


Gradient-Based Iterative Parameter Estimation Algorithms for Dynamical Systems from Observation Data

Feng Ding ^{1,2,3,*} , Jian Pan ¹, Ahmed Alsaedi ⁴ and Tasawar Hayat ⁴

¹ School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan 430068, China; jpan@163.com

² College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, China

³ School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China

⁴ Department of Mathematics, King Abdulaziz University, Jeddah 21589, Saudi Arabia; aalsaedi@hotmail.com (A.A.); fmgpak@gmail.com (T.H.)

* Correspondence: fding@jiangnan.edu.cn or fdingab@126.com

Received: 21 March 2019; Accepted: 29 April 2019; Published: 14 May 2019



Abstract: It is well-known that mathematical models are the basis for system analysis and controller design. This paper considers the parameter identification problems of stochastic systems by the controlled autoregressive model. A gradient-based iterative algorithm is derived from observation data by using the gradient search. By using the multi-innovation identification theory, we propose a multi-innovation gradient-based iterative algorithm to improve the performance of the algorithm. Finally, a numerical simulation example is given to demonstrate the effectiveness of the proposed algorithms.

Keywords: parameter estimation; iterative algorithm; gradient search; multi-innovation identification; stochastic system

1. Introduction

Parameter estimation deals with the problem of building mathematical models of systems [1–5] based on observation data [6–8] and is the basis for system identification [9–11]. It has widespread applications in many areas [12–16]. For instance, system identification is used to get the appropriate models for control system designs [17,18], simulations and predictions [19,20] in control and system modelling [21,22]. Recently, Bottegal et al. proposed a two-experiment method to identify Wiener systems by using the data acquired from two separate experiments, in which the first experiment estimates the static nonlinearity and the second experiment identifies the linear block based on the estimated nonlinearity [23]. For nonlinear errors-in-variables systems contaminated with outliers, a robust identification approach is presented by means of the expectation maximization under the framework of the maximum likelihood estimation [24,25].

The recursive identification and the iterative identification are two important types of parameter estimation methods [26,27]. In recursive identification methods, the parameter estimates can be computed recursively in real-time [28,29]. Differing from the recursive methods, the basic idea of the iterative methods is to update the parameter estimates by using batch data [30–32]. In recent years, a number of iterative identification methods have been proposed for all kinds of systems. A data filtering-based iterative estimation algorithm is studied for an infinite impulse response filter with colored noise [33]. For Box-Jenkins models, Liu et al. proposed a least squares-based iterative algorithm according to the auxiliary model identification idea and the iterative search [34]. Other related work includes the recursive algorithms [35,36] and the iterative algorithms [37–39].

The multi-innovation theory provides a new idea for system identification [40,41]. The innovation is the useful information that can improve the parameter estimation accuracy. The basic idea of the multi-innovation theory is to expand the dimension of the innovation and to enhance the data utilization. A brief review of the multi-innovation algorithms is listed as follows. An adaptive filtering based multi-innovation stochastic gradient algorithm was derived for bilinear systems with colored noise and can give small parameter estimation errors as the innovation length increases [42]. A multi-innovation gradient algorithm was developed based on the Kalman filtering to solve the joint state and parameter estimation problem for a nonlinear state space system with time-delay [43].

Identification is to fit a closest model of a practical system from observation data. The commonly used techniques are some optimization methods. These methods include linear programming and convex mixed integer nonlinear programs [44]. Some related optimization algorithms and solvers have been proposed for complete global optimization problems [45], such as the general algebraic modeling system (GAMS) and mixed-integer nonlinear programming (MINLP) solver [46] and the mixed-integer linear programming (MILP) and MINLP approaches for solving nonlinear discrete transportation problem [47]. This paper focuses on the parameter identification problems of controlled autoregressive systems by using the gradient search [48] and the multi-innovation identification theory [49]. The basic idea is to use the iterative technique for computing the parameter estimation through defining and minimizing two quadratic criterion functions. The main contributions of this paper are as follows.

- Based on the gradient search, a gradient-based iterative algorithm is presented for identifying the parameters of controlled autoregressive systems.
- A multi-innovation gradient-based iterative algorithm is derived for improving the performance of the algorithm by using the multi-innovation identification theory.

The outlines of this paper are organized as follows. Section 2 defines some definitions and derives the identification model of controlled autoregressive systems. Section 3 derives a gradient-based iterative algorithm. Section 4 proposes a multi-innovation gradient-based iterative algorithm. Section 5 offers an example to illustrate the effectiveness of the proposed algorithms. Finally, Section 6 gives some concluding remarks.

2. The System Description

Let us introduce some symbols used in this paper. The symbol I_n denotes an identity matrix of appropriate size ($n \times n$); $\mathbf{1}_n$ stands for an n -dimensional column vector whose elements are 1; the superscript T stands for the vector/matrix transpose; the norm of a matrix (or a column vector) X is defined by $\|X\|^2 := \text{tr}[XX^T]$.

Consider the dynamic stochastic systems described by the following controlled autoregressive (CAR) model:

$$A(z)y(t) = B(z)u(t) + v(t), \quad (1)$$

where $\{u(t)\}$ is the input sequence of the system and $\{y(t)\}$ is the output sequence of the system, $\{v(t)\}$ is a white noise sequence with zero mean, $A(z)$ and $B(z)$ are polynomials in the unit backward shift operator [$z^{-1}y(t) = y(t-1)$, $zy(t) = y(t+1)$], and defined as

$$\begin{aligned} A(z) &:= 1 + a_1z^{-1} + a_2z^{-2} + \cdots + a_{n_a}z^{-n_a}, \\ B(z) &:= b_1z^{-1} + b_2z^{-2} + \cdots + b_{n_b}z^{-n_b}. \end{aligned}$$

Assume that n_a and n_b are known, and $y(t) = 0$, $u(t) = 0$ and $v(t) = 0$ for $t \leq 0$.

Let $n := n_a + n_b$, define the parameter vectors:

$$\begin{aligned}\boldsymbol{\vartheta} &:= \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \in \mathbb{R}^n, \\ \mathbf{a} &:= [a_1, a_2, \dots, a_{n_a}]^T \in \mathbb{R}^{n_a}, \\ \mathbf{b} &:= [b_1, b_2, \dots, b_{n_b}]^T \in \mathbb{R}^{n_b},\end{aligned}$$

and the corresponding information vectors:

$$\begin{aligned}\boldsymbol{\varphi}(t) &:= \begin{bmatrix} \boldsymbol{\varphi}_a(t) \\ \boldsymbol{\varphi}_b(t) \end{bmatrix} \in \mathbb{R}^n, \\ \boldsymbol{\varphi}_a(t) &:= [-y(t-1), -y(t-2), \dots, -y(t-n_a)]^T \in \mathbb{R}^{n_a}, \\ \boldsymbol{\varphi}_b(t) &:= [u(t-1), u(t-2), \dots, u(t-n_b)]^T \in \mathbb{R}^{n_b}.\end{aligned}$$

Through the above definitions, the system in Equation (1) can be rewritten as

$$\begin{aligned}y(t) &= [1 - A(z)]y(t) + B(z)u(t) + v(t) \\ &= (-a_1z^{-1} - a_2z^{-2} - \dots - a_{n_a}z^{-n_a})y(t) + (b_1z^{-1} + b_2z^{-2} + \dots + b_{n_b}z^{-n_b})u(t) + v(t) \\ &= -a_1y(t-1) - a_2y(t-2) - \dots - a_{n_a}y(t-n_a) \\ &\quad + b_1u(t-1) + b_2u(t-2) + \dots + b_{n_b}u(t-n_b) + v(t) \\ &= [-y(t-1), -y(t-2), \dots, -y(t-n_a)]\mathbf{a} + [u(t-1), u(t-2), \dots, u(t-n_b)]\mathbf{b} + v(t) \\ &= \boldsymbol{\varphi}_a^T(t)\mathbf{a} + \boldsymbol{\varphi}_b^T(t)\mathbf{b} + v(t) \\ &= \boldsymbol{\varphi}^T(t)\boldsymbol{\vartheta} + v(t).\end{aligned}\tag{2}$$

This identification model for System in (1) involves two different parameter vectors \mathbf{a} and \mathbf{b} , which are merged to a new vector $\boldsymbol{\vartheta}$, and two different information vectors $\boldsymbol{\varphi}_a(t)$ and $\boldsymbol{\varphi}_b(t)$, which are merged to a new vector $\boldsymbol{\varphi}(t)$ composed of available input-output data $u(t-i)$ and $y(t-i)$. The objective of this paper is by means of the gradient search and based on the multi-innovation identification theory to derive new identification algorithms for estimating the system parameter vectors from available data $\{u(t), y(t)\}$.

3. The Gradient-Based Iterative Algorithm

In this section, we first define a criterion function and present a gradient-based iterative algorithm by using the gradient search. In addition, a brief discussion about the choice of the iterative step-size is given in this section.

Let $k = 1, 2, 3, \dots$ be an iterative variable and $\hat{\boldsymbol{\vartheta}}_k := \begin{bmatrix} \hat{\mathbf{a}}_k \\ \hat{\mathbf{b}}_k \end{bmatrix}$ be the iterative estimate of $\boldsymbol{\vartheta} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}$ at iteration k . $\lambda_{\max}[\mathbf{X}]$ is the largest eigenvalue of the symmetric matrix \mathbf{X} .

The observation data are basic for parameter identification algorithms. This paper uses a batch of data with the length L and is based on Model (2), we define the stacked output data vector $\mathbf{Y}(L)$ and the stacked data matrix $\boldsymbol{\Phi}(L)$ as

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

Define a static criterion function

$$J_1(\boldsymbol{\vartheta}) := \frac{1}{2} \sum_{t=1}^L [y(t) - \boldsymbol{\varphi}^T(t)\boldsymbol{\vartheta}]^2,$$

which can be equivalently expressed as

$$J_1(\boldsymbol{\vartheta}) := \frac{1}{2} \|\mathbf{Y}(L) - \boldsymbol{\Phi}(L)\boldsymbol{\vartheta}\|^2.$$

By means of the negative gradient search, computing the partial derivative of $J_1(\boldsymbol{\vartheta})$ with respect to $\boldsymbol{\vartheta}$, we can obtain the iterative relation:

$$\begin{aligned} \hat{\boldsymbol{\vartheta}}_k &= \hat{\boldsymbol{\vartheta}}_{k-1} - \mu \text{grad}[J_1(\hat{\boldsymbol{\vartheta}}_{k-1})] \\ &= \hat{\boldsymbol{\vartheta}}_{k-1} + \mu \boldsymbol{\Phi}^T(L) [\mathbf{Y}(L) - \boldsymbol{\Phi}(L)\hat{\boldsymbol{\vartheta}}_{k-1}] \\ &= [\mathbf{I}_n - \mu \boldsymbol{\Phi}^T(L)\boldsymbol{\Phi}(L)]\hat{\boldsymbol{\vartheta}}_{k-1} + \mu \boldsymbol{\Phi}^T(L)\mathbf{Y}(L), \end{aligned}$$

where $\mu > 0$ is an iterative step-size or a convergence factor. The above equation can be seen as a discrete-time system. To ensure the convergence of $\hat{\boldsymbol{\vartheta}}_k$, all the eigenvalues of matrix $[\mathbf{I} - \mu \hat{\boldsymbol{\Phi}}^T(L)\hat{\boldsymbol{\Phi}}(L)]$ must be in the unit circle, that is to say μ should satisfy $-\mathbf{I}_n \leq \mathbf{I}_n - \mu \boldsymbol{\Phi}^T(L)\boldsymbol{\Phi}(L) \leq \mathbf{I}_n$, or $0 \leq \mu \boldsymbol{\Phi}^T(L)\boldsymbol{\Phi}(L) \leq 2\mathbf{I}_n$, so a conservative choice of μ is

$$\mu \leq \frac{2}{\lambda_{\max}[\boldsymbol{\Phi}^T(L)\boldsymbol{\Phi}(L)]} = 2\lambda_{\max}^{-1}[\boldsymbol{\Phi}^T(L)\boldsymbol{\Phi}(L)].$$

In order to avoid computing the complicated eigenvalues of a square matrix and to reduce computational cost, we use the trace of the matrix and take another way for selecting the step-size

$$\mu \leq \frac{2}{\|\boldsymbol{\Phi}(L)\|^2} = 2\|\boldsymbol{\Phi}(L)\|^{-2}.$$

Then we can obtain the gradient-based iterative (GI) algorithm for estimating the parameter vector $\boldsymbol{\vartheta}$ of the CAR system in (1) [4]:

$$\hat{\boldsymbol{\vartheta}}_k = \hat{\boldsymbol{\vartheta}}_{k-1} + \mu \boldsymbol{\Phi}^T(L) [\mathbf{Y}(L) - \boldsymbol{\Phi}(L)\hat{\boldsymbol{\vartheta}}_{k-1}], \quad k = 1, 2, 3, \dots \quad (3)$$

$$\mu = 2\lambda_{\max}^{-1}[\boldsymbol{\Phi}^T(L)\boldsymbol{\Phi}(L)], \quad \text{or} \quad \mu = 2\|\boldsymbol{\Phi}(L)\|^{-2}, \quad (4)$$

$$\mathbf{Y}(L) = [y(1), y(2), \dots, y(L)]^T, \quad (5)$$

$$\boldsymbol{\Phi}(L) = [\boldsymbol{\varphi}(1), \boldsymbol{\varphi}(2), \dots, \boldsymbol{\varphi}(L)]^T, \quad (6)$$

$$\boldsymbol{\varphi}(t) = \begin{bmatrix} \boldsymbol{\varphi}_a(t) \\ \boldsymbol{\varphi}_b(t) \end{bmatrix}, \quad t = 1, 2, \dots, L, \quad (7)$$

$$\boldsymbol{\varphi}_a(t) = [-y(t-1), -y(t-2), \dots, -y(t-n_a)]^T, \quad (8)$$

$$\boldsymbol{\varphi}_b(t) = [u(t-1), u(t-2), \dots, u(t-n_b)]^T. \quad (9)$$

The steps of computing $\hat{\boldsymbol{\vartheta}}_k$ involved in the GI algorithm in Equations (3)–(9) are summarized in the following. The pseudo-code of implementing the gradient-based iterative algorithm is shown in Algorithm 1.

1. For $t \leq 0$, all the variables are set to zero. Let $k = 1$, give the data length L ($L \gg n$) and set the initial values: $\hat{\boldsymbol{\vartheta}}_0 = \mathbf{1}_n/p_0$, $p_0 = 10^6$, and the parameter estimation accuracy ε .
2. Collect the input and output data $u(t)$ and $y(t)$, $t = 1, 2, \dots, L$.
3. Form the information vectors $\boldsymbol{\varphi}_a(t)$, $\boldsymbol{\varphi}_b(t)$ and $\boldsymbol{\varphi}(t)$ using Equations (8)–(9) and (7).
4. Construct the stacked output vector $\mathbf{Y}(L)$ by Equation (5) and the stacked information matrix $\boldsymbol{\Phi}(L)$ by Equation (6), select a large μ according to Equation (4).

5. Update the parameter vector estimate $\hat{\boldsymbol{\theta}}_k$ using Equation (3).
6. Compare $\hat{\boldsymbol{\theta}}_k$ with $\hat{\boldsymbol{\theta}}_{k-1}$: If $\|\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_{k-1}\| > \varepsilon$, increase k by 1 and go to step 5; otherwise, obtain the iteration k and the parameter estimation vector $\hat{\boldsymbol{\theta}}_k$.

Algorithm 1 The pseudo-code of implementing the gradient-based iterative algorithm.

Data: $\{u(t), y(t), t = 1, 2, \dots, L\}, n_a, n_b, \varepsilon, N$

Result: $\hat{\boldsymbol{\theta}}_k$

```

1: Initialization:  $\hat{\boldsymbol{\theta}}_0 = [1, 1, \dots, 1]^T / p_0 \in \mathbb{R}^{n_a+n_b}, p_0 = 10^6$ .
2: for  $t = 1 : L$  do
3:   Form the information vectors  $\boldsymbol{\varphi}_a(t), \boldsymbol{\varphi}_b(t)$  and  $\boldsymbol{\varphi}(t)$  using (8)–(9) and (7).
4: end for
5: Construct the stacked output vector  $\mathbf{Y}(L)$  by (5).
6: Construct the stacked information matrix  $\boldsymbol{\Phi}(L)$  by (6).
7: for  $k = 1$  do
8:   Compute the step-size  $\mu$  using (4):
9:    $\mu = 2\lambda_{\max}^{-1}[\boldsymbol{\Phi}^T(L)\boldsymbol{\Phi}(L)]$  or  $\mu = 2\|\boldsymbol{\Phi}(L)\|^{-2}$ ;
10:  Update the parameter vector estimate  $\hat{\boldsymbol{\theta}}_k$  using (3):
11:   $\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k + \mu\boldsymbol{\Phi}^T(L)[\mathbf{Y}(L) - \boldsymbol{\Phi}(L)\hat{\boldsymbol{\theta}}_k]$ ;
12:  If  $\|\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_{k-1}\| > \varepsilon$ 
13:     $k = k + 1$ ;
14:  else
15:    Obtain the iteration  $k$  and the parameter estimation vector  $\hat{\boldsymbol{\theta}}_k$ , break;
16:  end
17: end for

```

The flowchart of computing $\hat{\boldsymbol{\theta}}_k$ in the GI algorithm is shown in Figure 1.

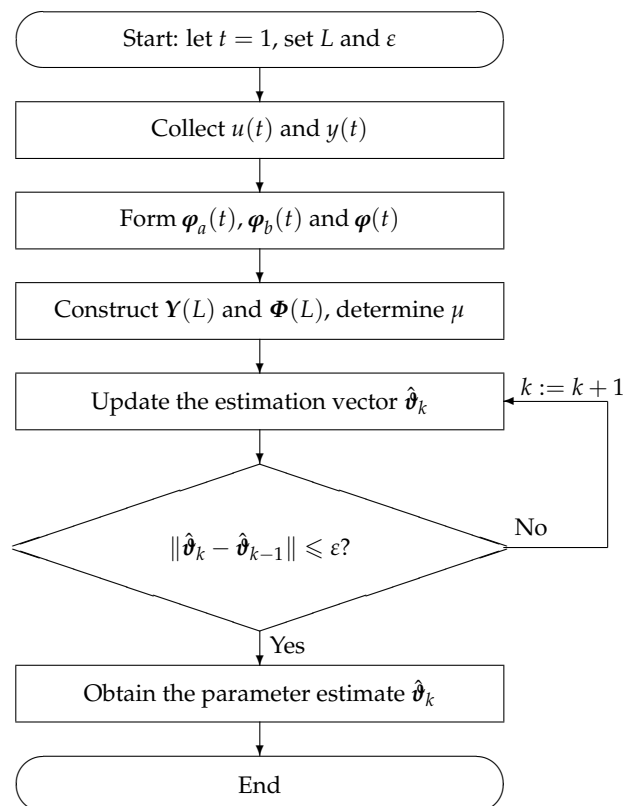


Figure 1. The flowchart of the gradient-based iterative (GI) algorithm.

Remark 1. The criterion function is quadratic and a convex optimization problem. For a static optimization problem, a constant step-size is often used for simplification in parameter estimation algorithms.

Remark 2. The computation is an important property of an algorithm, and is represented by the floating-point operation of multiplications/divisions and additions/subtractions (the flop for short). Table 1 gives the computational efficiency of the GI algorithm.

Table 1. The computational efficiency of the gradient-based iterative (GI) algorithm.

Variables	Expressions	Multiplications	Additions
$\hat{\boldsymbol{\theta}}_k$	$\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_{k-1} + \mu \boldsymbol{\Phi}^T(L)[\mathbf{Y}(L) - \boldsymbol{\Phi}(L)\hat{\boldsymbol{\theta}}_{k-1}] \in \mathbb{R}^n$	$(2nL + n)k$	$2nLk$
μ	$\mu \leq 2\ \boldsymbol{\Phi}(L)\ ^{-2} \in \mathbb{R}$	$nL + 1$	$nL - 1$
Sum		$(2nL + n)k + nL + 1$	$2nLk + nL - 1$
Total flops		$N_1 := (4nL + n)k + 2nL$	

Remark 3. The gradient algorithm can be used to find the optimal solution for quadratic optimization problems and nonlinear optimization problems. It can handle not only the linear regression systems with the known information vector, but also the linear and nonlinear systems with the unknown information vector. However, Under the same data length L , the estimation accuracy of the GI algorithm becomes high as the iterative index k increases, and when the data length L increases, the estimation accuracy also becomes high. Thus, we choose a sufficient large data length in practice

Remark 4. If we choose the step-size as $\mu = 2\|\boldsymbol{\Phi}(L)\|^{-2}$ and eliminate the intermediate variables $\mathbf{Y}(L)$ and $\boldsymbol{\Phi}(L)$, the GI algorithm can be equivalently transformed into

$$\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_{k-1} + 2 \left[\sum_{j=1}^L \|\boldsymbol{\varphi}(j)\|^2 \right]^{-1} \left[\sum_{j=1}^L \boldsymbol{\varphi}(j)[y(j) - \boldsymbol{\varphi}^T(j)\hat{\boldsymbol{\theta}}_{k-1}] \right]. \quad (10)$$

In order to improve the performance of the GI algorithm, we introduce a forgetting factor $0 \leq \lambda \leq 1$ and obtain the forgetting factor GI algorithm:

$$\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_{k-1} + 2 \left[\sum_{t=1}^L \lambda^{L-t} \|\boldsymbol{\varphi}(t)\|^2 \right]^{-1} \left[\sum_{t=1}^L \lambda^{L-t} \boldsymbol{\varphi}(t)[y(t) - \boldsymbol{\varphi}^T(t)\hat{\boldsymbol{\theta}}_{k-1}] \right]. \quad (11)$$

4. The Multi-Innovation Gradient-Based Iterative Algorithm

The multi-innovation identification algorithm uses the data in a moving data window to update the parameter estimates. The moving data window, which is also called the dynamical data window, moves along with t . The length of the moving data window can be variable or invariable. In this section, we derive a multi-innovation gradient-based iterative algorithm with the constant window length.

Consider the newest p data from $j = t - p + 1$ to $j = t$ (p represents the data length, i.e., the length of the moving data window), and define the stacked output vector $\mathbf{Y}(p, t)$ and the stacked information matrix $\boldsymbol{\Phi}(p, t)$ as

$$\mathbf{Y}(p, t) := \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-p+1) \end{bmatrix} \in \mathbb{R}^p, \quad \boldsymbol{\Phi}(p, t) := \begin{bmatrix} \boldsymbol{\varphi}^T(t) \\ \boldsymbol{\varphi}^T(t-1) \\ \vdots \\ \boldsymbol{\varphi}^T(t-p+1) \end{bmatrix} \in \mathbb{R}^{p \times n}. \quad (12)$$

According to the identification model in (2), define a dynamic data window criterion function:

$$J_2(\boldsymbol{\theta}) := \frac{1}{2} \sum_{j=t-p+1}^t [y(j) - \boldsymbol{\varphi}^T(j)\boldsymbol{\theta}]^2 = \frac{1}{2} \|\mathbf{Y}(p, t) - \boldsymbol{\Phi}(p, t)\boldsymbol{\theta}\|^2.$$

Using the negative gradient search to minimize $J_2(\hat{\boldsymbol{\theta}})$, we can obtain the following iterative relation:

$$\begin{aligned}\hat{\boldsymbol{\theta}}_k(t) &= \hat{\boldsymbol{\theta}}_{k-1}(t) - \mu(t) \text{grad}[J_2(\hat{\boldsymbol{\theta}}_{k-1}(t))] \\ &= \hat{\boldsymbol{\theta}}_{k-1}(t) + \mu(t) \boldsymbol{\Phi}^T(p, t) [\mathbf{Y}(p, t) - \boldsymbol{\Phi}(p, t) \hat{\boldsymbol{\theta}}_{k-1}(t)] \\ &= [\mathbf{I}_n - \mu(t) \boldsymbol{\Phi}^T(p, t) \boldsymbol{\Phi}(p, t)] \hat{\boldsymbol{\theta}}_{k-1}(t) + \mu(t) \boldsymbol{\Phi}^T(p, t) \mathbf{Y}(p, t),\end{aligned}\quad (13)$$

where $\mu(t) \geq 0$ is the step-size of the time t at iteration k . Similarly, in order to guarantee the convergence of the parameter estimation vector $\hat{\boldsymbol{\theta}}_k(t)$, $\mu(t)$ can be conservatively chosen as

$$\mu(t) \leq \frac{2}{\lambda_{\max}[\boldsymbol{\Phi}^T(p, t) \boldsymbol{\Phi}(p, t)]} = 2\lambda_{\max}^{-1}[\boldsymbol{\Phi}^T(p, t) \boldsymbol{\Phi}(p, t)]. \quad (14)$$

In consideration of the computational cost of the eigenvalue calculation, by using the norm $\mu(t)$ is to take as

$$\mu(t) \leq \frac{2}{\|\boldsymbol{\Phi}(p, t)\|^2} = 2\|\boldsymbol{\Phi}(p, t)\|^{-2}. \quad (15)$$

Equations (12)–(15) and (7)–(9) form the multi-innovation gradient-based iterative (MIGI) algorithm for estimating the parameter vector $\boldsymbol{\theta}$ [4]:

$$\hat{\boldsymbol{\theta}}_k(t) = \hat{\boldsymbol{\theta}}_{k-1}(t) + \mu(t) \boldsymbol{\Phi}^T(p, t) [\mathbf{Y}(p, t) - \boldsymbol{\Phi}(p, t) \hat{\boldsymbol{\theta}}_{k-1}(t)], \quad k = 1, 2, 3, \dots \quad (16)$$

$$\mu(t) \leq 2\lambda_{\max}^{-1}[\boldsymbol{\Phi}^T(p, t) \boldsymbol{\Phi}(p, t)], \quad \text{or} \quad \mu(t) \leq 2\|\boldsymbol{\Phi}(p, t)\|^{-2}, \quad (17)$$

$$\mathbf{Y}(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (18)$$

$$\boldsymbol{\Phi}(p, t) = [\boldsymbol{\varphi}(t), \boldsymbol{\varphi}(t-1), \dots, \boldsymbol{\varphi}(t-p+1)]^T, \quad (19)$$

$$\boldsymbol{\varphi}(t) = \begin{bmatrix} \boldsymbol{\varphi}_a(t) \\ \boldsymbol{\varphi}_b(t) \end{bmatrix}, \quad (20)$$

$$\boldsymbol{\varphi}_a(t) = [-y(t-1), -y(t-2), \dots, -y(t-n_a)]^T, \quad (21)$$

$$\boldsymbol{\varphi}_b(t) = [u(t-1), u(t-2), \dots, u(t-n_b)]^T. \quad (22)$$

The identification steps of the MIGI algorithm in Equations (16)–(22) for computing $\hat{\boldsymbol{\theta}}_k(t)$ are listed as follows.

1. For $t \leq 0$, all the variables are set to zero. Let $t = 1$, give the data length p ($p \gg n$) and set the initial values: $\hat{\boldsymbol{\theta}}_0(t) = \mathbf{1}_n/p_0$, $p_0 = 10^6$, the maximum iteration k_{\max} and the accuracy ε .
2. Let $k = 1$, collect the input and output data $u(t)$ and $y(t)$.
3. Form the information vectors $\boldsymbol{\varphi}_a(t)$, $\boldsymbol{\varphi}_b(t)$ and $\boldsymbol{\varphi}(t)$ using Equations (21)–(22) and (20).
4. Construct the stacked output vector $\mathbf{Y}(p, t)$ by Equation (18) and the stacked information matrix $\boldsymbol{\Phi}(p, t)$ by Equation (19), select a large $\mu(t)$ according to Equation (17).
5. Update the parameter vector estimate $\hat{\boldsymbol{\theta}}_k(t)$ using Equation (16).
6. If $k < k_{\max}$, increase k by 1 and go to Step 5; otherwise, proceed with the next step.
7. Compare $\hat{\boldsymbol{\theta}}_k(t)$ with $\hat{\boldsymbol{\theta}}_{k-1}(t)$: If $\|\hat{\boldsymbol{\theta}}_k(t) - \hat{\boldsymbol{\theta}}_{k-1}(t)\| > \varepsilon$, set $\hat{\boldsymbol{\theta}}_0(t+1) := \hat{\boldsymbol{\theta}}_k(t)$ and increase k by 1 and go to step 2; otherwise, obtain the parameter estimation vector $\hat{\boldsymbol{\theta}}_k(t)$.

The flowchart of computing the parameter estimation vector $\hat{\boldsymbol{\theta}}_k(t)$ is shown in Figure 2.

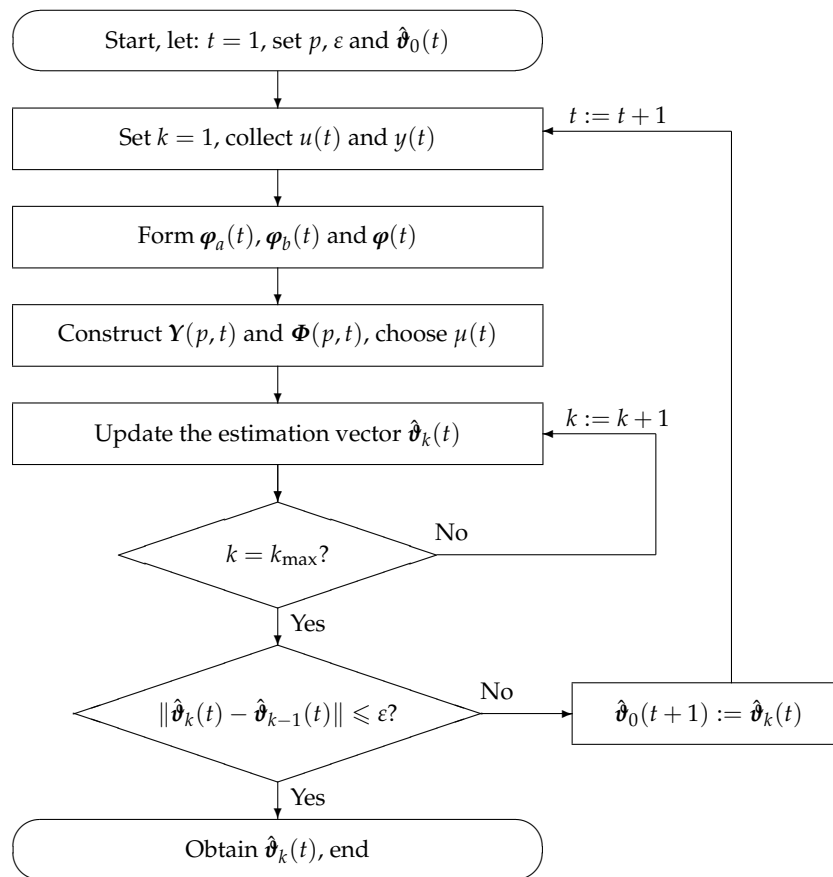


Figure 2. The flowchart of the multi-innovation gradient-based iterative (MIGI) algorithm.

Remark 5. The multi-innovation gradient identification method can improve the parameter estimation accuracy by expanding the dimension of the innovation and making full use of the system information. The MIGI algorithm in Equations (16)–(22) can be seen as an application of the multi-innovation theory in the iterative identification. In particular, we use the newest p data from $j = t - p + 1$ to $j = t$ to do the iterative estimation. When $k = k_{\max}$, move the data window forward to the next moment, introduce new observation data and eliminate the oldest data to keep p data in the data window. The MISG algorithm enhances the utilization of data and improves the parameter estimation accuracy.

Remark 6. When take $p = t = L$, the MIGI algorithm reduces to the GI algorithm. That is the MIGI algorithm is the extension of the GI algorithm, or the GI algorithm is the special case of the MIGI algorithm.

Furthermore, the proposed approaches in the paper can combine other mathematical tools [50–52] and statistical strategies [53–60] to study the parameter estimation problems of linear and nonlinear systems with different disturbance noises [61–64], and can be applied to other literature [65–68] such as signal processing [69–74] and neural networks [75–77].

5. Example

In industrial processes, some control task can be simplified into a water tank control plant, as is shown in Figure 3 [78].

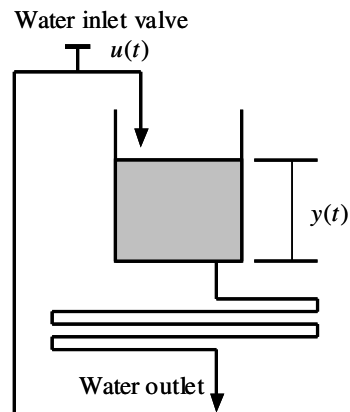


Figure 3. A water tank plant.

In this system, the manipulated variable is the position of the inlet water valve, denoted $u(t)$; the measured variable is the water level in the tank, denoted $y(t)$. For a two-level water tank plant, it can be described by a second-order controlled autoregressive system, we assume that its mathematical model is given by

$$\begin{aligned} A(z)y(t) &= B(z)u(t) + v(t), \\ A(z) &= 1 + a_1z^{-1} + a_2z^{-2} = 1 + 1.35z^{-1} + 0.75z^{-2}, \\ B(z) &= b_1z^{-1} + b_2z^{-2} = 1.68z^{-1} + 2.32z^{-2}. \end{aligned}$$

The parameter vector to be identified is given by

$$\boldsymbol{\vartheta} = [a_1, a_2, b_1, b_2]^T = [1.35, 0.75, 1.68, 2.32]^T.$$

In simulation, the input $\{u(t)\}$ is taken as an uncorrelated uniform distribution random signal sequence with zero mean and unit variance, $\{v(t)\}$ is taken as a normal distribution white noise sequence with zero mean and variance σ^2 . Using the simulation model parameters and the input signal generates the output signal $\{y(t)\}$.

Taking the noise variances $\sigma^2 = 0.20^2$, $\sigma^2 = 1.00^2$ and $\sigma^2 = 2.00^2$, respectively, the corresponding noise-to-signal ratios were $\delta_{\text{ns}} = 13.51\%$, $\delta_{\text{ns}} = 67.55\%$, and $\delta_{\text{ns}} = 135.11\%$. Setting the data length $L = 3000$, the convergence factor $\mu = \lambda_{\max}^{-1}[\boldsymbol{\Phi}^T(L)\boldsymbol{\Phi}(L)]$, applying the GI algorithm and the input-output data $\{u(t), y(t)\}$ to estimate the parameters of this example system, the GI parameter estimates and errors versus with different noise variances are shown in Tables 2–4, the GI estimation errors $\delta := \|\hat{\boldsymbol{\vartheta}}(t) - \boldsymbol{\vartheta}\| / \|\boldsymbol{\vartheta}\|$ versus k are shown in Figure 4 with $\sigma^2 = 0.20^2$, $\sigma^2 = 1.00^2$ and $\sigma^2 = 2.00^2$, and the GI estimates of the parameters a_1 , a_2 , b_1 and b_2 versus k are shown in Figure 5 for $\sigma^2 = 0.20^2$.

Table 2. The GI estimates and their errors with $\sigma^2 = 0.20^2$.

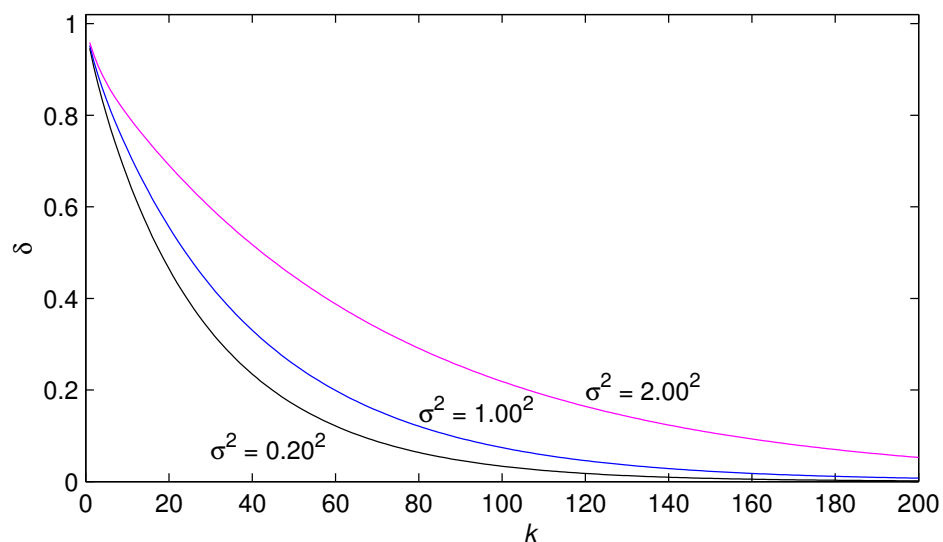
k	a_1	a_2	b_1	b_2	δ (%)
1	0.36787	−0.03138	0.08512	0.00447	94.61743
2	0.49364	0.09219	0.16729	0.03988	90.39550
5	0.72013	0.30680	0.39043	0.18999	80.09229
10	0.88253	0.44424	0.69306	0.47742	66.48636
20	1.02561	0.54320	1.10288	0.97583	46.48051
50	1.22656	0.67165	1.56551	1.80449	16.83784
100	1.32469	0.73450	1.67354	2.21545	3.34572
150	1.34439	0.74715	1.68095	2.29850	0.68915
200	1.34836	0.74970	1.68146	2.31527	0.16042
True values	1.35000	0.75000	1.68000	2.32000	−

Table 3. The GI estimates and their errors with $\sigma^2 = 1.00^2$.

k	a_1	a_2	b_1	b_2	δ (%)
1	0.39074	−0.07246	0.05593	0.00225	95.24273
2	0.51599	0.05176	0.11093	0.02612	91.71095
5	0.75720	0.28762	0.26578	0.12989	83.37742
10	0.93536	0.45376	0.49189	0.34387	72.56593
20	1.05985	0.55458	0.84302	0.75447	55.60085
50	1.20990	0.65849	1.39092	1.55649	25.60909
100	1.30606	0.72444	1.63629	2.08855	7.40476
150	1.33488	0.74426	1.67889	2.24897	2.23868
200	1.34356	0.75024	1.68613	2.29741	0.74605
True values	1.35000	0.75000	1.68000	2.32000	-

Table 4. The GI estimates and their errors with $\sigma^2 = 2.00^2$.

k	a_1	a_2	b_1	b_2	δ (%)
1	0.41321	−0.11502	0.02712	0.00069	95.88923
2	0.53665	0.00817	0.05426	0.01247	93.10387
5	0.79437	0.26457	0.13329	0.06494	87.03542
10	1.00553	0.47258	0.25709	0.18223	80.07592
20	1.13655	0.59653	0.47660	0.43831	69.11109
50	1.21869	0.66123	0.95452	1.06581	44.78973
100	1.28208	0.70753	1.37278	1.68366	21.85326
150	1.31416	0.73098	1.55522	1.99714	10.70952
200	1.33042	0.74288	1.63461	2.15631	5.25922
True values	1.35000	0.75000	1.68000	2.32000	-

**Figure 4.** The GI estimation errors δ versus k with $\sigma^2 = 0.20^2$, $\sigma^2 = 1.00^2$ and $\sigma^2 = 2.00^2$.

From Tables 2–4 and Figures 4 and 5, we can draw the following conclusions.

- As the noise levels decrease, the GI algorithm can give more accurate parameter estimates—see the parameter estimation errors in the last columns in Tables 2–4 and the parameter estimation error curves in Figure 4 under different noise variances.
- The GI parameter estimates approach to their true values for sufficiently large data length as the iteration k increases—see Figure 5.

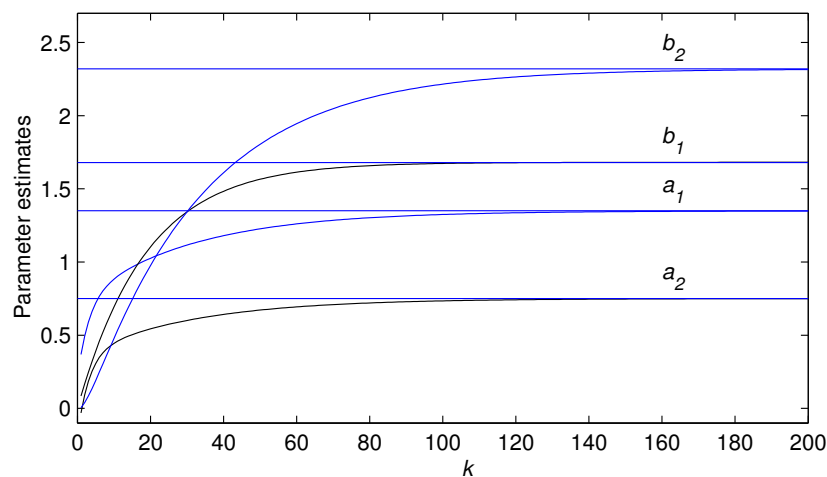


Figure 5. The GI estimates versus k with $\sigma^2 = 0.20^2$.

6. Conclusions

Modeling a dynamical system is the first step for system analysis and design in control engineering. Through defining and minimizing a quadratic criterion function by using observation data, this paper studies and derives a gradient-based iterative algorithm for stochastic systems described by the controlled autoregressive model based on the gradient search. Furthermore, in order to track time-varying parameters, a multi-innovation gradient-based iterative algorithm has been proposed by means of the multi-innovation identification theory. The proposed algorithms have the following advantages.

- For a lower noise level, the gradient-based algorithm can give more accurate parameter estimates. The parameter estimation errors become smaller as the iterative index increases.
- The gradient-based iterative parameter estimates approach their true values for sufficiently large data length and iterative index.
- The multi-innovation gradient-based iterative algorithm can track time-varying parameters of dynamical systems, improving the performance of the algorithms.
- The simulation results indicate that the proposed algorithms are effective for estimating the parameters of stochastic systems.
- The proposed methods in this paper can be extended to model industrial processes and network systems [79–84] by means of some other mathematical tools and approaches [85–90].

Author Contributions: Conceptualization and methodology, F.D.; software, J.P. and F.D.; validation and analysis, A.A. and T.H.

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

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

References

1. Ding, F. *System Identification—New Theory and Methods*; Science Press: Beijing, China, 2013. Available online: <http://www.bookask.com/book/631454.html> (accessed on 5 April 2019). (In Chinese)
2. Ding, F. *System Identification—Performances Analysis for Identification Methods*; Science Press: Beijing, China, 2014. Available online: http://product.dangdang.com/23511899.html?ddclick_reco_product_alsobuy (accessed on 5 April 2019). (In Chinese)

3. Ding, F. *System Identification—Auxiliary Model Identification Idea and Methods*; Science Press: Beijing, China, 2017. Available online: <http://www.zxhsd.com/kgsm/ts/2017/07/07/3878821.shtml> (accessed on 5 April 2019). (In Chinese)
4. Ding, F. *System Identification—Iterative Search Principle and Identification Methods*; Science Press: Beijing, China, 2018. Available online: <https://item.jd.com/12438606.html> (accessed on 5 April 2019). (In Chinese)
5. Ding, F. *System Identification—Multi-Innovation Identification Theory and Methods*; Science Press: Beijing, China, 2016. Available online: <http://product.dangdang.com/23933240.html> (accessed on 5 April 2019). (In Chinese)
6. Xu, L. The damping iterative parameter identification method for dynamical systems based on the sine signal measurement. *Signal Process.* **2016**, *120*, 660–667. [[CrossRef](#)]
7. Xu, L. A proportional differential control method for a time-delay system using the Taylor expansion approximation. *Appl. Math. Comput.* **2014**, *236*, 391–399. [[CrossRef](#)]
8. Xu, L.; Ding, F. Parameter estimation for control systems based on impulse responses. *Int. J. Control Autom. Syst.* **2017**, *15*, 2471–2479. [[CrossRef](#)]
9. Li, M.H.; Liu, X.M. The least squares based iterative algorithms for parameter estimation of a bilinear system with autoregressive noise using the data filtering technique. *Signal Process.* **2018**, *147*, 23–34. [[CrossRef](#)]
10. Xu, L.; Chen, L.; Xiong, W.L. Parameter estimation and controller design for dynamic systems from the step responses based on the Newton iteration. *Nonlinear Dyn.* **2015**, *79*, 2155–2163. [[CrossRef](#)]
11. Xu, L. The parameter estimation algorithms based on the dynamical response measurement data. *Adv. Mech. Eng.* **2017**, *9*, 1–12. [[CrossRef](#)]
12. Tian, X.P.; Niu, H.M. A bi-objective model with sequential search algorithm for optimizing network-wide train timetables. *Comput. Ind. Eng.* **2019**, *127*, 1259–1272. [[CrossRef](#)]
13. Yang, F.; Zhang, P.; Li, X.X. The truncation method for the Cauchy problem of the inhomogeneous Helmholtz equation. *Appl. Anal.* **2019**, *98*, 991–1004. [[CrossRef](#)]
14. Zhao, N.; Liu, R.; Chen, Y.; Wu, M.; Jiang, Y.; Xiong, W.; Liu, C. Contract design for relay incentive mechanism under dual asymmetric information in cooperative networks. *Wirel. Netw.* **2018**, *24*, 3029–3044. [[CrossRef](#)]
15. Xu, G.H.; Shekofteh, Y.; Akgul, A.; Li, C.B.; Panahi, S. A new chaotic system with a self-excited attractor: Entropy measurement, signal encryption, and parameter estimation. *Entropy* **2018**, *20*, 86. [[CrossRef](#)]
16. Li, X.Y.; Li, H.X.; Wu, B.Y. Piecewise reproducing kernel method for linear impulsive delay differential equations with piecewise constant arguments. *Appl. Math. Comput.* **2019**, *349*, 304–313. [[CrossRef](#)]
17. Noshadi, A.; Shi, J.; Lee, W.S.; Shi, P.; Kalam, A. System identification and robust control of multi-input multi-output active magnetic bearing systems. *IEEE Trans. Control. Syst. Technol.* **2016**, *24*, 1227–1239. [[CrossRef](#)]
18. Pan, J.; Li, W.; Zhang, H.P. Control algorithms of magnetic suspension systems based on the improved double exponential reaching law of sliding mode control. *Int. J. Control Autom. Syst.* **2018**, *16*, 2878–2887. [[CrossRef](#)]
19. Zhang, X.; Ding, F.; Xu, L.; Yang, E.F. Highly computationally efficient state filter based on the delta operator. *Int. J. Adapt. Control Signal Process.* **2019**. [[CrossRef](#)]
20. Luo, X.S.; Song, Y.D. Data-driven predictive control of Hammerstein-Wiener systems based on subspace identification. *Inf. Sci.* **2018**, *422*, 447–461. [[CrossRef](#)]
21. Ma, F.Y.; Yin, Y.K.; Li, M. Start-up process modelling of sediment microbial fuel cells based on data driven. *Math. Probl. Eng.* **2019**, *2019*, 7403732. [[CrossRef](#)]
22. Li, M.H.; Liu, X.M. Auxiliary model based least squares iterative algorithms for parameter estimation of bilinear systems using interval-varying measurements. *IEEE Access* **2018**, *6*, 21518–21529. [[CrossRef](#)]
23. Bottegal, G.; Castro-Garcia, R.; Suykens, J.A.K. A two-experiment approach to Wiener system identification. *Automatica* **2018**, *93*, 282–289. [[CrossRef](#)]
24. Guo, F.; Hariprasad, K.; Huang, B.; Ding, Y.S. Robust identification for nonlinear errors-in-variables systems using the EM algorithm. *J. Process. Control.* **2017**, *54*, 129–137. [[CrossRef](#)]
25. Li, M.H.; Liu, X.M.; Ding, F. Filtering-based maximum likelihood gradient iterative estimation algorithm for bilinear systems with autoregressive moving average noise. *Circuits Syst. Signal Process.* **2019**, *37*, 5023–5048. [[CrossRef](#)]

26. Zhang, X.; Ding, F.; Xu, L.; Yang, E.F. State filtering-based least squares parameter estimation for bilinear systems using the hierarchical identification principle. *IET Control Theory Appl.* **2018**, *12*, 1704–1713. [[CrossRef](#)]
27. Ding, F. Two-stage least squares based iterative estimation algorithm for CARARMA system modeling. *Appl. Math. Model.* **2013**, *37*, 4798–4808. [[CrossRef](#)]
28. Xu, H.; Ding, F.; Yang, E.F. Modeling a nonlinear process using the exponential autoregressive time series model. *Nonlinear Dyn.* **2019**, *95*, 2079–2092. [[CrossRef](#)]
29. Ding, F. Decomposition based fast least squares algorithm for output error systems. *Signal Process.* **2013**, *93*, 1235–1242. [[CrossRef](#)]
30. Ge, Z.W.; Ding, F.; Xu, L.; Alsaedi, A.; Hayat, T. Gradient-based iterative identification method for multivariate equation-error autoregressive moving average systems using the decomposition technique. *J. Frankl. Inst.* **2019**, *356*, 1658–1676. [[CrossRef](#)]
31. Pan, J.; Ma, H.; Jiang, X.; Ding, W.; Ding, F. Adaptive gradient-based iterative algorithm for multivariate controlled autoregressive moving average systems using the data filtering technique. *Complexity* **2018**, *2018*, 9598307. [[CrossRef](#)]
32. Xu, L. Application of the Newton iteration algorithm to the parameter estimation for dynamical systems. *J. Comput. Appl. Math.* **2015**, *288*, 33–43. [[CrossRef](#)]
33. Wang, Y.J.; Ding, F. Iterative estimation for a non-linear IIR filter with moving average noise by means of the data filtering technique. *IMA J. Math. Control Inf.* **2017**, *34*, 745–764. [[CrossRef](#)]
34. Liu, Y.J.; Wang, D.Q.; Ding, F. Least squares based iterative algorithms for identifying Box-Jenkins models with finite measurement data. *Digit. Signal Process.* **2010**, *20*, 1458–1467. [[CrossRef](#)]
35. Liu, Q.Y.; Ding, F. Auxiliary model-based recursive generalized least squares algorithm for multivariate output-error autoregressive systems using the data filtering. *Circuits Syst. Signal Process.* **2019**, *38*, 590–610. [[CrossRef](#)]
36. Xu, L.; Ding, F.; Gu, Y.; Alsaedi, A.; Hayat, T. A multi-innovation state and parameter estimation algorithm for a state space system with d-step state-delay. *Signal Process.* **2017**, *140*, 97–103. [[CrossRef](#)]
37. Xu, L.; Ding, F. Iterative parameter estimation for signal models based on measured data. *Circuits Syst. Signal Process.* **2018**, *37*, 3046–3069. [[CrossRef](#)]
38. Xu, L.; Xiong, W.L.; Alsaedi, A.; Hayat, T. Hierarchical parameter estimation for the frequency response based on the dynamical window data. *Int. J. Control Autom. Syst.* **2018**, *16*, 1756–1764. [[CrossRef](#)]
39. Ding, F.; Liu, X.P.; Liu, G. Gradient based and least-squares based iterative identification methods for OE and OEMA systems. *Digit. Signal Process.* **2010**, *20*, 664–677. [[CrossRef](#)]
40. Xu, L.; Ding, F. Recursive least squares and multi-innovation stochastic gradient parameter estimation methods for signal modeling. *Circuits Syst. Signal Process.* **2017**, *36*, 1735–1753. [[CrossRef](#)]
41. Pan, J.; Jiang, X.; Wan, X.K.; Ding, W. A filtering based multi-innovation extended stochastic gradient algorithm for multivariable control systems. *Int. J. Control. Autom. Syst.* **2017**, *15*, 1189–1197. [[CrossRef](#)]
42. Zhang, X.; Xu, L.; Ding, F.; Hayat, T. Combined state and parameter estimation for a bilinear state space system with moving average noise. *J. Frankl. Inst.* **2018**, *355*, 3079–3103. [[CrossRef](#)]
43. Zhang, X.; Ding, F.; Alsaadi, F.E.; Hayat, T. Recursive parameter identification of the dynamical models for bilinear state space systems. *Nonlinear Dyn.* **2017**, *89*, 2415–2429. [[CrossRef](#)]
44. Bonami, P.; Biegler, L.T.; Conn, A.R.; Cornuéjols, G.; Grossmann, I.E.; Lairde, C.D.; Lee, J.; Lodi, A.; Margot, F.; Sawaya, N.; et al. An algorithmic framework for convex mixed integer nonlinear programs. *Discret. Optim.* **2008**, *5*, 186–204. [[CrossRef](#)]
45. Neumaier, A.; Shcherbina, O.; Huyer, W.; Vinkó, T. A comparison of complete global optimization solvers. *Math. Program.* **2005**, *103*, 335–356. [[CrossRef](#)]
46. Lastusilta, T.; Bussieck, M.R.; Westerlund, T. An experimental study of the GAMS/AlphaECP MINLP solver. *Ind. Eng. Chem. Res.* **2009**, *48*, 7337–7345. [[CrossRef](#)]
47. Klanšek, U. A comparison between MILP and MINLP approaches to optimal solution of nonlinear discrete transportation problem. *Transport* **2015**, *30*, 135–144. [[CrossRef](#)]
48. Ding, F.; Liu, X.G.; Chu, J. Gradient-based and least-squares-based iterative algorithms for Hammerstein systems using the hierarchical identification principle. *IET Control Theory Appl.* **2013**, *7*, 176–184. [[CrossRef](#)]
49. Xu, L.; Ding, F. Parameter estimation algorithms for dynamical response signals based on the multi-innovation theory and the hierarchical principle. *IET Signal Process.* **2017**, *11*, 228–237. [[CrossRef](#)]

50. Wu, M.H.; Li, X.; Liu, C.; Liu, M.; Zhao, N.; Wang, J.; Wan, X.K.; Rao, Z.H.; Zhu, L. Robust global motion estimation for video security based on improved k-means clustering. *J. AmbientIntell. Humaniz. Comput.* **2019**, *10*, 439–448. [\[CrossRef\]](#)
51. Wan, X.K.; Wu, H.; Qiao, F.; Li, F.; Li, Y.; Yan, Y.; Wei, J. Electrocardiogram baseline wander suppression based on the combination of morphological and wavelet transformation based filtering. *Computat. Math. Methods Med.* **2019**, *2019*, 7196156. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Wen, Y.Z.; Yin, C.C. Solution of Hamilton-Jacobi-Bellman equation in optimal reinsurance strategy under dynamic VaR constraint. *J. Funct. Spaces* **2019**, *2019*, 6750892. [\[CrossRef\]](#)
53. Yin, C.C.; Zhao, J.S. Nonexponential asymptotics for the solutions of renewal equations, with applications. *J. Appl. Probab.* **2006**, *43*, 815–824. [\[CrossRef\]](#)
54. Yin, C.C.; Wang, C.W. The perturbed compound Poisson risk process with investment and debit interest. *Methodol. Comput. Appl. Probab.* **2010**, *12*, 391–413. [\[CrossRef\]](#)
55. Yin, C.C.; Yuen, K.C. Optimality of the threshold dividend strategy for the compound Poisson model. *Stat. Probab. Lett.* **2011**, *81*, 1841–1846. [\[CrossRef\]](#)
56. Yin, C.C.; Wen, Y.Z. Exit problems for jump processes with applications to dividend problems. *J. Comput. Appl. Math.* **2013**, *245*, 30–52. [\[CrossRef\]](#)
57. Yin, C.C.; Wen, Y.Z. Optimal dividend problem with a terminal value for spectrally positive Levy processes. *Insur. Math. Econ.* **2013**, *53*, 769–773. [\[CrossRef\]](#)
58. Yin, C.C.; Wen, Y.Z.; Zhao, Y.X. On the optimal dividend problem for a spectrally positive levy process. *Astin Bull.* **2014**, *44*, 635–651. [\[CrossRef\]](#)
59. Yin, C.C.; Yuen, K.C. Exact joint laws associated with spectrally negative Levy processes and applications to insurance risk theory. *Front. Math. China* **2014**, *9*, 1453–1471. [\[CrossRef\]](#)
60. Yin, C.C.; Yuen, K.C. Optimal dividend problems for a jump-diffusion model with capital injections and proportional transaction costs. *J. Ind. Manag. Optim.* **2015**, *11*, 1247–1262. [\[CrossRef\]](#)
61. Zhang, X.; Ding, F.; Xu, L.; Alsaedi, A.; Hayat, T. A hierarchical approach for joint parameter and state estimation of a bilinear system with autoregressive noise. *Mathematics* **2019**, *7*. [\[CrossRef\]](#)
62. Xu, L.; Ding, F.; Zhu, Q.M. Hierarchical Newton and least squares iterative estimation algorithm for dynamic systems by transfer functions based on the impulse responses. *Int. J. Syst. Sci.* **2019**, *50*, 141–151. [\[CrossRef\]](#)
63. Wang, Y.J.; Ding, F.; Xu, L. Some new results of designing an IIR filter with colored noise for signal processing. *Digit. Signal Process.* **2018**, *72*, 44–58. [\[CrossRef\]](#)
64. Wang, Y.J.; Ding, F.; Wu, M.H. Recursive parameter estimation algorithm for multivariate output-error systems. *J. Frankl. Inst.* **2018**, *355*, 5163–5181. [\[CrossRef\]](#)
65. Sun, Z.Y.; Zhang, D.; Meng, Q.; Chen, C.C. Feedback stabilization of time-delay nonlinear systems with continuous time-varying output function. *Int. J. Syst. Sci.* **2019**, *50*, 244–255. [\[CrossRef\]](#)
66. Zhan, X.S.; Cheng, L.L.; Wu, J.; Yang, Q.S.; Han, T. Optimal modified performance of MIMO networked control systems with multi-parameter constraints. *ISA Trans.* **2019**, *84*, 111–117. [\[CrossRef\]](#)
67. Zha, X.S.; Guan, Z.H.; Zhang, X.H.; Yuan, F.S. Optimal tracking performance and design of networked control systems with packet dropout. *J. Frankl. Inst.* **2013**, *350*, 3205–3216.
68. Jiang, C.M.; Zhang, F.F.; Li, T.X. Synchronization and antisynchronization of N-coupled fractional-order complex chaotic systems with ring connection. *Math. Methods Appl. Sci.* **2018**, *41*, 2625–2638. [\[CrossRef\]](#)
69. Wang, T.; Liu, L.; Zhang, J.; Schaeffer, E.; Wang, Y. A M-EKF fault detection strategy of insulation system for marine current turbine. *Mech. Syst. Signal Process.* **2019**, *115*, 269–280. [\[CrossRef\]](#)
70. Cao, Y.; Lu, H.; Wen, T. A safety computer system based on multi-sensor data processing. *Sensors* **2019**, *19*, 818. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Cao, Y.; Zhang, Y.; Wen, T.; Li, P. Research on dynamic nonlinear input prediction of fault diagnosis based on fractional differential operator equation in high-speed train control system. *Chaos* **2019**, *29*, 013130. [\[CrossRef\]](#) [\[PubMed\]](#)
72. Cao, Y.; Li, P.; Zhang, Y. Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing. *Future Gener. Comput. Syst.* **2018**, *88*, 279–283. [\[CrossRef\]](#)
73. Cao, Y.; Ma, L.C.; Xiao, S.; Zhang, X.; Xu, W. Standard analysis for transfer delay in CTCS-3. *Chin. J. Electron.* **2017**, *26*, 1057–1063. [\[CrossRef\]](#)
74. Zhao, N.; Wu, M.H.; Chen, J.J. Android-based mobile educational platform for speech signal processing. *Int. J. Electr. Eng. Edu.* **2017**, *54*, 3–16. [\[CrossRef\]](#)

75. Zhao, N.; Chen, Y.; Liu, R.; Wu, M.H.; Xiong, W. Monitoring strategy for relay incentive mechanism in cooperative communication networks. *Comput. Electr. Eng.* **2017**, *60*, 14–29. [[CrossRef](#)]
76. Ji, Y.; Ding, F. Multiperiodicity and exponential attractivity of neural networks with mixed delays. *Circuits Syst. Signal Process.* **2017**, *36*, 2558–2573. [[CrossRef](#)]
77. Ji, Y.; Liu, X.M. Unified synchronization criteria for hybrid switching-impulsive dynamical networks. *Circuits Syst. Signal Process.* **2015**, *34*, 1499–1517. [[CrossRef](#)]
78. Ding, F.; Chen, T. Performance bounds of the forgetting factor least squares algorithm for time-varying systems with finite measurement data. *IEEE Trans. Circuits Syst. Regul. Pap.* **2005**, *52*, 555–566. [[CrossRef](#)]
79. Li, N.; Guo, S.; Wang, Y. Weighted preliminary-summation-based principal component analysis for non-Gaussian processes. *Control. Eng. Pract.* **2019**. [[CrossRef](#)]
80. Wang, Y.; Si, Y.; Huang, B.; Lou, Z. Survey on the theoretical research and engineering applications of multivariate statistics process monitoring algorithms: 2008–2017. *Can. J. Chem.* **2018**, *96*, 2073–2085. [[CrossRef](#)]
81. Feng, L.; Li, Q.X.; Li, Y.F. Imaging with 3-D aperture synthesis radiometers. *IEEE Trans. Geosci. Remote. Sens.* **2019**, *57*, 2395–2406. [[CrossRef](#)]
82. Shi, W.X.; Liu, N.; Zhou, Y.M.; Cao, X.A. Effects of postannealing on the characteristics and reliability of polyfluorene organic light-emitting diodes. *IEEE Trans. Electron. Devices* **2019**, *66*, 1057–1062. [[CrossRef](#)]
83. Fu, B.; Ouyang, C.X.; Li, C.S.; Wang, J.W.; Gul, E. An improved mixed integer linear programming approach based on symmetry diminishing for unit commitment of hybrid power system. *Energies* **2019**, *12*, 833. [[CrossRef](#)]
84. Wu, T.Z.; Shi, X.; Liao, L.; Zhou, C.J.; Zhou, H.; Su, Y.H. A capacity configuration control strategy to alleviate power fluctuation of hybrid energy storage system based on improved particle swarm optimization. *Energies* **2019**, *12*, 642. [[CrossRef](#)]
85. Liu, F.; Xue, Q.; Yabuta, K. Boundedness and continuity of maximal singular integrals and maximal functions on Triebel-Lizorkin spaces. *Scie. China-Math.* **2019**. [[CrossRef](#)]
86. Liu, F. Boundedness and continuity of maximal operators associated to polynomial compound curves on Triebel-Lizorkin spaces. *Math. Inequalities Appl.* **2019**, *22*, 25–44. [[CrossRef](#)]
87. Liu, F.; Fu, Z.; Jhang, S. Boundedness and continuity of Marcinkiewicz integrals associated to homogeneous mappings on Triebel-Lizorkin spaces. *Front. Math. China* **2019**, *14*, 95–122. [[CrossRef](#)]
88. Ding, J.L. The hierarchical iterative identification algorithm for multi-input-output-error systems with autoregressive noise. *Complexity* **2017**, *2017*, 1–11. [[CrossRef](#)]
89. Wang, D.Q.; Yan, Y.R.; Liu, Y.J.; Ding, J.H. Model recovery for Hammerstein systems using the hierarchical orthogonal matching pursuit method. *J. Computat. Appl. Math.* **2019**, *345*, 135–145. [[CrossRef](#)]
90. Wang, D.Q.; Li, L.W.; Ji, Y.; Yan, Y.R. Model recovery for Hammerstein systems using the auxiliary model based orthogonal matching pursuit method. *Appl. Math. Modell.* **2018**, *54*, 537–550. [[CrossRef](#)]

