


Article

# Iterative Identification for Multivariable Systems with Time-Delays Based on Basis Pursuit De-Noising and Auxiliary Model

Junyao You <sup>1</sup>  and Yanjun Liu <sup>1,2,\*</sup>

<sup>1</sup> School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China; yjy804521297@163.com

<sup>2</sup> Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122, China

\* Correspondence: yjl@jiangnan.edu.cn

Received: 30 September 2018 ; Accepted: 31 October 2018; Published: 6 November 2018



**Abstract:** This paper focuses on the joint estimation of parameters and time-delays of the multiple-input single-output output-error systems. Since the time-delays are unknown, an effective identification model with a high dimensional and sparse parameter vector is established based on overparameterization. Then, the identification problem is converted to a sparse optimization problem. Based on the basis pursuit de-noising criterion and the auxiliary model identification idea, an auxiliary model based basis pursuit de-noising iterative algorithm is presented. The parameters are estimated by solving a quadratic program, and the unavailable terms in the information vector are updated by the auxiliary model outputs iteratively. The time-delays are estimated according to the sparse structure of the parameter vector. The proposed method can obtain effective estimates of the parameters and time-delays from few sampled data. The simulation results illustrate the effectiveness of the proposed algorithm.

**Keywords:** multivariable system; parameter identification; time-delay estimation; basis pursuit de-noising; auxiliary model; quadratic program

## 1. Introduction

This paper focuses on the identification of the multiple-input single-output output-error systems with unknown time-delays. The investigation is introduced from its background, the formulation of the problem, the literature survey, and the scope and the contributions in this section.

### 1.1. Background

System identification is a process of developing mathematical models of dynamical systems based on observed data [1]. The dynamical system to be modeled can be a mechanical system, an industrial system or other types of system. Many identification methods have been proposed for the specific systems. For example, some identification techniques for mechanical systems can be found in [2–4]. In this paper, we study the identification of multivariable systems because most of the industrial processes can be modeled as multivariable systems.

Multivariable systems are a class of systems which often present significant interactions between the multiple-input and multiple-output channels. A wealth of literature is available on effective identification algorithms for such systems. For example, an expectation-maximization technique was employed for the estimation of the linear time-invariant (LTI) multiple-input multiple-output (MIMO) systems [5]. A matchable-observable linear identification with zero-order oracle filter tuning method was developed for the LTI MIMO systems by comprising the problems of parameter estimation,

filter tuning and model structure selection [6]. The methods proposed in [5,6] are applicable for the MIMO systems without time-delays. In many industry processes, time-delays are usually unavoidable. Therefore, it is of great interest to study the identification of multivariable systems with unknown time-delays.

For the estimation of time-delays in multivariable systems, several methods have been proposed. In [7], a parametric model is first estimated by using generalized orthonormal basis filters. Then, the time-delay estimates are generated by analyzing the simulated step response of the resulting model. However, it is required that the model structure is available a priori. To address this problem, a non-parametric method was proposed in [8]. The central idea is to break up a general  $m \times n$  MIMO system into  $mn$  decoupled single-input single-output systems using partial coherence functions, and then to apply the frequency-domain method to estimate the time-delay of each subsystem. Since the time-delay estimation procedure was repeated for  $mn$  times, this method requires a large amount of computation and is limited for high order systems. In this paper, we attempt to find a simple and efficient method to address the joint estimation of time-delays and parameters of multivariable systems.

### 1.2. Formulation of the Problem of Interest for this Investigation

For simplicity, we consider the identification problem of multiple-input single-output output-error (MISO-OE) systems with unknown input time-delays. The main body of the MISO-OE model is the impulse transfer functions between the inputs and output. The identification challenge lies in that the system output is nonlinear in the polynomial parameters.

For the identification of the multivariable systems with time-delays, it is needed to form an appropriate identification model first. Taking into account the unknown time-delays, a high dimensional identification model is formed with a sparse parameter vector, which contains many zeros [9,10]. For the identification of high dimensional models, traditional identification methods, such as the least squares method and the stochastic gradient method, require many observations as well as heavy computational burden. However, there are many situations in reality where only limited observations are available such as setpoint-operated processes, online estimation and linear time-variant system identification [11]. It has proven that traditional identification is unreliable in these cases and the estimates often tend to be biased or have large variances [12]. Therefore, it is necessary to develop new identification methods and reduce the identification cost.

The compressive sensing theory originally arose in the signal processing community, which enables the reconstruction of sparse or compressible signals under the undersampling condition [13]. In view of the sparse characteristic of the identification model, the estimation of the sparse parameter vector can be addressed as a reconstruction problem of the sparse signal.

### 1.3. Literature Survey

Greedy algorithms and convex optimization approaches are most widely used reconstruction methods. Greedy algorithms are commonly applied for solving minimum  $l_0$  norm problem with advantages of fast speed and easy implement [14]. Convex optimization approaches are a class of global optimization methods, which address minimum  $l_1$  norm problems with high stability and strong applicability [15]. Basis pursuit, basis pursuit de-noising (BPDN) [16], the least absolute shrinkage and selection operator (LASSO) [17], and the least angle regression [18] are typical convex optimization approaches. These approaches can be converted to linear programming or quadratic programming form, and then the optimal solutions can be obtained [19]. Recently, the reconstruction methods have been applied for system identification. For example, the orthogonal matching pursuit algorithm was applied to obtain the joint estimation of parameters and time-delays of the MISO finite impulse response systems and MISO controlled autoregressive systems [20,21]. The compressive sampling matching pursuit algorithm was modified by employing the instrumental variable method to identify a class of closed-loop systems [22]. The LASSO approach is investigated for computing efficient

model structure of the overparameterized nonlinear autoregressive, moving average exogenous systems [23]. In this paper, we aim to identify the MISO-OE systems with time-delays based on a convex optimization approach.

Compared with the systems studied in [20,21], the structure of the MISO-OE systems is more general and complex. Considering that the system output is nonlinear in the polynomial parameters, the MISO-OE system model can be transformed into a linear regression form by introducing intermediate variables. However, the intermediate variables are unmeasurable and make the identification difficult. Fortunately, the auxiliary model identification idea can solve this problem by replacing unmeasured variables with outputs of the auxiliary model which is constructed of measurable variables [24]. Combining the auxiliary model idea with conventional methods, many new identification methods have been developed such as the auxiliary model based stochastic gradient algorithm [25], and the auxiliary model based least squares (AM-RLS) algorithm [26]. It is indicated that the AM-RLS algorithm can obtain unbiased and consistent estimation of parameters under a persistent excitation condition. Inspired by the algorithms in [25,26], we attempt to combine the auxiliary model idea with the convex optimization approach to deal with the joint estimation of parameters and time-delays of MISO-OE systems.

#### 1.4. Scope and Contribution of this Study

This investigation deals with the identification of the MISO-OE systems with unknown time-delays based on convex optimization and auxiliary model. The structure of the MISO-OE systems is different from the systems studied in [20,21], which means that different methods need to be applied to form the identification model. In this paper, the auxiliary model technique is employed to address the nonlinearity of parameters, and an effective identification model with a high dimensional and sparse parameter vector is derived due to the unknown time-delays. In addition, a different identification method is applied here. The methods proposed in [20,21] are modifications of a greedy algorithm. In this paper, a convex optimization approach, the BPDN approach, is modified for identification due to its robustness. By converting the BPDN approach to a quadratic programming form, the parameters are calculated from the optimal solutions, and the unmeasurable terms in the information vector are updated iteratively. The unknown time-delays can be read from the estimated parameter vector. The effectiveness of the proposed algorithm is tested by simulation examples.

#### 1.5. Organization of the Paper

The rest of this paper is organized as follows. Section 2 introduces the MISO-OE systems with time-delays and describes the identification problem. Section 3 presents an auxiliary model-basis pursuit de-noising iterative (AM-BPDNI) algorithm. Section 4 provides a simulation example to show the effectiveness of the AM-BPDNI algorithm. Finally, some summaries are given in Section 5.

## 2. Problem Description

Consider an MISO-OE system

$$y(t) = \sum_{i=1}^r z^{-d_i} \frac{B_i(z)}{A_i(z)} u_i(t) + v(t), \quad (1)$$

where  $u_i(t)$  and  $d_i$  are the input and time-delay of the  $i$ th input channel,  $y(t)$  is the output,  $v(t)$  is a white noise vector with zero mean and variance  $\sigma^2$ , and  $A_i(z)$  and  $B_i(z)$  are the time-invariant polynomials with constant coefficients in the unit backward shift operator  $z^{-1}$  [i.e.,  $u(t)z^{-1} = u(t-1)$ ],

$$\begin{aligned} A_i(z) &:= 1 + a_{i1}z^{-1} + a_{i2}z^{-2} + \dots + a_{in_{ai}}z^{-n_{ai}}, \\ B_i(z) &:= b_{i1}z^{-1} + b_{i2}z^{-2} + \dots + b_{in_{bi}}z^{-n_{bi}}. \end{aligned}$$

Assume that the orders  $n_{ai}$  and  $n_{bi}$  are known,  $y(t) = 0$ ,  $u_i(t) = 0$  and  $v(t) = 0$  for  $t < 0$ . To form the identification model, an intermediate variable is introduced [24],

$$x_i(t) := z^{-d_i} \frac{B_i(z)}{A_i(z)} u_i(t) = [1 - A_i(z)]x_i(t) + z^{-d_i} B_i(z) u_i(t).$$

Since the time-delay  $d_i$  of each input channel is unknown, an overparameterization method is applied by setting a maximum input regression length  $l$  which satisfies  $l \geq \max(d_i + n_{bi})$  [20]. Then,  $x_i(t)$  can be written in an impact form

$$x_i(t) = \boldsymbol{\varphi}_i^T(t) \boldsymbol{\theta}_i,$$

where

$$\boldsymbol{\varphi}_i(t) := [-x_i(t-1), -x_i(t-2), \dots, -x_i(t-n_{ai}), u_i(t-1), \dots, u_i(t-d_i), u_i(t-d_i-1), \dots, u_i(t-d_i-n_{bi}), \dots, u_i(t-l)]^T \in \mathbb{R}^{n_{ai}+l}, \tag{2}$$

$$\boldsymbol{\theta}_i := [a_{i1}, a_{i2}, \dots, a_{in_{ai}}, \underbrace{0, \dots, 0}_{d_i}, b_{i1}, b_{i2}, \dots, b_{in_{bi}}, \underbrace{0, \dots, 0}_{l-d_i-n_{bi}}]^T \in \mathbb{R}^{n_{ai}+l}. \tag{3}$$

The identification model of the system in Equation (1) can be formed as

$$y(t) = \sum_{i=1}^r x_i(t) + v(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\theta} + v(t), \tag{4}$$

where

$$\boldsymbol{\varphi}(t) := [\boldsymbol{\varphi}_1^T(t), \boldsymbol{\varphi}_2^T(t), \dots, \boldsymbol{\varphi}_r^T(t)]^T \in \mathbb{R}^n, \tag{5}$$

$$\boldsymbol{\theta} := [\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, \dots, \boldsymbol{\theta}_r^T]^T \in \mathbb{R}^n, \tag{6}$$

$$n := \sum_{i=1}^r (n_{ai} + l). \tag{7}$$

It can be seen from Equations (3) and (6) that the parameter vector  $\boldsymbol{\theta}$  contains many zeros, therefore  $\boldsymbol{\theta}$  is a sparse vector and the system in Equation (4) is a sparse system. The sparsity level can be measured by  $K := \sum_{i=1}^r (n_{ai} + n_{bi})$ , which denotes the number of non-zero elements in  $\boldsymbol{\theta}$ . The identification objective is to estimate the unknown parameters  $a_{ij}$ ,  $b_{ij}$  as well as the time-delays  $d_i$  from observations.

### 3. Identification Algorithm

If we have  $m$  observations from  $t = 1$  to  $t = m$ , Equation (4) can be written in a stacked form

$$\mathbf{Y} = \Phi \boldsymbol{\theta} + \mathbf{V}, \tag{8}$$

where

$$\mathbf{Y} := [y(1), y(2), \dots, y(m)]^T \in \mathbb{R}^m, \tag{9}$$

$$\Phi := [\boldsymbol{\varphi}(1), \boldsymbol{\varphi}(2), \dots, \boldsymbol{\varphi}(m)]^T \in \mathbb{R}^{m \times n}, \tag{10}$$

$$\mathbf{V} := [v(1), v(2), \dots, v(m)]^T \in \mathbb{R}^m.$$

From Equations (2), (5) and (10), we can see that the information matrix  $\Phi$  contains many unknown intermediate terms. Therefore, it is difficult to perform the identification directly. According to the auxiliary model identification idea [24,27], the information matrix  $\Phi$  can be replaced with its estimate

$$\hat{\Phi}_k := [\hat{\varphi}_k(1), \hat{\varphi}_k(2), \dots, \hat{\varphi}_k(m)]^T, \tag{11}$$

where

$$\hat{\varphi}_k(t) := [\hat{\varphi}_{1,k}^T(t), \hat{\varphi}_{2,k}^T(t), \dots, \hat{\varphi}_{r,k}^T(t)]^T, \tag{12}$$

$$\hat{\varphi}_{i,k}(t) := [-\hat{x}_{i,k-1}(t-1), \dots, -\hat{x}_{i,k-1}(t-n_{ai}), u_i(t-1), \dots, u_i(t-l)]^T. \tag{13}$$

Note that the unmeasurable terms  $x_i(t-j)$  are replaced with their auxiliary model output estimates  $\hat{x}_i(t-j)$ . Then, the parameter vector  $\theta$  can be estimated by the auxiliary model based least squares iterative (AM-LSI) algorithm [28],

$$\hat{\theta}_k = [\hat{\Phi}_k^T \hat{\Phi}_k]^{-1} \hat{\Phi}_k^T Y, \quad k = 1, 2, 3, \dots \tag{14}$$

$$\hat{x}_{1,k}(t) = \hat{\varphi}_{1,k}^T(t) \hat{\theta}_k(1 : n_{a1} + l), \tag{15}$$

$$\hat{x}_{i,k}(t) = \hat{\varphi}_{i,k}^T(t) \hat{\theta}_k \left( 1 + \sum_{j=1}^{i-1} (n_{aj} + l) : \sum_{j=1}^i (n_{aj} + l) \right), \quad i = 2, 3, \dots, r, \tag{16}$$

where  $\hat{\theta}_k$  denotes the parameter vector estimate at the  $k$ th iteration. According to the least squares (LS) theory, the AM-LSI algorithm is efficient if it is satisfied that  $m \gg n$ . However, from Equation (7), we can see that the dimension of the system in Equation (8) is high. Therefore, it would take a lot of time and efforts to obtain enough observations to meet the identification requirement. Moreover, Equation (14) shows that the AM-LSI algorithm requires computing the inverse matrix  $[\hat{\Phi}_k^T \hat{\Phi}_k]^{-1}$  at each iteration, which leads to a heavy computational burden. Furthermore, the sparse solution cannot be obtained [23], and the time-delays cannot be effectively estimated. Thus, the AM-LSI algorithm is infeasible for the high dimensional and sparse system identification.

Inspired by the compressive sensing theory, the identification of the sparse system in Equation (8) can be further expressed as an optimization problem

$$\hat{\theta} = \arg \min \|\theta\|_0, \quad \text{s.t.} \quad \|Y - \Phi\theta\|_2^2 \leq \varepsilon,$$

where  $\|\cdot\|_0$  represents the  $l_0$  norm,  $\|\cdot\|_2$  the  $l_2$  norm, and  $\varepsilon$  the error tolerance, which is a priori chosen. However, it is a non-convex problem and is difficult to solve in practice. A commonly used alternative is to replace the  $l_0$  norm with a relaxed convex  $l_1$  norm [9],

$$\hat{\theta} = \arg \min \|\theta\|_1, \quad \text{s.t.} \quad \|Y - \Phi\theta\|_2^2 \leq \varepsilon, \tag{17}$$

where  $\|\cdot\|_1$  is the  $l_1$  norm. It is proved that the minimization of  $l_1$  norm is equivalent to the minimization of  $l_0$  norm when the Restricted Isometry Property (RIP) is satisfied [13]. Then, the convex optimization problem in Equation (17) is solvable with the BPDN criterion [16]. The BPDN approach can effectively reduce the interference of noise and has good robustness. Inspired by the auxiliary model and the iterative identification idea in the AM-LSI algorithm, we apply the BPDN criterion to modify the AM-LSI algorithm by replacing the LS step with the BPDN criterion. The parameters are estimated by the BPDN criterion and the unmeasurable terms are replaced by the outputs of the auxiliary model in each iteration.

Let  $k = 1, 2, \dots$  be the iterative number. To get accurate reconstruction from Equation (17), the information matrix should satisfy some conditions, such as the RIP [29] or the exact recovery condition (ERC) [9]. The ERC guarantee and the consistency properties for the identification of the

controlled autoregressive models have been investigated in [12]. To meet the ERC, we normalize the information matrix  $\hat{\Phi}_k$  defined in Equations (11)–(13) by dividing the elements in each column by the  $l_2$  norm of that column [30,31]. Denote the element of the  $i$ th row and  $j$ th column in  $\hat{\Phi}_k$  by  $\hat{\Phi}_{k,ij}$ , and the normalized information matrix  $\hat{\Phi}_{k,n}$  is constructed by

$$\hat{\Phi}_k := \hat{\Phi}_{k,n} \hat{\Phi}_{l_2}, \quad \hat{\Phi}_{l_2} := \begin{bmatrix} \hat{\Phi}_k(1) & 0 & \cdots & 0 \\ 0 & \hat{\Phi}_k(2) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \hat{\Phi}_k(n) \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad (18)$$

where  $\hat{\Phi}_k(j) := \sqrt{\sum_{i=1}^m (\hat{\Phi}_{k,ij})^2}$ . Similarly, the normalized parameter vector  $\theta_{k,n}$  can be defined as

$$\theta_n := \hat{\Phi}_{l_2} \theta. \quad (19)$$

Note that the location of non-zeros in  $\theta_n$  is identical to that in  $\theta$ . Accordingly, the constrained optimization problem in Equation (17) equals

$$\hat{\theta}_n = \arg \min \|\theta_n\|_1, \quad \text{s.t.} \quad \|Y - \hat{\Phi}_{k,n} \theta_n\|_2^2 \leq \varepsilon. \quad (20)$$

The problem in Equation (20) is closely related to the following unconstrained convex optimization problem

$$\hat{\theta}_n = \min_{\theta_n} \frac{1}{2} \|Y - \hat{\Phi}_{k,n} \theta_n\|_2^2 + \lambda \|\theta_n\|_1, \quad (21)$$

where  $\lambda$  is a nonnegative parameter. Since the information matrix  $\hat{\Phi}_k$  is normalized, we can set  $\lambda$  to the value  $\lambda = \sigma \sqrt{2 \log(n)}$  [16].

The key step of the BPDN approach is to express Equation (21) as a quadratic program (QP) problem [32]. To begin with, two nonnegative vectors  $u_n$  and  $v_n$  are introduced to express  $\theta_n$ . Let  $\theta_{nj}$ ,  $u_{nj}$  and  $v_{nj}$  be the  $j$ th element of the vectors  $\theta_n$ ,  $u_n$  and  $v_n$ , respectively, where  $u_{nj} \geq 0$  and  $v_{nj} \geq 0$  for all  $j = 1, 2, \dots, n$ . Define

$$u_{nj} := (\theta_{nj})_+, \quad v_{nj} := (-\theta_{nj})_+, \quad (*)_+ := \max\{0, *\}.$$

Then,  $\theta_n$  can be rewritten as

$$\theta_n = u_n - v_n.$$

Accordingly, the  $l_1$  regularization term  $\|\theta_n\|_1$  can be expressed as

$$\|\theta_n\|_1 = \mathbf{1}_n^T u_n + \mathbf{1}_n^T v_n = \mathbf{1}_{2n}^T z_n, \quad (22)$$

where  $\mathbf{1}_n^T := [1, \dots, 1] \in \mathbb{R}^n$ ,  $\mathbf{1}_{2n}^T := [1, \dots, 1] \in \mathbb{R}^{2n}$ , and

$$z_n := [u_n^T, v_n^T]^T \in \mathbb{R}^{2n}. \quad (23)$$

Note that all elements in  $z_n$  are nonnegative. Similarly, the quadratic error term can be written as

$$\begin{aligned} & \|Y - \hat{\Phi}_{k,n} \theta_n\|_2^2 \\ &= \|Y - [\hat{\Phi}_{k,n} - \hat{\Phi}_{k,n}] z_n\|_2^2 \\ &= Y^T Y - Y^T [\hat{\Phi}_{k,n}, -\hat{\Phi}_{k,n}] z_n - z_n^T [\hat{\Phi}_{k,n}, -\hat{\Phi}_{k,n}]^T Y + z_n^T [\hat{\Phi}_{k,n}, -\hat{\Phi}_{k,n}]^T [\hat{\Phi}_{k,n}, -\hat{\Phi}_{k,n}] z_n. \end{aligned} \quad (24)$$

Since  $Y^T[\hat{\Phi}_{k,n'} - \hat{\Phi}_{k,n}]\mathbf{z}_n$  is a scalar, it follows that

$$Y^T[\hat{\Phi}_{k,n'} - \hat{\Phi}_{k,n}]\mathbf{z}_n = (Y^T[\hat{\Phi}_{k,n'} - \hat{\Phi}_{k,n}]\mathbf{z}_n)^T = \mathbf{z}_n^T[\hat{\Phi}_{k,n'} - \hat{\Phi}_{k,n}]^T Y = [Y^T \hat{\Phi}_{k,n'} - Y^T \hat{\Phi}_{k,n}]\mathbf{z}_n.$$

Let

$$b := \hat{\Phi}_{k,n}^T Y \in \mathbb{R}^n. \tag{25}$$

Equation (24) can be further written as

$$\begin{aligned} \|Y - \hat{\Phi}_{k,n}\theta_n\|_2^2 &= Y^T Y - 2[b^T, -b^T]\mathbf{z}_n + \mathbf{z}_n^T[\hat{\Phi}_{k,n'} - \hat{\Phi}_{k,n}]^T[\hat{\Phi}_{k,n'} - \hat{\Phi}_{k,n}]\mathbf{z}_n \\ &= Y^T Y - 2[b^T, -b^T]\mathbf{z}_n + \mathbf{z}_n^T \begin{bmatrix} \hat{\Phi}_{k,n}^T \hat{\Phi}_{k,n} & -\hat{\Phi}_{k,n}^T \hat{\Phi}_{k,n} \\ -\hat{\Phi}_{k,n}^T \hat{\Phi}_{k,n} & \hat{\Phi}_{k,n}^T \hat{\Phi}_{k,n} \end{bmatrix} \mathbf{z}_n \\ &= Y^T Y - 2[b^T, -b^T]\mathbf{z}_n + \mathbf{z}_n^T b \mathbf{z}_n, \end{aligned} \tag{26}$$

where

$$b := \begin{bmatrix} \hat{\Phi}_{k,n}^T \hat{\Phi}_{k,n} & -\hat{\Phi}_{k,n}^T \hat{\Phi}_{k,n} \\ -\hat{\Phi}_{k,n}^T \hat{\Phi}_{k,n} & \hat{\Phi}_{k,n}^T \hat{\Phi}_{k,n} \end{bmatrix} \in \mathbb{R}^{2n \times 2n}. \tag{27}$$

From Equations (22) and (26), we have

$$\begin{aligned} &\min_{\theta_n} \frac{1}{2} \|Y - \hat{\Phi}_{k,n}\theta_n\|_2^2 + \lambda \|\theta_n\|_1 \\ &= \min_{\mathbf{z}_n} \frac{1}{2} Y^T Y - [b^T, -b^T]\mathbf{z}_n + \frac{1}{2} \mathbf{z}_n^T b \mathbf{z}_n + \lambda \mathbf{1}_{2n}^T \mathbf{z}_n \\ &= \min_{\mathbf{z}_n} \frac{1}{2} Y^T Y + C^T \mathbf{z}_n + \frac{1}{2} \mathbf{z}_n^T b \mathbf{z}_n, \end{aligned} \tag{28}$$

where

$$C := \lambda \mathbf{1}_{2n} + [-b^T, b^T]^T = \sigma \sqrt{2 \log(n)} \mathbf{1}_{2n} + [-b^T, b^T]^T \in \mathbb{R}^{2n}. \tag{29}$$

Since  $\frac{1}{2} Y^T Y$  is a constant, Equation (28) can be constructed in a standard QP framework,

$$\min_{\mathbf{z}_n} C^T \mathbf{z}_n + \frac{1}{2} \mathbf{z}_n^T b \mathbf{z}_n, \text{ s.t. } z_{nj} \geq 0, j = 1, 2, \dots, n. \tag{30}$$

For the inequality constrained QP problem in Equation (30), the common solution is the active set method [33]. While for the sake of simplicity, the QP problem can be directly solved by calling the relevant function from the standard scientific software toolbox. For example, the MATLAB toolbox provides a function “quadprog”. Then, the parameter vector  $\hat{\theta}_n$  can be obtained from the optimum solution  $\hat{\mathbf{z}}_n$ ,

$$\hat{\theta}_n = \hat{\mathbf{z}}_n(1 : n) - \hat{\mathbf{z}}_n(n + 1 : 2n), \tag{31}$$

where  $\hat{\mathbf{z}}_n(1 : n)$  represents a vector formed by the first  $n$  elements of  $\hat{\mathbf{z}}_n$ , and  $\hat{\mathbf{z}}_n(n + 1 : 2n)$  a vector formed by the last  $n$  elements of  $\hat{\mathbf{z}}_n$ . Considering that the system in Equation (8) is contaminated with noise, the parameter estimation error can be large. To further reduce the estimation error, a small threshold  $\text{TH} = \epsilon$  can be set to filter the elements close to zero in  $\hat{\theta}_n$ . Let  $\hat{\theta}_{nj} = 0$  if  $\hat{\theta}_{nj} < \epsilon$  and denote

the filtered parameter vector as  $\hat{\theta}_{n,\epsilon}$ . Then, the parameter vector estimate  $\hat{\theta}_k$  can be recovered according to Equation (19),

$$\hat{\theta}_k = \hat{\Phi}_{l_2}^{-1} \hat{\theta}_{n,\epsilon}. \tag{32}$$

The estimates of the intermediate variables  $x_{i,k}(t)$  can be refreshed by  $\hat{\theta}_k$  as shown in Equations (15) and (16).

Equations (9), (11)–(13), (15), (16), (18), (23), (25), (27), and (29)–(32) form the auxiliary model-basis pursuit de-noising iterative (AM-BPDNI) algorithm for the MISO-OE system. The implementation procedures are listed as follows:

1. Collect the input–output data  $\{u_i(t), y(t): i = 1, 2, \dots, r; t = 1, 2, \dots, m\}$  and set the parameter estimation accuracy  $\epsilon_0$ .
2. Construct the stacked output vector  $\mathbf{Y}$  by Equation (9).
3. Initialize the iteration: let  $k = 1$  and  $\hat{x}_{i,0}(t)$  be random sequences.
4. Construct the information matrix  $\hat{\Phi}_k$  by Equations (11)–(13) and normalize  $\hat{\Phi}_k$  by (18).
5. Form the vectors and matrix  $\mathbf{z}_n, \mathbf{b}, \mathbf{b}$  and  $\mathbf{C}$  by Equations (23), (25), (27), and (29) and formulate the QP by Equation (30).
6. Call the function to obtain the optimum solution  $\hat{\mathbf{z}}_n$  and compute  $\hat{\theta}_n$  by Equation (31).
7. Set a threshold to obtain  $\hat{\theta}_{n,\epsilon}$  and recover the parameter vector estimate  $\hat{\theta}_k$  by (32).
8. Compare  $\hat{\theta}_k$  with  $\hat{\theta}_{k-1}$ : if  $\|\hat{\theta}_k - \hat{\theta}_{k-1}\| > \epsilon_0$ , update the auxiliary model outputs  $x_{i,k}(t)$  by Equations (15) and (16) and go to Step 4. Otherwise, stop the iteration and obtain the parameter vector estimate  $\hat{\theta}$ .

The unknown time-delay of each input channel can be estimated according to the location of zero-blocks and the number of zeros in  $\hat{\theta}$ . It can be seen from Equations (3) and (6) that there are  $2r$  zero-blocks in  $\hat{\theta}$ . Denote the number of zeros in each zero-block by  $z_i$  ( $i = 1, 2, \dots, 2r$ ). Then, time-delays can be estimated by

$$\hat{d}_i = z_{2i-1}, \quad i = 1, 2, \dots, 2r. \tag{33}$$

#### 4. Simulation Example

**Example 1.** Consider the following MISO-OE system with time-delays,

$$y(t) = \sum_{i=1}^3 z^{-d_i} \frac{B_i(z)}{A_i(z)} u_i(t) + v(t),$$

$$\begin{aligned} A_1(z) &= 1 - 0.1z^{-1} + 0.7z^{-2}, & B_1(z) &= 1.5z^{-1} + 0.9z^{-2}, \\ A_2(z) &= 1 + 0.3z^{-1} + 0.5z^{-2}, & B_2(z) &= 0.2z^{-1} + 1.8z^{-2}, \\ A_3(z) &= 1 - 0.2z^{-1} - 0.4z^{-2}, & B_3(z) &= -0.1z^{-1} + 2z^{-2}, \end{aligned} \tag{34}$$

$$d_1 = 20, \quad d_2 = 10, \quad d_3 = 30.$$

The system in Equation (34) is a second order system with three inputs and one output. The inputs  $\{u_i(t)\}$ ,  $i = 1, 2, 3$  are taken as random uncorrelated signal sequences with zero mean and unit variances, and  $\{v(t)\}$  as a white noise sequence with zero mean and variances  $\sigma^2$ . Let the maximum input regression length be  $l = 50$ . Then, the parameter vector to be identified is

$$\theta_1 = [-0.1, 0.7, \mathbf{0}_{20}, 1.5, 0.9, \mathbf{0}_{28}, 0.3, 0.5, \mathbf{0}_{10}, 0.2, 1.8, \mathbf{0}_{38}, -0.2, -0.4, \mathbf{0}_{30}, -0.1, 2, \mathbf{0}_{18}]^T \in \mathbb{R}^{156},$$

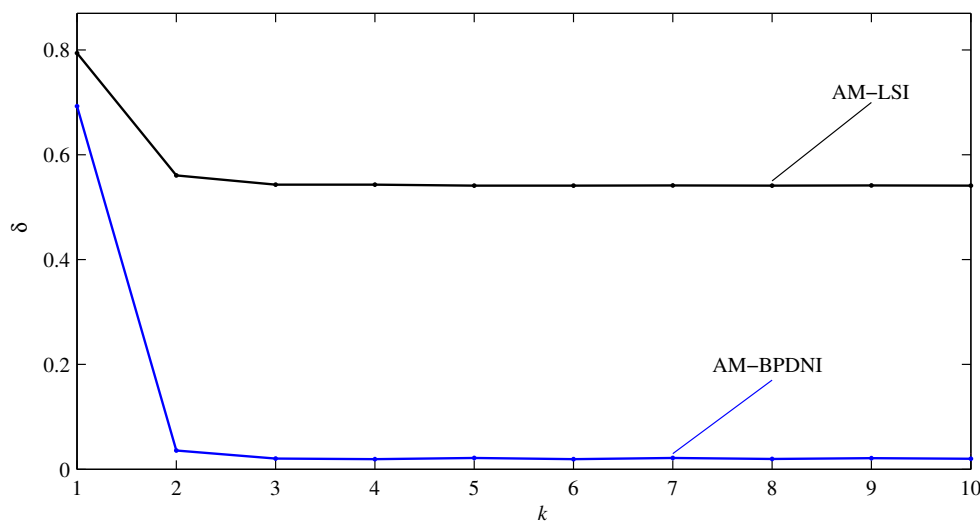
where  $\mathbf{0}_i$  denotes the zero-block with  $i$  zeros. Note that the number of non-zero elements is  $K = \|\theta\|_0 = \sum_{i=1}^3 (n_{ai} + n_{bi}) = 12$ .



Taking  $m = 130$  and  $TH = 0.001$ , apply the AM-LSI algorithm and the AM-BPDNI algorithm to perform the identification, respectively. The parameter estimation errors  $\delta := \|\theta - \hat{\theta}\| / \|\theta\|$  versus different noise levels are shown in Table 1. When  $\sigma^2 = 0.10^2$ , the estimation errors  $\delta$  versus the iterative number  $k$  are shown in Figure 1. It can be seen that the AM-BPDNI algorithm performs better than the AM-LSI algorithm and is insensitive to noise.

**Table 1.** The parameter estimation errors  $\delta$  (%) versus the noise variances  $\sigma^2$  ( $m = 130$ ).

$\sigma$	0.10	0.15	0.20	0.25	0.30	0.40
AM-LSI	53.4284	53.7175	54.2753	55.3337	55.8435	57.2966
AM-BPDNI	1.9974	2.9825	4.7985	8.6665	8.9512	9.8808

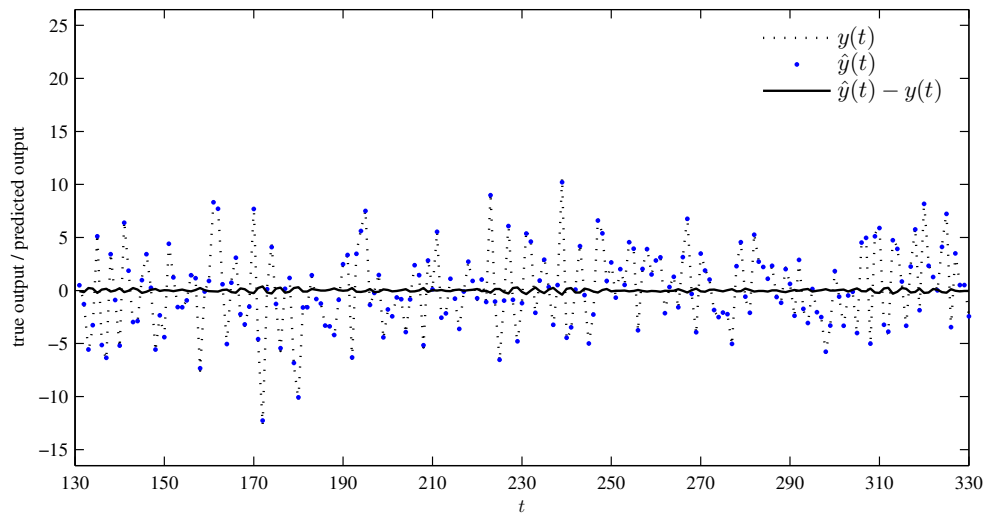


**Figure 1.** The parameter estimation error  $\delta$  versus  $k$  ( $m = 130, \sigma^2 = 0.10^2$ ).

Let the variance of  $\{v(t)\}$  be  $\sigma^2 = 0.10^2$ . Apply the AM-BPDNI algorithm to obtain the estimated model of the system in Equation (34) with the first  $m = 130$  data. Then, validate the estimated model by using  $m_e = 200$  samples from  $t = 131$  to 330. The predicted output of the estimated model, the true output of the system and their errors are plotted in Figure 2. It is shown that the predicted outputs  $\hat{y}(t)$  are close to the true outputs  $y(t)$ . Moreover, the average predicted output error

$$\delta_v := \sqrt{\frac{1}{m_e} \sum_{t=131}^{330} [\hat{y}(t) - y(t)]^2} = 0.1422$$

is small and close to the standard deviation of the noise  $\sigma = 0.10$ . It follows that the estimated model can well capture the system dynamics.

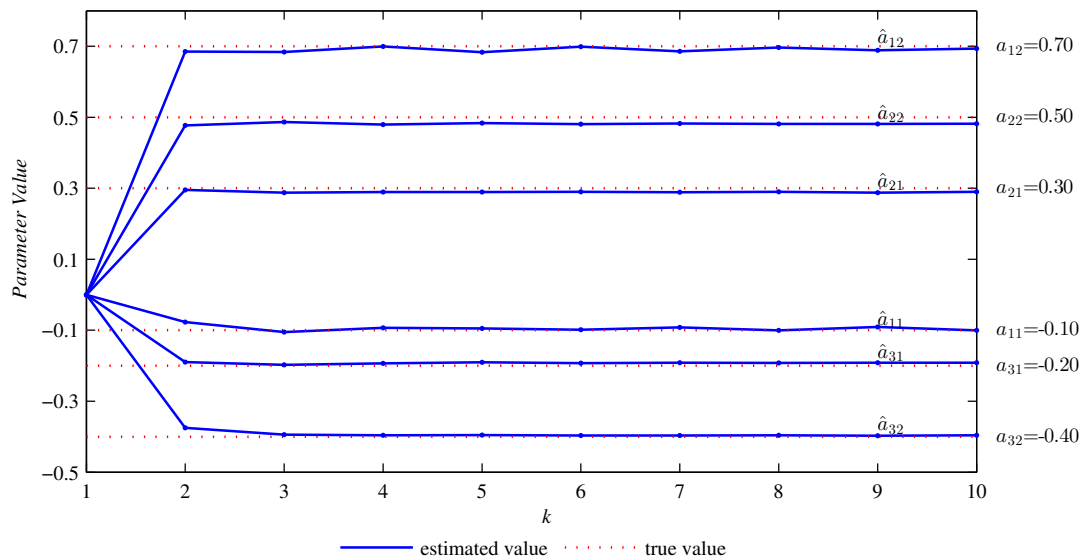


**Figure 2.** Predicted output  $\hat{y}(t)$ , true output  $y(t)$  and errors (from  $t = 131$  to  $330$ ).

Let  $m = 130$  and  $\sigma^2 = 0.10^2$ . Using the AM-BPDNI algorithm to estimate the sparse parameter vector  $\theta$ , the non-zero parameter estimates versus  $k$  are shown in Table 2 and Figure 3.

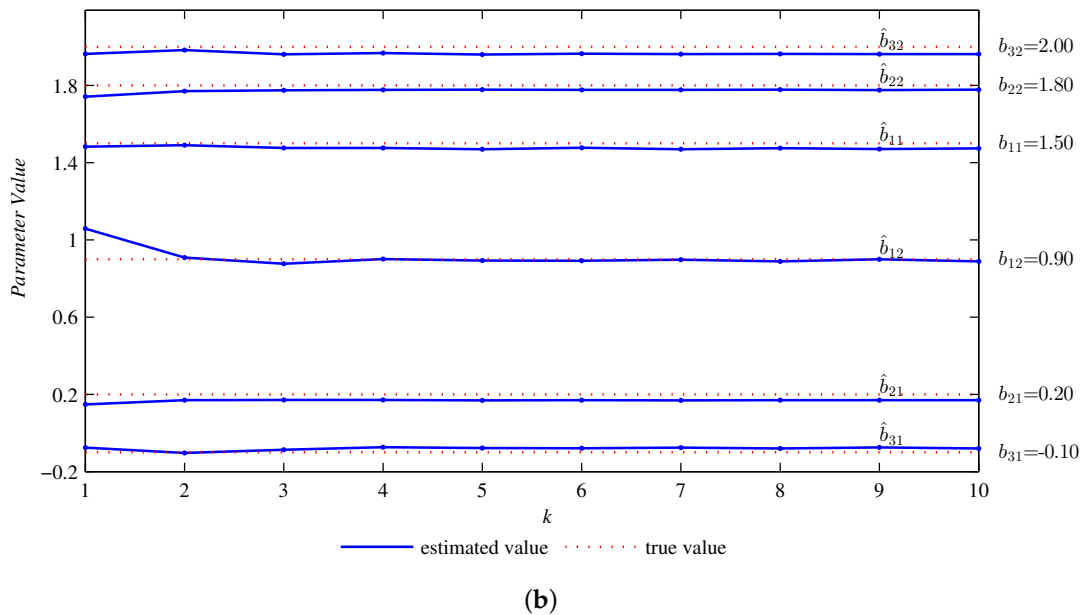
**Table 2.** The non-zero parameter estimates and estimation error  $\delta$  versus  $k$  ( $m = 130, \sigma^2 = 0.10^2$ ).

$k$	$a_{11}$	$a_{12}$	$b_{11}$	$b_{12}$	$a_{21}$	$a_{22}$	$b_{21}$	$b_{22}$	$a_{31}$	$a_{32}$	$b_{31}$	$b_{32}$	$\delta(\%)$
1	0.000	0.000	1.483	1.058	0.000	0.000	0.148	1.742	0.000	0.000	-0.076	1.963	69.2788
2	-0.077	0.685	1.491	0.908	0.296	0.477	0.170	1.771	-0.190	-0.375	-0.104	1.983	3.5707
5	-0.095	0.684	1.470	0.893	0.289	0.484	0.168	1.778	-0.190	-0.396	-0.078	1.960	2.1595
10	-0.100	0.694	1.474	0.888	0.290	0.482	0.170	1.778	-0.192	-0.396	-0.080	1.963	1.9974
True value	-0.100	0.700	1.500	0.900	0.300	0.500	0.200	1.800	-0.200	-0.400	-0.100	2.000	



**(a)**

**Figure 3.** Cont.



**Figure 3.** (a) The parameter estimates  $\hat{a}_{ij}$  versus  $k$ . (b) The parameter estimates  $\hat{b}_{ij}$  versus  $k$ . The non-zero parameter estimates versus  $k$  with  $m = 130, \sigma^2 = 0.10^2$ .

With 10 iterations, the estimated parameter vector is

$$\hat{\theta}_1 = [-0.1001, 0.6938, \mathbf{0}_{20}, 1.4737, 0.8882, \mathbf{0}_{28}, 0.2902, 0.4818, \mathbf{0}_{10}, 0.1699, 1.7783, \mathbf{0}_{38}, -0.1917, -0.3961, \mathbf{0}_{30}, -0.08002, 1.9628, \mathbf{0}_{18}]^T. \tag{35}$$

It can be seen from Equation (35) that there are six zero-blocks in  $\hat{\theta}$  and the number of zeros in each zero-block are  $z_1 = 20, z_2 = 28, z_3 = 10, z_4 = 38, z_5 = 30$  and  $z_6 = 18$ . Then, the time-delay of each input channel can be estimated according to Equation (33),

$$\hat{d}_1 = z_1 = 20, \hat{d}_2 = z_3 = 10, \hat{d}_3 = z_5 = 30.$$

Obviously, the time-delay estimates are agreement with the true time-delays.

**Example 2.** Consider the following MISO-OE system with time-delays,

$$y(t) = \sum_{i=1}^4 z^{-d_i} \frac{B_i(z)}{A_i(z)} u_i(t) + v(t),$$

$$\begin{aligned} A_1(z) &= 1 - 0.1z^{-1} + 0.7z^{-2}, & B_1(z) &= 1.5z^{-1} + 0.9z^{-2}, \\ A_2(z) &= 1 + 0.3z^{-1} + 0.5z^{-2}, & B_2(z) &= 0.2z^{-1} + 1.8z^{-2}, \\ A_3(z) &= 1 - 0.2z^{-1} - 0.4z^{-2}, & B_3(z) &= -0.1z^{-1} + 2z^{-2}, \\ A_4(z) &= 1 - 0.4z^{-1} - 0.1z^{-2}, & B_4(z) &= 1.1z^{-1} + 0.8z^{-2}, \end{aligned} \tag{36}$$

$$d_1 = 20, \quad d_2 = 10, \quad d_3 = 30, \quad d_4 = 40.$$

Compared with the system in Equation (34), the system in Equation (36) has one more input. Thus, the number of parameters is increased. Let  $l = 50$  and the true parameter vector is

$$\theta_2 = [-0.1, 0.7, \mathbf{0}_{20}, 1.5, 0.9, \mathbf{0}_{28}, 0.3, 0.5, \mathbf{0}_{10}, 0.2, 1.8, \mathbf{0}_{38}, -0.2, -0.4, \mathbf{0}_{30}, -0.1, 2, \mathbf{0}_{18}, -0.4, -0.1, \mathbf{0}_{40}, 1.1, 0.8, \mathbf{0}_8]^T \in \mathbb{R}^{208}.$$

Taking  $m = 130$ ,  $\sigma^2 = 0.10^2$ , and  $\text{TH} = 0.001$ , employ the AM-BPDNI algorithm to identify the system in Equation (36). The estimated parameter vector is

$$\hat{\theta}_2 = [-0.0874, 0.6849, \mathbf{0}_{20}, 1.4682, 0.9218, \mathbf{0}_{28}, 0.2880, 0.4711, \mathbf{0}_{10}, 0.1658, 1.7708, \mathbf{0}_{38}, -0.1842, -0.3968, \mathbf{0}_{30}, -0.0759, 1.9503, \mathbf{0}_{18}, -0.4563, -0.0556, \mathbf{0}_{40}, 1.0429, 0.7457, \mathbf{0}_8]^T.$$

The time-delay estimates are

$$\hat{d}_1 = z_1 = 20, \hat{d}_2 = z_3 = 10, \hat{d}_3 = z_5 = 30, \hat{d}_4 = z_7 = 40,$$

which are identical to the true time-delays.

The parameter estimation errors of the systems in Equations (34) and (36) versus  $k$  are shown in Figure 4.

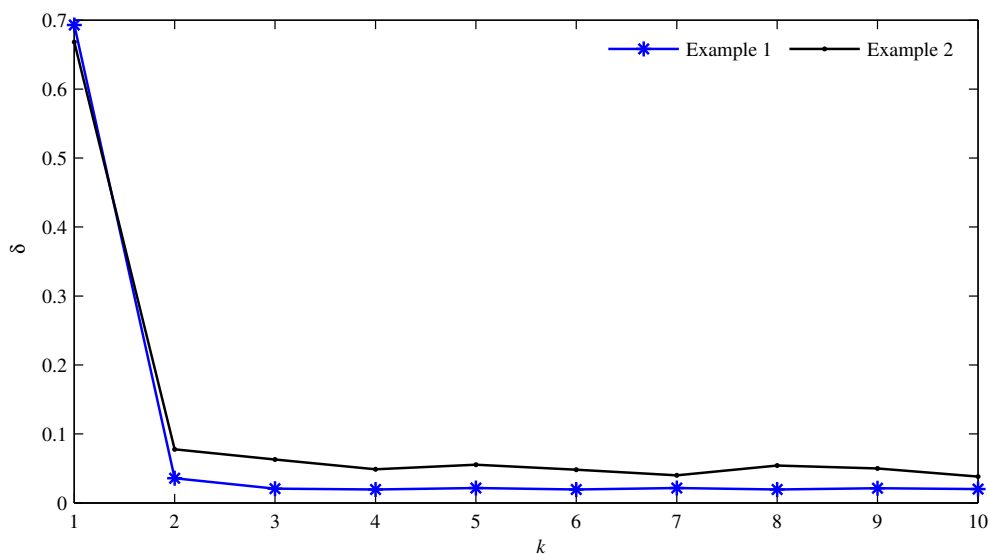


Figure 4. The parameter estimation errors  $\delta$  of different systems versus  $k$ .

The running time of the proposed method for the systems is  $t_1 = 1.374637$  s and  $t_2 = 2.398602$  s. It can be concluded that the computational burden increases as the dimension of the parameter vector increases.

The simulation results show that for the MISO-OE model, the proposed AM-BPDNI algorithm can obtain efficient estimation of parameters from few observations ( $m < n$ ) with good robustness. Moreover, the AM-BPDNI algorithm can effectively estimate the time-delays according to the sparse characteristic of the estimated parameter vector. However, as the number of the input channels increases, the computational burden of the proposed algorithm increases.

## 5. Conclusions

This paper proposes an AM-BPDNI algorithm for the identification of MISO-OE systems with unknown time-delays. Based on the BPDN criterion and the auxiliary model identification idea, the sparse parameters and multiple-input time-delays can be effectively and simultaneously estimated. The proposed algorithm requires few sampled data and is robust to noise.

**Author Contributions:** J.Y. wrote the manuscript. Y.L. revised the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (Nos. 61304138 and 61473136) and the Jiangsu Province Industry University Prospective Joint Research Project (BY2015019-29).

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

## References

1. Prochazka, A.; Kingsbury, N.; Payner, P.J.W.; Uhler, J. *Signal Analysis and Prediction*; Birkhäuser Basel: Boston, MA, USA, 1998.
2. Pappalardo, C.M.; Guida, D. System identification algorithm for computing the modal parameters of linear mechanical systems. *Machines* **2018**, *6*, 12. [[CrossRef](#)]
3. Pappalardo, C.M.; Guida, D. System identification and experimental modal analysis of a frame structure. *Eng. Lett.* **2018**, *26*, 56–68.
4. Pappalardo, C.M.; Guida, D. A time-domain system identification numerical procedure for obtaining linear dynamical models of multibody mechanical systems. *Arch. Appl. Mech.* **2018**, *88*, 1325–1347. [[CrossRef](#)]
5. Gibson, S.; Ninness, B. Robust maximum-likelihood estimation of multivariable dynamic systems. *Automatica* **2005**, *41*, 1667–1682. [[CrossRef](#)]
6. Romano, R.A.; Pait, F. Matchable-observable linear models and direct filter tuning: an approach to multivariable identification. *IEEE Trans. Autom. Control* **2017**, *62*, 2180–2193. [[CrossRef](#)]
7. Patwardhan, S.C.; Shah, S.L. From data to diagnosis and control using generalized orthonormal basis filters. Part I: Development of state observers. *J. Process Control* **2005**, *15*, 819–835. [[CrossRef](#)]
8. Selvanathan, S.; Tangirala, A.K. Time-delay estimation in multivariate systems using Hilbert transform relation and partial coherence functions. *Chem. Eng. Sci.* **2010**, *65*, 660–674. [[CrossRef](#)]
9. Tropp, J.A. Just relax: Convex programming methods for identifying sparse signals in noise. *IEEE Trans. Inf. Theory* **2006**, *52*, 1030–1051. [[CrossRef](#)]
10. Elad, M. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*; Springer: New York, NY, USA, 2010.
11. Sanandaji, B.M.; Vincent, T.L.; Wakin, M.B.; Tóth, R.; Poolla, K. Compressive system identification of LTI and LTV ARX models. In Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, USA, 12–15 December 2011; pp. 791–798.
12. Tóth, R.; Sanandaji, B.M.; Poolla, K.; Vincent, T.L. Compressive system identification in the Linear Time-Invariant framework. In Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, USA, 12–15 December 2011; pp. 783–790.
13. Donoho, D.L. Compressed Sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306. [[CrossRef](#)]
14. Tropp, J.A.; Gilbert, A.C.; Strauss, M.J. Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit. *Signal Process.* **2006**, *86*, 572–588. [[CrossRef](#)]
15. Tropp, J.A. Algorithms for simultaneous sparse approximation. Part II: Convex relaxation. *Signal Process.* **2006**, *86*, 589–602. [[CrossRef](#)]
16. Chen, S.S.; Donoho, D.L.; Saunders, M.A. Atomic decomposition by basis pursuit. *SIAM Rev.* **2001**, *43*, 129–159. [[CrossRef](#)]
17. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Statist. Soc. Ser. B (Methodological)* **1996**, *58*, 267–288.
18. Efron, B.; Hastie, T.; Johnstone, I.; Tibshirani, R. Least angle regression. *Ann. Statist.* **2004**, *32*, 407–451.
19. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: New York, NY, USA, 2004.
20. Liu, Y.J.; Tao, T.Y. A CS recovery algorithm for model and time delay identification of MISO-FIR systems. *Algorithms* **2015**, *8*, 743–753. [[CrossRef](#)]
21. Liu, Y.J.; Tao, T.Y.; Ding, F. Parameter and time-delay identification for MISO systems based on orthogonal matching pursuit algorithm. *Control Decis.* **2015**, *30*, 2103–2107.
22. Liu, Y.J.; H, X.; Ding, F. An instrumental variable based compressed sampling matching pursuit method for closed-loop identification. *Control Decis.* **2017**, *32*, 1837–1843.
23. Sánchez-Peña, R.S.; Casín, J.Q.; Cayuela, V.P. *Identification and Control: The Gap Between Theory and Practice*; Springer: London, UK, 2007.
24. Ding, F.; Chen, T. Combined parameter and output estimation of dual-rate systems using an auxiliary model. *Automatica* **2004**, *40*, 1739–1748. [[CrossRef](#)]
25. Liu, Q.Y.; Ding, F. The data filtering based generalized stochastic gradient parameter estimation algorithms for multivariate output-error autoregressive systems using the auxiliary model. *Multidimens. Syst. Signal Process.* **2018**, *29*, 1781–1800. [[CrossRef](#)]

26. Wang, Y.J.; Ding, F. Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model. *Automatica* **2016**, *71*, 308–313. [[CrossRef](#)]
27. Wang, Y.J.; Ding, F. The filtering based iterative identification for multivariable systems. *IET Control Theory Appl.* **2016**, *10*, 894–902. [[CrossRef](#)]
28. Ma, J.X.; Ding, F.; Yang, E.F. Data filtering-based least squares iterative algorithm for Hammerstein nonlinear systems by using the model decomposition. *Nonlinear Dyn.* **2016**, *83*, 1895–1908. [[CrossRef](#)]
29. Candès, E.J.; Romberg, J.; Tao, T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory* **2006**, *52*, 489–509. [[CrossRef](#)]
30. Wang, W.X.; Yang, R.; Lai, Y.C.; Kovanis, V.; Grebogi, C. Predicting catastrophes in nonlinear dynamical systems by compressive sensing. *Phys. Rev. Lett.* **2011**, *106*, 154101. [[CrossRef](#)] [[PubMed](#)]
31. Naik, M.; Cochran, D. Nonlinear system identification using compressed sensing. In Proceedings of the 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 4–7 November 2012; pp. 426–430.
32. Figueiredo, M.A.T.; Nowak, R.D.; Wright, S.J. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE J. Sel. Top. Signal Process.* **2007**, *1*, 586–597. [[CrossRef](#)]
33. Nocedal, J.; Wright, S.J. *Numerical Optimization*, 2nd ed.; Springer: New York, NY, USA, 2006.



© 2018 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/>).