, et al., Adaptive gradient-based iterative algorithm for multivariate controlled autoregressive moving average systems using the data filtering technique. *Complexity* (2018). <https://doi.org/10.1155/2018/9598307> (Article ID 9598307)
35. H. Ramos, M.T.T. Monteiro, A new approach based on the Newton's method to solve systems of nonlinear equations. *J. Comput. Appl. Math.* **318**, 3–13 (2017)
36. Z.H. Rao, C.Y. Zeng, M.H. Wu et al., Research on a handwritten character recognition algorithm based on an extended nonlinear kernel residual network. *KSII Trans. Internet Inf. Syst.* **12**(1), 413–435 (2018)
37. P. Stoica, H.B. Li, J. Li, Amplitude estimation of sinusoidal signals: survey, new results, and an application. *IEEE Trans. Signal Process.* **48**(2), 338–352 (2000)
38. L.J. Wan, F. Ding, Decomposition-based gradient iterative identification algorithms for multivariable systems using the multi-innovation theory. *Circuits Syst. Signal Process.* **38**, 1–20 (2019). <https://doi.org/10.1007/s00034-018-1014-2>
39. Y. Wang, Y. Si, B. Huang, S.X. Ding, Survey on the theoretical research and engineering applications of multivariate statistics process monitoring algorithms: 2008–2017. *Can. J. Chem. Eng.* **96**(10), 2073–2085 (2018)
40. L. Xu, Application of the Newton iteration algorithm to the parameter estimation for dynamical systems. *J. Comput. Appl. Math.* **288**, 33–43 (2015)
41. L. Xu, The parameter estimation algorithms based on the dynamical response measurement data. *Adv. Mech. Eng.* **9**(11), 1–12 (2017). <https://doi.org/10.1177/1687814017730003>
42. L. Xu, F. Ding, Iterative parameter estimation for signal models based on measured data. *Circuits Syst. Signal Process.* **37**(7), 3046–3069 (2018)
43. L. Xu, F. Ding, Parameter estimation for control systems based on impulse responses. *Int. J. Control Autom. Syst.* **15**(6), 2471–2479 (2017)
44. L. Xu, F. Ding, Y. Gu et al., A multi-innovation state and parameter estimation algorithm for a state space system with d-step state-delay. *Signal Process.* **140**, 97–103 (2017)
45. H. Xu, F. Ding, E.F. Yang, Modeling a nonlinear process using the exponential autoregressive time series model. *Nonlinear Dyn.* (2019). <https://doi.org/10.1007/s11071-018-4677-0>
46. G.H. Xu, Y. Shekofteh, A. Akgul et al., A new chaotic system with a self-excited attractor: entropy measurement, signal encryption, and parameter estimation. *Entropy* **20**(2), 1–23 (2018). <https://doi.org/10.3390/e20020086>
47. L. Xu, W.L. Xiong et al., Hierarchical parameter estimation for the frequency response based on the dynamical window data. *Int. J. Control Autom. Syst.* **16**(4), 1756–1764 (2018)
48. C. Yang, G. Wei, A noniterative frequency estimator with rational combination of three spectrum lines. *IEEE Trans. Signal Process.* **59**(10), 5065–5070 (2011)
49. C. Yin, S.S. Wu et al., Design and stability analysis of multivariate extremum seeking with Newton method. *J. Frankl. Inst.* **355**(4), 1559–1578 (2018)
50. Q.G. Zhang, The Rife frequency estimation algorithm based on real-time FFT. *Signal Process.* **25**(6), 1002–1004 (2009)
51. Y.Z. Zhang, Y. Cao, Y.H. Wen, L. Liang, F. Zou, Optimization of information interaction protocols in cooperative vehicle-infrastructure systems. *Chin. J. Electron.* **27**(2), 439–444 (2018)
52. X. Zhang, F. Ding, F.E. Alsaadi, T. Hayat, Recursive parameter identification of the dynamical models for bilinear state space systems. *Nonlinear Dyn.* **89**(4), 2415–2429 (2017)
53. X. Zhang, F. Ding, L. Xu et al., State filtering-based least squares parameter estimation for bilinear systems using the hierarchical identification principle. *IET Control Theory Appl.* **12**(12), 1704–1713 (2018)
54. X. Zhang, L. Xu, F. Ding et al., Combined state and parameter estimation for a bilinear state space system with moving average noise. *J. Frankl. Inst.* **355**(6), 3079–3103 (2018)
55. W.H. Zhang, L. Xue, X. Jiang, Global stabilization for a class of stochastic nonlinear systems with SISS-like conditions and time delay. *Int. J. Robust Nonlinear Control* **28**(13), 3909–3926 (2018)

56. N. Zhao, R. Liu, Y. Chen et al., Contract design for relay incentive mechanism under dual asymmetric information in cooperative networks. *Wirel. Netw.* **24**(8), 3029–3044 (2018)
57. D. Zhao, Y. Wang, Y. Li, S.X. Ding, H-infinity fault estimation for two-dimensional linear discrete time-varying systems based on Krein space method. *IEEE Trans. Syst. Man Cybern. Syst.* **48**(12), 2070–2079 (2018)
58. Z.P. Zhou, X.F. Liu, State and fault estimation of sandwich systems with hysteresis. *Int. J. Robust Nonlinear Control* **28**(13), 3974–3986 (2018)

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