

Article

# Iterative Parameter Estimation Algorithms for Dual-Frequency Signal Models

Siyu Liu <sup>1</sup>, Ling Xu <sup>2</sup> and Feng Ding <sup>1,\*</sup>

<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, China; syliu12@126.com

<sup>2</sup> School of Internet of Things Technology, Wuxi Vocational Institute of Commerce, Wuxi 214153, China; lingxu0848@163.com

\* Correspondence: fding@jiangnan.edu.cn

Received: 15 August 2017; Accepted: 11 October 2017; Published: 14 October 2017

**Abstract:** This paper focuses on the iterative parameter estimation algorithms for dual-frequency signal models that are disturbed by stochastic noise. The key of the work is to overcome the difficulty that the signal model is a highly nonlinear function with respect to frequencies. A gradient-based iterative (GI) algorithm is presented based on the gradient search. In order to improve the estimation accuracy of the GI algorithm, a Newton iterative algorithm and a moving data window gradient-based iterative algorithm are proposed based on the moving data window technique. Comparative simulation results are provided to illustrate the effectiveness of the proposed approaches for estimating the parameters of signal models.

**Keywords:** signal processing; parameter estimation; moving data window; gradient search; Newton search

## 1. Introduction

Parameter estimation is used widely in system identification [1–3] and signal processing [4,5]. The existing parameter estimation methods for signal models can be classified into the following basic categories: the frequency-domain methods and the time-domain methods. The frequency-domain methods based on the fast Fourier transform (FFT) mainly include the Rife method, the phase difference method, etc. The accuracy of the Rife method is high in the case of noiseless or higher signal-to-noise ratio with adaptive sampling points; however, the error given by the Rife method is large if the signal frequency is near the DFT quantization frequency point. Therefore, some improved methods were proposed [6,7]. For example, Jacobsen and Kootsookos used three largest spectral lines of the FFT spectrum to calibrate the frequency estimation [8]. Deng et al. proposed a modified Rife method for the frequency shift of signals [9]. However, there is a risk that the shift direction of the frequency may be error. The time-domain methods mainly include the maximum likelihood algorithm, the subspace method and the self-correlation phase method. The maximum likelihood method is effective estimation to minimize the average risk though imposing significant computational costs. The methods in [10–12] used multiple autocorrelation coefficients or multi-steps autocorrelation functions to estimate the frequency so the amount of computations increase.

In practice, actual signals are usually disturbed by various stochastic noise, and the time series signals like vibration signals or biomedical signals are subjected to dynamic excitations, including nonlinear and non-stationary properties. In order to solve the difficulties, time-frequency representations (TFR) provide a powerful tool since a TFR can give information about the frequencies contained in signal over time, such as short-time Fourier transform, wavelet transform, and Hilbert–Huang transform [13]. Recently, Amezcua-Sanchez and Adeli presented an adaptive multiple signal classification-empirical wavelet transform methodology for accurate time-frequency

representation of noisy non-stationary and nonlinear signals [14]. Daubechies et al. used an empirical mode decomposition-like tool to decompose into the blocks functions, with slowly varying amplitudes and frequencies [15]. These existing approaches can obtain signal models indirectly, because they are realized based on the transform techniques like the Fourier transform and wavelet transform. In this paper, we propose the direct parameter estimation algorithms for signal modeling.

The iterative methods and/or the recursive 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 [16–22]. Furthermore, the iterative algorithms can give more accurate parameter estimates because of making full use of the observed data. Yun utilized the iterative methods to find the roots of nonlinear equations [23]. Dehghan and Hajarian presented an iterative algorithm for solving the generalized coupled Sylvester matrix equations over the generalized centro-symmetric matrices [24]. Wang et al. proposed an iterative method for a class of complex symmetric linear systems [25]. Xu derived a Newton iterative algorithm to the parameter estimation for dynamical systems [26]. Pei et al. used a monotone iterative technique to get the existence of positive solutions and to seek the positive minimal and maximal solutions for a Hadamard type fractional integro-differential equation [27].

As an optimization tool, the Newton method is useful for solving roots of nonlinear problems or deriving parameter estimation algorithms from observed data [28–30]. For a long time, the Newton method has been utilized in much literature, such as transcendental equations, minimization and maximization problems, and numerical verification for solutions of nonlinear equations. Simpson noted that the Newton method can be used to give the generalization to systems of two equations and to solve optimization problems [31]. Dennis provided Newton's method or quasi-Newton methods for multidimensional unconstrained optimization and nonlinear equation problems [32]. Jürgen studied the accelerated convergence of the Newton method by molding a given function into a new one that the roots remain unchanged, but it looks nearly linear in a neighborhood of the root [33]. Djoudi et al. presented a guided recursive Newton method involving inverse iteration to solve the transcendental eigenproblems by reducing it to a generalised linear eigenproblem [34]. Benner described the numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems [35]. Seinfeld et al. used a quasi-Newton search technique and a barrier modification to enforce closed-loop stability for the H-infinity control problem [36]. Liu et al. presented an iterative identification algorithm for Wiener nonlinear systems using the Newton method [37]. The gradient method with the search directions defined by the gradient of the function at the current point has been developed for optimization problems. For instance, Curry used the gradient descent method for minimizing a nonlinear function of  $n$  real variables [38]. Vranhatis et al. studied the development, convergence theory and numerical testing of a class of gradient unconstrained minimization algorithms with adaptive step-size [39]. Hajarian proposed a gradient-based iterative algorithm to find the solutions of the general Sylvester discrete-time periodic matrix equations.

The moving data window with a fixed length is moved as time, which is a first-in-first-out sequence. When a new observation arrives, the data in this moving window are updated by including the new observation and eliminating the oldest one. The length of the moving window remains fixed. The algorithm computes parameter estimates using the observed data in the current window. Recently, Wang et al. presented a moving-window second order blind identification for time-varying transient operational modal parameter identification of linear time-varying structures [41]. Al-Matouq and Vincent developed a multiple-window moving horizon estimation strategy that exploits constraint inactivity to reduce the problem size in long horizon estimation problems [42]. This paper focuses on the parameter estimation problems of dual-frequency signal models. The main contributions of this paper are twofold. The basic idea is to present a gradient-based iterative (GI) algorithm and to estimate the parameters for signal models. Several estimation errors obtained by the Newton iterative and the moving data window based GI algorithms are compared to the errors given by the GI algorithm.

To close this section, we give the outline of this paper. Section 2 derives a GI parameter estimation algorithm. Sections 3 and 4 propose the Newton and moving data window gradient based iterative parameter estimation algorithms. Section 5 provides an example to verify the effectiveness of the proposed algorithms. Finally, some concluding remarks are given in Section 6.

## 2. The Gradient-Based Iterative Parameter Estimation Algorithm

Consider the following dual-frequency cosine signal model:

$$y(t) = a \cos(\omega_1 t) + b \cos(\omega_2 t) + v(t), \tag{1}$$

where  $\omega_1 > 0$  and  $\omega_2 > 0$  are the angular frequencies,  $a > 0$  and  $b > 0$  are the amplitudes,  $t$  is a continuous-time variable,  $y(t)$  is the observation, and  $v(t)$  is a stochastic disturbance with zero mean and variance  $\sigma^2$ . In actual engineering, we can only get discrete observed data. Suppose that the sampling data are  $y(t_i), i = 1, 2, 3, \dots, L$ , where  $L$  is the data length and  $t_i$  is the sampling time.

As we all know, signals include the sine signal, cosine signal, Gaussian signal, exponential signal, complex exponential signal, etc. Among them, the sine signal and the cosine signal are typical periodic signals whose waveforms are sine and cosine curves in mathematics. Many periodic signals can be decomposed into the sun of multiple sinusoidal signals with different frequencies and different amplitudes by the Fourier series [43]. The cosine signal differs from the sine signal by  $\pi/2$  in the initial phase. This paper takes the double-frequency cosine signal model as an example and derives the parameter estimation algorithms, which are also applicable to the sinusoidal signal models.

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

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

The criterion function contains the parameters to be estimated, and it represents the error between the observed data and the model output. We hope that this error is as small as possible, which is equivalent to minimizing  $J_1(\theta)$  and obtaining the estimates of the parameter vector  $\theta := [a, b, \omega_1, \omega_2]^T \in \mathbb{R}^4$ .

Letting the partial derivative of  $J_1(\theta)$  with respect to  $\theta$  be zero gives

$$\text{grad}[J_1(\theta)] := \frac{\partial J_1(\theta)}{\partial \theta} = \left[ \frac{\partial J_1(\theta)}{\partial a}, \frac{\partial J_1(\theta)}{\partial b}, \frac{\partial J_1(\theta)}{\partial \omega_1}, \frac{\partial J_1(\theta)}{\partial \omega_2} \right]^T \in \mathbb{R}^4. \tag{3}$$

Define the stacked output vector  $Y(L)$  and the stacked information matrix  $\Phi(\theta, L)$  as

$$Y(L) := \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(L) \end{bmatrix} \in \mathbb{R}^L, \quad \Phi(\theta, L) := \begin{bmatrix} \varphi^T(\theta, 1) \\ \varphi^T(\theta, 2) \\ \vdots \\ \varphi^T(\theta, L) \end{bmatrix} \in \mathbb{R}^{L \times 4}. \tag{4}$$

Define the information vector

$$\varphi(\theta, t_i) := [-\cos(\omega_1 t_i), -\cos(\omega_2 t_i), at_i \sin(\omega_1 t_i), bt_i \sin(\omega_2 t_i)]^T \in \mathbb{R}^4. \tag{5}$$

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

$$F(\theta, L) := \begin{bmatrix} f(\theta, 1) \\ f(\theta, 2) \\ \vdots \\ f(\theta, L) \end{bmatrix} \in \mathbb{R}^L. \tag{6}$$

Then, the gradient vector can be expressed as

$$\text{grad}[J_1(\theta)] := \Phi^T(\theta, L)[Y(L) - F(\theta, L)]. \tag{7}$$

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

$$\begin{aligned} \text{grad}[J_1(\hat{\theta}_{k-1})] &= \left[ \frac{\partial J_1(\hat{\theta}_{k-1})}{\partial a}, \frac{\partial J_1(\hat{\theta}_{k-1})}{\partial b}, \frac{\partial J_1(\hat{\theta}_{k-1})}{\partial \omega_1}, \frac{\partial J_1(\hat{\theta}_{k-1})}{\partial \omega_2} \right]^T \in \mathbb{R}^4, \\ \frac{\partial J_1(\hat{\theta}_{k-1})}{\partial a} &= - \sum_{i=1}^L \cos(\hat{\omega}_{1,k-1} t_i) [y(t_i) - \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1} t_i) - \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1} t_i)], \\ \frac{\partial J_1(\hat{\theta}_{k-1})}{\partial b} &= - \sum_{i=1}^L \cos(\hat{\omega}_{2,k-1} t_i) [y(t_i) - \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1} t_i) - \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1} t_i)], \\ \frac{\partial J_1(\hat{\theta}_{k-1})}{\partial \hat{\omega}_1} &= \sum_{i=1}^L \hat{a}_{k-1} t_i \sin(\hat{\omega}_{1,k-1} t_i) [y(t_i) - \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1} t_i) - \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1} t_i)], \\ \frac{\partial J_1(\hat{\theta}_{k-1})}{\partial \hat{\omega}_2} &= \sum_{i=1}^L \hat{b}_{k-1} t_i \sin(\hat{\omega}_{2,k-1} t_i) [y(t_i) - \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1} t_i) - \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1} t_i)]. \end{aligned} \tag{8}$$

Using the negative gradient search and minimizing  $J_1(\theta)$ , introducing an iterative step-size  $\mu_k$ , we can get the gradient-based iterative (GI) algorithm for dual-frequency signal models:

$$\hat{\theta}_k = \hat{\theta}_{k-1} - \mu_k \text{grad}[J_1(\hat{\theta}_{k-1})] \tag{9}$$

$$= \hat{\theta}_{k-1} - \mu_k \Phi^T(\hat{\theta}_{k-1}, L)[Y(L) - F(\hat{\theta}_{k-1}, L)],$$

$$Y(L) = [y(1), y(2), \dots, y(L)]^T, \tag{10}$$

$$\hat{\Phi}_k := \Phi(\hat{\theta}_{k-1}, L) = [\varphi(\hat{\theta}_{k-1}, 1), \varphi(\hat{\theta}_{k-1}, 2), \dots, \varphi(\hat{\theta}_{k-1}, L)]^T, \tag{11}$$

$$\hat{F}_k := F(\hat{\theta}_{k-1}, L) = [f(\hat{\theta}_{k-1}, 1), f(\hat{\theta}_{k-1}, 2), \dots, f(\hat{\theta}_{k-1}, L)]^T, \tag{12}$$

$$\begin{aligned} \hat{\varphi}_k &:= \varphi(\hat{\theta}_{k-1}, t_i) \\ &= [-\cos(\hat{\omega}_{1,k-1} t_i), -\cos(\hat{\omega}_{2,k-1} t_i), \hat{a}_{k-1} t_i \sin(\hat{\omega}_{1,k-1} t_i), \hat{b}_{k-1} t_i \sin(\hat{\omega}_{2,k-1} t_i)]^T, \end{aligned} \tag{13}$$

$$\begin{aligned} \hat{f}_k &:= f(\hat{\theta}_{k-1}, t_i) \\ &= \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1} t_i) + \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1} t_i), \end{aligned} \tag{14}$$

$$\hat{\theta}_k = [\hat{a}_k, \hat{b}_k, \hat{\omega}_{1,k}, \hat{\omega}_{2,k}]^T, \tag{15}$$

$$0 < \mu_k \leq \lambda_{\max}^{-1}[\hat{\Phi}_k^T \hat{\Phi}_k]. \tag{16}$$

The steps of the GI algorithm to compute  $\hat{\theta}_k$  are listed as follows:

1. Let  $k = 1$ , give a small number  $\varepsilon$  and set the initial value  $\hat{\theta}_0 = \mathbf{1}_n / p_0$ ,  $p_0$  is generally taken to be a large positive number, e.g.,  $p_0 = 10^6$ .
2. Collect the observed data  $y(t_i)$ ,  $i = 1, 2, \dots, L$ , where  $L$  is the data length, form  $Y(L)$  by Equation (10).
3. Form  $\hat{\varphi}_k$  by Equation (13) and  $\hat{\Phi}_k$  by Equation (11).
4. Form  $\hat{f}_k$  by Equation (14) and  $\hat{F}_k$  by Equation (12).

5. Choose a larger  $\mu_k$  satisfying Equation (16), and update the estimate  $\hat{\theta}_k$  by Equation (15).
6. Compare  $\hat{\theta}_k$  with  $\hat{\theta}_{k-1}$ , if  $\|\hat{\theta}_k - \hat{\theta}_{k-1}\| \leq \varepsilon$ , then terminate this procedure and obtain the iteration time  $k$  and the estimate  $\hat{\theta}_k$ ; otherwise, increase  $k$  by 1 and go to Step 3.

### 3. The Newton Iterative Parameter Estimation Algorithm

The gradient method uses the first derivative, and its convergence rate is slow. The Newton method is discussed here, which requires the second derivative, with fast convergence speed. The following derives the Newton iterative algorithm to estimate the parameters of dual-frequency signal models.

According to the criterion function  $J_1(\theta)$  in Equation (2). Calculating the second partial derivative of the criterion function  $J_1(\theta)$  with respect to the parameter  $\theta$  gives the Hessian matrix

$$H(\theta) := \frac{\partial^2 J_1(\theta)}{\partial \theta \partial \theta^T} = \frac{\partial \text{grad}[J_1(\theta)]}{\partial \theta^T}. \tag{17}$$

Let  $k = 1, 2, 3, \dots$  be an iterative variable and  $\hat{\theta}_{k-1}$  be the estimate of  $\theta$  at iteration  $k - 1$ . Based on the Newton search, we can derive the Newton iterative (NI) parameter estimation algorithm of dual-frequency signal models:

$$\hat{\theta}_k = \hat{\theta}_{k-1} - H^{-1}(\hat{\theta}_{k-1}) \text{grad}[J_1(\hat{\theta}_{k-1})], \tag{18}$$

$$Y(L) = [y(1), y(2), \dots, y(L)]^T, \tag{19}$$

$$\hat{\Phi}_k := \Phi(\hat{\theta}_{k-1}, L) = [\varphi(\hat{\theta}_{k-1}, 1), \varphi(\hat{\theta}_{k-1}, 2), \dots, \varphi(\hat{\theta}_{k-1}, L)]^T, \tag{20}$$

$$\hat{F}_k := F(\hat{\theta}_{k-1}, L) = [f(\hat{\theta}_{k-1}, 1), f(\hat{\theta}_{k-1}, 2), \dots, f(\hat{\theta}_{k-1}, L)]^T, \tag{21}$$

$$\hat{\varphi}_k := \varphi(\hat{\theta}_{k-1}, t_i) = [-\cos(\hat{\omega}_{1,k-1} t_i), -\cos(\hat{\omega}_{2,k-1} t_i), \hat{a}_{k-1} t_i \sin(\hat{\omega}_{1,k-1} t_i), \hat{b}_{k-1} t_i \sin(\hat{\omega}_{2,k-1} t_i)]^T, \tag{22}$$

$$\hat{f}_k := f(\hat{\theta}_{k-1}, t_i) = \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1} t_i) + \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1} t_i), \tag{23}$$

$$H(\hat{\theta}_{k-1}) = \begin{bmatrix} h_{11}(\hat{\theta}_{k-1}) & h_{12}(\hat{\theta}_{k-1}) & h_{13}(\hat{\theta}_{k-1}) & h_{14}(\hat{\theta}_{k-1}) \\ h_{21}(\hat{\theta}_{k-1}) & h_{22}(\hat{\theta}_{k-1}) & h_{23}(\hat{\theta}_{k-1}) & h_{24}(\hat{\theta}_{k-1}) \\ h_{31}(\hat{\theta}_{k-1}) & h_{32}(\hat{\theta}_{k-1}) & h_{33}(\hat{\theta}_{k-1}) & h_{34}(\hat{\theta}_{k-1}) \\ h_{41}(\hat{\theta}_{k-1}) & h_{42}(\hat{\theta}_{k-1}) & h_{43}(\hat{\theta}_{k-1}) & h_{44}(\hat{\theta}_{k-1}) \end{bmatrix}, \tag{24}$$

$$h_{11}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \cos^2(\hat{\omega}_{1,k-1} t_i), \tag{25}$$

$$h_{12}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \cos(\hat{\omega}_{1,k-1} t_i) \cos(\hat{\omega}_{2,k-1} t_i), \tag{26}$$

$$h_{13}(\hat{\theta}_{k-1}) = \sum_{i=1}^L t_i \sin(\hat{\omega}_{1,k-1} t_i) [y(t_i) - \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1} t_i)] - \hat{a}_{k-1} t_i \sin(2\hat{\omega}_{1,k-1} t_i), \tag{27}$$

$$h_{14}(\hat{\theta}_{k-1}) = -\sum_{i=1}^L \hat{b}_{k-1} t_i \cos(\hat{\omega}_{1,k-1} t_i) \sin(\hat{\omega}_{2,k-1} t_i), \tag{28}$$

$$h_{22}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \cos^2(\hat{\omega}_{2,k-1} t_i), \tag{29}$$

$$h_{23}(\hat{\theta}_{k-1}) = -\sum_{i=1}^L \hat{a}_{k-1} t_i \cos(\hat{\omega}_{2,k-1} t_i) \sin(\hat{\omega}_{1,k-1} t_i), \tag{30}$$

$$h_{24}(\hat{\theta}_{k-1}) = \sum_{i=1}^L t_i \sin(\hat{\omega}_{2,k-1} t_i) [y(t_i) - \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1} t_i)] - \hat{b}_{k-1} t_i \sin(2\hat{\omega}_{2,k-1} t_i), \tag{31}$$

$$h_{33}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \hat{a}_{k-1}^2 t_i^2 \cos(\hat{\omega}_{1,k-1} t_i) [y(t_i) - \hat{b}_{k-1} \cos(\hat{\omega}_{2,k-1} t_i)] - \hat{a}_{k-1}^2 t_i^2 \cos(2\hat{\omega}_{1,k-1} t_i), \tag{32}$$

$$h_{34}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \hat{a}_{k-1} \hat{b}_{k-1} t_i^2 \sin(\hat{\omega}_{1,k-1} t_i) \sin(\hat{\omega}_{2,k-1} t_i), \tag{33}$$

$$h_{44}(\hat{\theta}_{k-1}) = \sum_{i=1}^L \hat{b}_{k-1} t_i^2 \cos(\hat{\omega}_{2,k-1} t_i) [y(t) - \hat{a}_{k-1} \cos(\hat{\omega}_{1,k-1} t_i)] - \hat{b}_{k-1}^2 t_i^2 \cos(2\hat{\omega}_{2,k-1} t_i). \tag{34}$$

$$\hat{\theta}_k = [\hat{a}_k, \hat{b}_k, \hat{\omega}_{1,k}, \hat{\omega}_{2,k}]^T. \tag{35}$$

The procedure for computing the parameter estimation vector  $\hat{\theta}_k$  using the NI algorithm in Equations (18)–(35) is listed as follows:

1. Let  $k = 1$ , give the parameter estimation accuracy  $\varepsilon$ , set the initial value  $\hat{\theta}_0 = \mathbf{1}_n/p_0$ .
2. Collect the observed data  $y(t_i), i = 1, 2, \dots, L$ , form  $Y(L)$  by Equation (19).
3. Form  $\hat{\phi}_k$  by Equation (22), form  $\hat{\Phi}_k$  by Equation (20).
4. Form  $\hat{f}_k$  by Equation (23), form  $\hat{F}_k$  by Equation (21).
5. Compute  $h_{mn}(\hat{\theta}_{k-1})$  by Equations (25)–(34),  $m = 1, 2, 3, 4, n = 1, 2, 3, 4$ , and form  $H(\hat{\theta}_{k-1})$  by Equation (24).
6. Update the estimate  $\hat{\theta}_k$  by Equation (35).
7. Compare  $\hat{\theta}_k$  with  $\hat{\theta}_{k-1}$ , if  $\|\hat{\theta}_k - \hat{\theta}_{k-1}\| \leq \varepsilon$ , then terminate this procedure and obtain the estimate  $\hat{\theta}_k$ ; otherwise, increase  $k$  by 1 and go to Step 3.

#### 4. The Moving Data Window Gradient-Based Iterative Algorithm

The gradient method and the Newton method use a batch of data, which is the data from  $t = 1$  to  $t = L$ . Here, the moving window data from  $y(t_{i-p+1})$  to  $y(t_i)$ ,  $p$  used here is the length of the moving data window. Let  $t_i$  be the current sampling time. The moving window data can be represented as  $y(t_i), y(t_{i-1}), \dots, y(t_{i-p+1})$ . These sampling data change with the sampling time  $t_i$ . With the increasing of  $i$ , the data window moves forward constantly. New data are collected and old data are removed from the window. The following derives the moving data window gradient-based iterative algorithm of dual-frequency signal models.

Define the moving data window criterion function

$$J_2(\theta) := \frac{1}{2} \sum_{j=0}^{p-1} [y(t_{i-j}) - a \cos(\omega_1 t_{i-j}) - b \cos(\omega_2 t_{i-j})]^2. \tag{36}$$

Define the information vector

$$\varphi(\theta, t_{i-j}) := [-\cos(\omega_1 t_{i-j}), -\cos(\omega_2 t_{i-j}), a t_{i-j} \sin(\omega_1 t_{i-j}), b t_{i-j} \sin(\omega_2 t_{i-j})]^T \in \mathbb{R}^4, \tag{37}$$

and the stacked information matrix

$$\Phi(p, \theta, t_i) := [\varphi^T(\theta, t_i), \varphi^T(\theta, t_{i-1}), \dots, \varphi^T(\theta, t_{i-p+1})]^T \in \mathbb{R}^{p \times 4}. \tag{38}$$

Define  $f(p, t_{i-j}) := a \cos(\omega_1 t_{i-j}) + b \cos(\omega_2 t_{i-j}) \in \mathbb{R}$ , the vector  $F(p, \theta, t_i)$  and  $Y(p, t_i)$  as

$$F(p, \theta, t_i) := \begin{bmatrix} f(p, t_i) \\ f(p, t_{i-1}) \\ \vdots \\ f(p, t_{i-p+1}) \end{bmatrix} \in \mathbb{R}^p, \quad Y(p, t_i) := \begin{bmatrix} y(t_i) \\ y(t_{i-1}) \\ \vdots \\ y(t_{i-p+1}) \end{bmatrix} \in \mathbb{R}^p. \tag{39}$$

Then, the gradient vector can be expressed as

$$\text{grad}[J_2(\theta)] := \Phi^T(p, \theta, t_i)[Y(p, t_i) - F(p, \theta, t_i)]. \tag{40}$$

Let  $k = 1, 2, 3, \dots$  be an iterative variable, and  $\hat{\boldsymbol{\theta}}_k(t_i) := [\hat{a}_k(t_i), \hat{b}_k(t_i), \hat{\omega}_{1,k}(t_i), \hat{\omega}_{2,k}(t_i)]^T \in \mathbb{R}^4$  be the estimate of  $\boldsymbol{\theta}$  at iteration  $k$  and the sampling time  $t = t_i$ . Minimizing  $J_2(\boldsymbol{\theta})$  and using the negative gradient search, we can obtain the moving data window gradient-based iterative (MDW-GI) algorithm:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_k(t_i) &= \hat{\boldsymbol{\theta}}_{k-1}(t_i) - \mu_k(t_i) \text{grad}[J_2(\hat{\boldsymbol{\theta}}_{k-1}(t_i))] \\ &= \hat{\boldsymbol{\theta}}_{k-1}(t_i) - \mu_k(t_i) \boldsymbol{\Phi}^T(p, \hat{\boldsymbol{\theta}}_{k-1}(t_i), t_i) [\mathbf{Y}(p, t_i) - \mathbf{F}(p, \hat{\boldsymbol{\theta}}_{k-1}(t_i), t_i)], \end{aligned} \tag{41}$$

$$\begin{aligned} \hat{\boldsymbol{\Phi}}_k(p, t_i) &= \hat{\boldsymbol{\Phi}}_k(p, \hat{\boldsymbol{\theta}}_{k-1}(t_i), t_i) \\ &= [\boldsymbol{\varphi}(\hat{\boldsymbol{\theta}}_{k-1}(t_i), t_i), \boldsymbol{\varphi}(\hat{\boldsymbol{\theta}}_{k-1}(t_i), t_{i-1}), \dots, \boldsymbol{\varphi}(\hat{\boldsymbol{\theta}}_{k-1}(t_i), t_{i-p+1})]^T, \end{aligned} \tag{42}$$

$$\mathbf{Y}(p, t_i) = [y(t_i), y(t_{i-1}), \dots, y(t_{i-p+1})]^T, \tag{43}$$

$$\begin{aligned} \hat{\mathbf{F}}_k(p, t_i) &= \hat{\mathbf{F}}_k(p, \hat{\boldsymbol{\theta}}_{k-1}(t_i), t_i) \\ &= [f(\hat{\boldsymbol{\theta}}_{k-1}(t_i), t_i), f(\hat{\boldsymbol{\theta}}_{k-1}(t_i), t_{i-1}), \dots, f(\hat{\boldsymbol{\theta}}_{k-1}(t_i), t_{i-p+1})]^T, \end{aligned} \tag{44}$$

$$\begin{aligned} \hat{\boldsymbol{\varphi}}_k(t_{i-j}) &:= \boldsymbol{\varphi}(\hat{\boldsymbol{\theta}}_{k-1}(t_i), t_{i-j}) \\ &= [-\cos(\hat{\omega}_{1,k-1}(t_i)t_{i-j}), -\cos(\hat{\omega}_{2,k-1}(t_i)t_{i-j}), \hat{a}_{k-1}(t_i)t_{i-j} \sin(\hat{\omega}_{1,k-1}(t_i)t_{i-j}), \\ &\quad \hat{b}_{k-1}(t_i)t_{i-j} \sin(\hat{\omega}_{2,k-1}(t_i)t_{i-j})]^T, \end{aligned} \tag{45}$$

$$\begin{aligned} \hat{f}_k(t_{i-j}) &:= f(\hat{\boldsymbol{\theta}}_{k-1}(t_i), t_{i-j}) \\ &= \hat{a}_{k-1}(t_i) \cos(\hat{\omega}_{1,k-1}(t_i)t_{i-j}) + \hat{b}_{k-1}(t_i) \cos(\hat{\omega}_{2,k-1}(t_i)t_{i-j}), \end{aligned} \tag{46}$$

$$\hat{\boldsymbol{\theta}}_k(t_i) = [\hat{a}_k(t_i), \hat{b}_k(t_i), \hat{\omega}_{1,k}(t_i), \hat{\omega}_{2,k}(t_i)]^T, \tag{47}$$

$$0 < \mu_k(t_i) \leq \lambda_{\max}^{-1}[\hat{\boldsymbol{\Phi}}_k^T(p, t_i) \hat{\boldsymbol{\Phi}}_k(p, t_i)]. \tag{48}$$

The steps of the MDW-GI algorithm in Equations (41)–(48) to compute  $\hat{\boldsymbol{\theta}}_k$  are listed as follows.

1. Pre-set the length of  $p$ , let  $i = p$ , give the parameter estimation accuracy  $\varepsilon$  and the iterative length  $k_{\max} = 500$ .
2. To initialize, let  $k = 1$ ,  $\hat{\boldsymbol{\theta}}_0(t_i) = \mathbf{1}_n / p_0$ ,  $p_0 = 10^6$ .
3. Collect the observed data  $y(t_i)$ , form  $\mathbf{Y}(p, t_i)$  by Equation (43).
4. Form  $\hat{\boldsymbol{\varphi}}_k(t_{i-j})$  by Equation (45), form  $\hat{\boldsymbol{\Phi}}_k(p, t_i)$  by Equation (42).
5. Form  $\hat{f}_k(t_{i-j})$  by Equation (46), form  $\hat{\mathbf{F}}_k(p, t_i)$  by Equation (44).
6. Select a larger  $\mu_k(t_i)$  satisfying Equation (48), update the estimate  $\hat{\boldsymbol{\theta}}_k(t_i)$  by Equation (47).
7. If  $k < k_{\max}$ , increase  $k$  by 1 and go to Step 4; otherwise, go to the next step.
8. Compare  $\hat{\boldsymbol{\theta}}_k(t_i)$  with  $\hat{\boldsymbol{\theta}}_{k-1}(t_i)$ , if  $\|\hat{\boldsymbol{\theta}}_k(t_i) - \hat{\boldsymbol{\theta}}_{k-1}(t_i)\| > \varepsilon$ , then let  $\hat{\boldsymbol{\theta}}_0(t_{i+1}) := \hat{\boldsymbol{\theta}}_k(t_i)$ ,  $i := i + 1$ , go to Step 2; otherwise, obtain the parameter estimate  $\hat{\boldsymbol{\theta}}_k(t_i)$ , terminate this procedure.

### 5. Examples

Case 1: The numerical simulations of the three iterative algorithms. Consider the following dual-frequency cosine signal model:

$$y(t) = 1.48 \cos(0.06t) + 2.25 \cos(0.55t) + v(t).$$

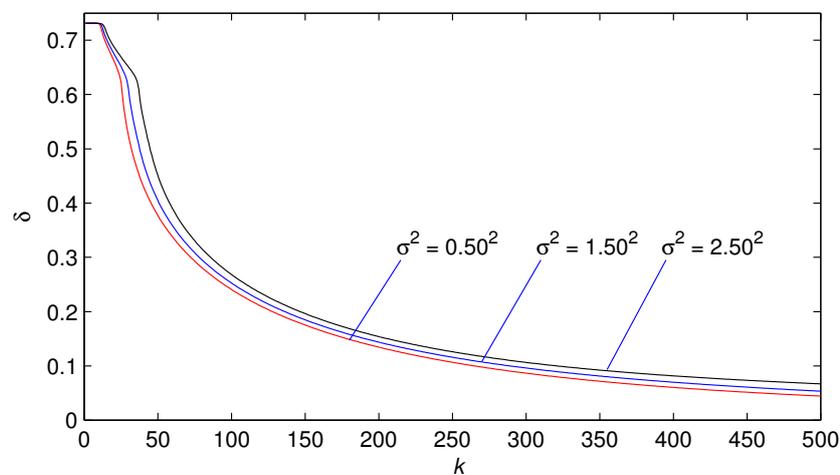
The parameter vector to be estimated is

$$\boldsymbol{\theta} = [a, b, \omega_1, \omega_2]^T = [1.48, 2.25, 0.06, 0.55]^T.$$

In simulation,  $\{v(t)\}$  is taken as a white noise sequence (stationary stochastic noise) with zero mean and variances  $\sigma^2 = 0.50^2$ ,  $\sigma^2 = 1.50^2$  and  $\sigma^2 = 2.50^2$ , respectively. Let  $t = kT$ , the sampling period  $T = 0.005$ ,  $k = 2000$ , the data length  $L = 2000$ . Apply the proposed GI algorithm using the observed data to estimate the parameters of this signal model. The parameter estimates and their estimation errors  $\delta := \|\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}\| / \|\boldsymbol{\theta}\| \times 100\%$  versus  $k$  are as shown in Table 1 and Figure 1.

**Table 1.** The GI parameter estimates and errors.

$\sigma^2$	$k$	$a$	$b$	$\omega_1$	$\omega_2$	$\delta$ (%)
$2.50^2$	10	0.55350	0.55350	0.00064	0.00148	73.11514
	20	0.61211	0.61528	0.03325	0.14768	68.89718
	50	1.06694	1.08276	0.00000	0.58116	45.10172
	100	1.20074	1.56991	0.00002	0.56483	26.83480
	200	1.28232	1.88059	0.00018	0.56030	15.39800
	500	1.34925	2.12940	0.01473	0.55767	6.68183
$1.50^2$	10	0.55290	0.55290	0.00212	0.00487	73.10844
	20	0.63483	0.63629	0.02549	0.17931	67.62638
	50	1.10299	1.20478	0.00002	0.56478	40.47663
	100	1.21921	1.60904	0.00014	0.55812	25.26461
	200	1.29994	1.90343	0.00137	0.55428	14.36541
	500	1.37665	2.14898	0.03582	0.55443	5.33204
$0.50^2$	10	0.55237	0.55240	0.00604	0.01387	73.04062
	20	0.65958	0.65653	0.01624	0.21377	66.34478
	50	1.09460	1.28918	0.00009	0.55852	37.71802
	100	1.22332	1.64271	0.00026	0.55228	24.07888
	200	1.31323	1.92538	0.00158	0.54864	13.44312
	500	1.39090	2.17489	0.02450	0.54741	4.43172
True values		1.48000	2.25000	0.06000	0.55000	



**Figure 1.** The GI estimation errors  $\delta$  versus  $k$ .

Apply the NI parameter estimation algorithm in Equations (18)–(35) with the finite observed data to estimate the parameters. The data length and the variances are the same as the condition in the GI algorithm. The parameter estimates and their estimation errors are given in Table 2 and the estimation errors  $\delta$  versus  $k$  are shown in Figure 2.

Apply the MDW-GI parameter estimation algorithm in Equations (41)–(48) to estimate the parameters of signal models. The length  $p$  of the moving data window is 100, 200 and 300, and  $\{v(t)\}$  is taken as a white noise sequence with zero mean and variance  $\sigma^2 = 0.50^2$ , respectively. The simulation results are shown in Table 3 and Figure 3.

**Table 2.** The NI parameter estimates and errors.

$\sigma^2$	$k$	$a$	$b$	$\omega_1$	$\omega_2$	$\delta$ (%)
2.50 <sup>2</sup>	1	1.83060	1.78268	0.09923	0.52087	21.32333
	2	1.84736	1.93621	0.13360	0.55852	17.77782
	10	1.68078	2.16420	0.11020	0.56216	8.16074
	20	1.65675	2.19656	0.10597	0.56159	6.93411
	100	1.62036	2.23024	0.09748	0.55889	5.34256
	200	1.60860	2.23778	0.09416	0.55763	4.86795
1.50 <sup>2</sup>	1	1.80786	1.79728	0.09883	0.52380	20.40210
	2	1.78233	1.95654	0.12241	0.55564	15.49323
	10	1.63563	2.16656	0.10287	0.56022	6.61985
	20	1.61619	2.19588	0.09943	0.56005	5.53200
	100	1.58796	2.22530	0.09256	0.55823	4.20939
	200	1.57898	2.23150	0.08988	0.55730	3.82961
0.50 <sup>2</sup>	1	1.78529	1.81118	0.09845	0.52670	19.51195
	2	1.73467	1.97038	0.11519	0.55453	13.90292
	10	1.60314	2.16384	0.09847	0.55978	5.65377
	20	1.58703	2.19039	0.09570	0.55991	4.65522
	100	1.56461	2.21617	0.09023	0.55871	3.50636
	200	1.55766	2.22129	0.08812	0.55803	3.19386
True values		1.48000	2.25000	0.06000	0.55000	

**Table 3.** The MDW-GI parameter estimates and errors ( $\sigma^2 = 0.50^2$ ).

$p$	$k$	$a$	$b$	$\omega_1$	$\omega_2$	$\delta$ (%)
100	1	1.58254	1.64339	0.07785	0.57750	22.40827
	2	1.87594	1.97597	0.11125	0.57907	17.64461
	5	1.83227	2.00369	0.11020	0.57801	15.77345
	10	1.77887	2.05136	0.10879	0.57605	13.20660
	20	1.69892	2.12339	0.10526	0.57220	9.37923
	50	1.57751	2.23398	0.09153	0.56174	3.79684
200	1	1.78604	1.85970	0.10256	0.57011	18.12068
	2	1.81880	1.95055	0.10931	0.57243	16.56384
	5	1.73791	2.01245	0.10997	0.57420	12.91251
	10	1.64148	2.08926	0.10871	0.57449	8.52118
	20	1.53916	2.17165	0.09906	0.56944	3.90746
	50	1.47938	2.22055	0.08073	0.56051	1.36478
300	1	1.74606	1.85586	0.10881	0.56979	17.40188
	2	1.78854	1.97309	0.11616	0.57303	15.23989
	5	1.69704	2.05038	0.11528	0.57458	10.94891
	10	1.59705	2.12840	0.10877	0.57251	6.44220
	20	1.50943	2.19678	0.09034	0.56412	2.52463
	50	1.46715	2.23197	0.07030	0.55621	0.91639
True values		1.48000	2.25000	0.06000	0.55000	

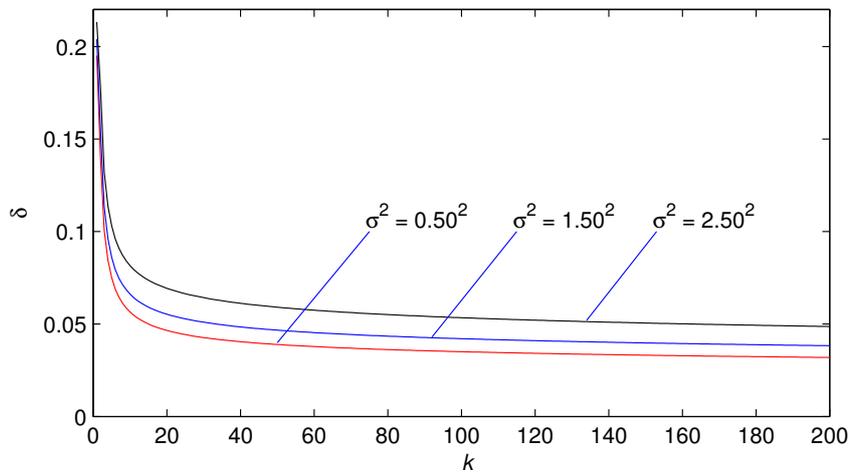


Figure 2. The NI estimation errors  $\delta$  versus  $k$ .

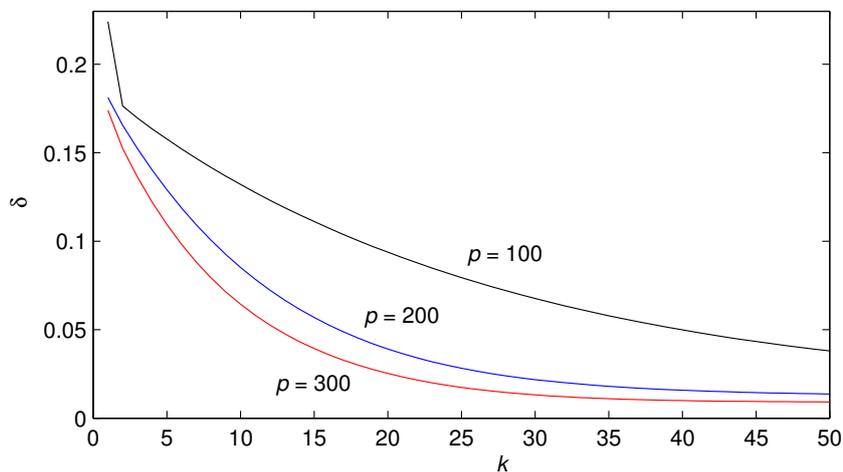


Figure 3. The MDW-GI estimation errors  $\delta$  versus  $k$  ( $\sigma^2 = 0.50^2$ ).

In order to validate the obtained models, we use the MDW-GI parameter estimates from next-to-last row in Table 3 with  $p = 300$  to construct the MDW-GI estimated model

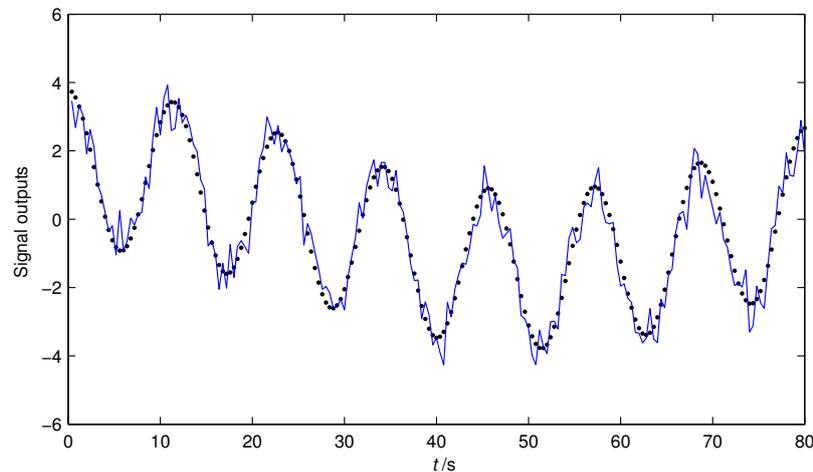
$$\hat{y}(t) = 1.46715 \cos(0.07030t) + 2.23197 \cos(0.55621t).$$

The estimated outputs  $\hat{y}(t)$  and the observed signal model  $y(t)$  are plotted in Figure 4, where the solid-line denotes the actual signal  $y(t)$  and the dot-line denotes the estimated signal  $\hat{y}(t)$ .

From the simulation results, we can draw the following conclusions:

- The parameter estimation errors obtained by the presented three algorithms gradually decreasing trend as the iterative variable  $k$  increases.
- The parameter estimation errors given by three algorithms become smaller with the noise variance decreasing.
- Tables 1 and 2 and Figures 1 and 2 show that the proposed NI algorithm has faster convergence speed and more accurate parameter estimates than the GI algorithm under the same simulation conditions.
- The larger the length  $p$  of the moving data window is, the faster the parameter estimation converge speed is (see Figure 3). Moreover, the larger the length  $p$ , the higher the accuracy is (see Table 3).

- In the simulation, these three algorithms are fulfilled in the same conditions ( $\sigma^2 = 0.50^2$ ), and the estimated models obtained by the MDW-GI and NI algorithms have higher accuracy than the GI algorithm.
- The outputs of estimated signal  $\hat{y}(t)$  are very close to the actual signal model  $y(t)$ . In other words, the estimated model can capture the dynamics of the signal.



**Figure 4.** The estimated outputs  $\hat{y}(t)$  and the actual signal model  $y(t)$  versus  $t$ . Solid line: the actual output  $y(t)$ , dot-line: the estimated signal  $\hat{y}(t)$ .

## 6. Conclusions

This paper studies the direct parameter estimation algorithms for signal models only using the discrete observed data. By using the gradient search, we need to select a small step-size in order to ensure the convergence, and this will increase the search time and decrease the convergence rate. The proposed NI and MDW-GI algorithms have higher accuracy than the GI estimation algorithm for estimating the unknown parameters. The MWD-GI is used to obtain the parameter estimates at the current moment based on the estimates of previous data obtained at the moment time. Furthermore, the MDW-GI algorithm based on the moving data window technique can estimate the parameters of signal models in real time. The proposed algorithms can be extended to multi-frequency signal models. In the next work, we will consider the estimation of the initial phase of the signal models, that is to say, the signal model is a highly nonlinear function in regard to the frequencies and phases, and estimate all the parameters of dual-frequency signals including the unknown amplitudes, frequencies and initial phases simultaneously.

**Acknowledgments:** This work was supported by the Science Research of Colleges Universities in Jiangsu Province (No. 16KJB120007, China), and sponsored by Qing Lan Project and the National Natural Science Foundation of China (No. 61492195).

**Author Contributions:** Feng Ding conceived the idea and supervised his student Siyu Liu to write the paper and Siyu Liu designed and performed the simulation experiments. Ling Xu was involved in the writing of this research article. All authors have read and approved the final manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ding, F. Complexity, convergence and computational efficiency for system identification algorithms. *Control Decis.* **2016**, *31*, 1729–1741.
2. Ding, F.; Wang, F.F. Recursive least squares identification algorithms for linear-in-parameter systems with missing data. *Control Decis.* **2016**, *31*, 2261–2266.

3. Xu, L. Moving data window based multi-innovation stochastic gradient identification method for transfer functions. *Control Decis.* **2017**, *32*, 1091–1096.
4. Xu, L.; Ding, F. Recursive least squares and multi-innovation stochastic gradient parameter estimation methods for signal modeling. *Syst. Signal Process.* **2017**, *36*, 1735–1753.
5. Xu, L. The parameter estimation algorithms based on the dynamical response measurement data. *Adv. Mech. Eng.* **2017**, *9*, doi:10.1177/1687814017730003.
6. Zhang, Q.G. The Rife frequency estimation algorithm based on real-time FFT. *Signal Process.* **2009**, *25*, 1002–1004.
7. Yang, C.; Wei, G. A noniterative frequency estimator with rational combination of three spectrum lines. *IEEE Trans. Signal Process.* **2011**, *59*, 5065–5070.
8. Jacobsen, E.; Kootsookos, P. Fast accurate frequency estimators. *IEEE Signal Process. Mag.* **2007**, *24*, 123–125.
9. Deng, Z.M.; Liu, H.; Wang, Z.Z. Modified Rife algorithm for frequency estimation of sinusoid wave. *J. Data Acquis. Process.* **2006**, *21*, 473–477.
10. Elasm-Ksibi, R.; Besbes, H.; López-Valcarce, R.; Cherif, S. Frequency estimation of real-valued single-tone in colored noise using multiple autocorrelation lags. *Signal Process.* **2010**, *90*, 2303–2307.
11. So, H.C.; Chan, K.W. Reformulation of Pisarenko harmonic decomposition method for single-tone frequency estimation. *IEEE Trans. Signal Process.* **2004**, *52*, 1128–1135.
12. Cao, Y.; Wei, G.; Chen, F.J. An exact analysis of modified covariance frequency estimation algorithm based on correlation of single-tone. *Signal Process.* **2012**, *92*, 2785–2790.
13. Boashash, B.; Khan, N.A.; Ben-Jabeur, T. Time-frequency features for pattern recognition using high-resolution TFDs: A tutorial review. *Digit. Signal Process.* **2015**, *40*, 1–30.
14. Amezcua-Sanchez, J.P.; Adeli, H. A new music-empirical wavelet transform methodology for time-frequency analysis of noisy nonlinear and non-stationary signals. *Digit. Signal Process.* **2015**, *45*, 56–68.
15. Daubechies, I.; Lu, J.F.; Wu, H.T. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Appl. Comput. Harmon. Anal.* **2011**, *30*, 243–261.
16. Ding, F.; Xu, L.; Liu, X.M. Signal modeling—Part A: Single-frequency signals. *J. Qingdao Univ. Sci. Technol. (Nat. Sci. Ed.)* **2017**, *38*, 1–13. (In Chinese)
17. Ding, F.; Xu, L.; Liu, X.M. Signal modeling—Part B: Dual-frequency signals. *J. Qingdao Univ. Sci. Technol. (Nat. Sci. Ed.)* **2017**, *38*, 1–17. (In Chinese)
18. Ding, F.; Xu, L.; Liu, X.M. Signal modeling—Part C: Recursive parameter estimation for multi-frequency signal models. *J. Qingdao Univ. Sci. Technol. (Nat. Sci. Ed.)* **2017**, *38*, 1–12. (In Chinese)
19. Ding, F.; Xu, L.; Liu, X.M. Signal modeling—Part D: Iterative parameter estimation for multi-frequency signal models. *J. Qingdao Univ. Sci. Technol. (Nat. Sci. Ed.)* **2017**, *38*, 1–11. (In Chinese)
20. Ding, F.; Xu, L.; Liu, X.M. Signal modeling—Part E: Hierarchical parameter estimation for multi-frequency signal models. *J. Qingdao Univ. Sci. Technol. (Nat. Sci. Ed.)* **2017**, *38*, 1–15. (In Chinese)
21. Ding, F.; Xu, L.; Liu, X.M. Signal modeling—Part F: Hierarchical iterative parameter estimation for multi-frequency signal models. *J. Qingdao Univ. Sci. Technol. (Nat. Sci. Ed.)* **2017**, *38*, 1–12. (In Chinese)
22. Ding, J.L. Data filtering based recursive and iterative least squares algorithms for parameter estimation of multi-input output systems. *Algorithms* **2016**, *9*, doi:10.3390/a9030049.
23. Yun, B.I. Iterative methods for solving nonlinear equations with finitely many roots in an interval. *J. Comput. Appl. Math.* **2013**, *236*, 3308–3318.
24. Dehghan, M.; Hajarian, M. Analysis of an iterative algorithm to solve the generalized coupled Sylvester matrix equations. *Appl. Math. Model.* **2011**, *35*, 3285–3300.
25. Wang, T.; Zheng, Q.Q.; Lu, L.Z. A new iteration method for a class of complex symmetric linear systems. *J. Comput. Appl. Math.* **2017**, *325*, 188–197.
26. Xu, L. Application of the Newton iteration algorithm to the parameter estimation for dynamical systems. *J. Comput. Appl. Math.* **2015**, *288*, 33–43.
27. Pei, K.; Wang, G.T.; Sun, Y.Y. Successive iterations and positive extremal solutions for a Hadamard type fractional integro-differential equations on infinite domain. *Appl. Math. Comput.* **2017**, *312*, 158–168.
28. Dehghan, M.; Hajarian, M. Fourth-order variants of Newtons method without second derivatives for solving nonlinear equations. *Eng. Comput.* **2012**, *29*, 356–365.

29. Gutiérrez, J.M. Numerical properties of different root-finding algorithms obtained for approximating continuous Newton's method. *Algorithms* **2015**, *8*, 1210–1216.
30. Wang, X.F.; Qin, Y.P.; Qian, W.Y.; Zhang, S.; Fan, X.D. A family of Newton type iterative methods for solving nonlinear equations. *Algorithms* **2015**, *8*, 786–798.
31. Simpson, T. *The Nature and Laws of Chance*; University of Michigan Library: Ann Arbor, MI, USA, 1740.
32. Dennis, J.E.; Schnable, R.B. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1983.
33. Jürgen, G. Accelerated convergence in Newton's method. *Soc. Ind. Appl. Math.* **1994**, *36*, 272–276.
34. Djoudi, M.S.; Kennedy, D.; Williams, F.W. Exact substructuring in recursive Newton's method for solving transcendental eigenproblems. *J. Sound Vib.* **2005**, *280*, 883–902.
35. Benner, P.; Li, J.R.; Penzl, T. Numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems. *Numer. Linear Algebra Appl.* **2008**, *15*, 755–777.
36. Seinfeld, D.R.; Haddad, W.M.; Bernstein, D.S.; Nett, C.N.  $H_2/H_\infty$  controller synthesis: Illustrative numerical results via quasi-newton methods. *Numer. Linear Algebra Appl.* **2008**, *15*, 755–777.
37. Liu, M.M.; Xiao, Y.S.; Ding, R.F. Iterative identification algorithm for Wiener nonlinear systems using the Newton method. *Appl. Math. Model.* **2013**, *37*, 6584–6591.
38. Curry, H.B. The method of steepest descent for non-linear minimization problems. *Q. Appl. Math.* **1944**, *2*, 258–261.
39. Vrahatis, M.N.; Androulakis, G.S.; Lambrinos, J.N.; Magoulas, G.D. A class of gradient unconstrained minimization algorithms with adaptive stepsize. *J. Comput. Appl. Math.* **2000**, *114*, 367–386.
40. Hajarian, M. Solving the general Sylvester discrete-time periodic matrix equations via the gradient based iterative method. *Appl. Math. Lett.* **2016**, *52*, 87–95.
41. Wang, C.; Wang, J.Y.; Zhang, T.S. Operational modal analysis for slow linear time-varying structures based on moving window second order blind identification. *Signal Process.* **2017**, *133*, 169–186.
42. Al-Matouq, A.A.; Vincent, T.L. Multiple window moving horizon estimation. *Automatica* **2015**, *53*, 264–274.
43. Boashash, B. Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals. *Proc. IEEE* **1992**, *80*, 520–538.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).