ICIC Express Letters ICIC International ©2016 ISSN 1881-803X
Volume **10**, Number **12**, December **2016** pp. **2927–2933**

MULTI-INNOVATION STOCHASTIC GRADIENT ALGORITHM FOR MULTI-INPUT OUTPUT-ERROR SYSTEMS USING THE FILTERING TECHNIQUE

Jiling Ding[1,2]

[1]Department of Mathematics
Jining University
No. 1, Xingtan Road, Qufu 273155, P. R. China
jlding@jnxy.edu.cn; dingjl08@163.com

[2]School of Internet of Things Engineering
Jiangnan University
No. 1800, Lihu Road, Wuxi 214122, P. R. China

Received May 2016; accepted August 2016

Abstract. *This paper discusses the identification problem of multi-input output-error moving average systems. A filtering based multi-innovation stochastic gradient algorithm is derived by using the filtering technique and the multi-innovation identification theory. The simulation results confirm that the proposed algorithm can generate highly accurate parameter estimates compared with the multi-innovation stochastic gradient algorithm.*
Keywords: Multi-innovation identification, Stochastic gradient algorithm, Filtering technique, Parameter estimation

1. **Introduction.** System identification is the methodology of modeling and identification of systems [1, 2, 3]. Many physical systems have multiple inputs and multiple outputs in industrial processes [4, 5, 6], so much work has been performed on the study of modeling and identification of multivariable systems [7, 8, 9]. Recently, Nasirin et al. studied an adaptive scheme of designing sliding mode control for a class of multi-input multi-output nonlinear systems [10]. Tao provided some fundamental theoretical aspects and technical issues of multivariable adaptive control [11].

Many parameter estimation methods have been developed for multivariable systems [12, 13, 14], e.g., the recursive parameter estimation algorithms [15, 16, 17], the auxiliary model identification methods [18, 19, 20], and the iterative identification methods [21, 22, 23].

The multi-innovation identification theory has been proved to be effective in the field of system identification [24, 25, 26]. Many research results have been performed for linear regression systems [27], and nonlinear systems [28, 29, 30]. The main role is to improve the parameter estimation accuracy by expanding a scalar innovation to an innovation vector. The filtering technique is another effective method to improve the convergence rates and the parameter estimation accuracy [31]. By adopting the filtering technique and the multi-innovation identification theory, this paper discusses the identification problem for multi-input output-error moving average (OEMA) systems and presents a data filtering based multi-innovation extended stochastic gradient (F-MI-ESG) identification algorithm. Compared with the multi-innovation extended stochastic gradient (MI-ESG) identification algorithm, the F-MI-ESG algorithm can obtain more accurate parameter estimates.

The rest of this paper is organized as follows. Section 2 discusses the identification model of multi-input OEMA systems. Section 3 presents a data filtering based multi-innovation extended stochastic gradient algorithm. Section 4 provides an illustrative example. Finally, Section 5 gives some concluding remarks.

2. **The System Description and Identification Model.** Consider the multi-input OEMA system:

$$y(t)=\sum_{j=1}^{r}\frac{B_j(q)}{A_j(q)}u_j(t)+D(q)v(t), \tag{1}$$

where $u_j(t)\in\mathbb{R}$, $j=1,2,\cdots,r$, are the inputs, $y(t)\in\mathbb{R}$ is the system output, $v(t)\in\mathbb{R}$ is the random white noise with zero mean and variance σ^2, and $A_j(q)$, $B_j(q)$ and $D(q)$ are polynomials in the unit backward shift operator q^{-1} with

$$\begin{aligned}A_j(q) &:= 1+a_{j1}q^{-1}+a_{j2}q^{-2}+\cdots+a_{jn_j}q^{-n_j},\\ B_j(q) &:= b_{j1}q^{-1}+b_{j2}q^{-2}+\cdots+b_{jn_j}q^{-n_j},\\ D(q) &:= 1+d_1q^{-1}+d_2q^{-2}+\cdots+d_{n_d}q^{-n_d}.\end{aligned}$$

Assume that the orders n_j and n_d are known, $y(t)=0$, $u_j(t)=0$ and $v(t)=0$ as $t<0$, and a_{ji}, b_{ji} and d_j are the unknown parameters to be identified from available input-output data $\{u_1(t),u_2(t),\cdots,u_r(t),y(t)\}$.

We use a linear filter $D^{-1}(q)$ to filter the input-output data. Define the filtered input $u_{j\mathrm{f}}(t)$, the filtered noise-free output $x_{j\mathrm{f}}(t)$ and the filtered output $y_{\mathrm{f}}(t)$ as

$$u_{j\mathrm{f}}(t) := D^{-1}(q)u_j(t)=-\sum_{i=1}^{n_d}d_iu_{j\mathrm{f}}(t-i)+u_j(t), \tag{2}$$

$$x_{j\mathrm{f}}(t) := \frac{B_j(q)}{A_j(q)}u_{j\mathrm{f}}(t)=-\sum_{i=1}^{n_j}a_{ji}x_{j\mathrm{f}}(t-i)+\sum_{i=1}^{n_j}b_{ji}u_{j\mathrm{f}}(t-i), \tag{3}$$

$$y_{\mathrm{f}}(t) := D^{-1}(q)y(t)=-\sum_{i=1}^{n_d}d_iy_{\mathrm{f}}(t-i)+y(t). \tag{4}$$

Multiplying both sides of (1) by $D^{-1}(q)$, Equation (1) can be expressed as

$$y_{\mathrm{f}}(t)=\sum_{j=1}^{r}x_{j\mathrm{f}}(t)+v(t)=-\sum_{j=1}^{r}\sum_{i=1}^{n_j}a_{ji}x_{j\mathrm{f}}(t-i)+\sum_{j=1}^{r}\sum_{i=1}^{n_j}b_{ji}u_{j\mathrm{f}}(t-i)+v(t). \tag{5}$$

Substituting (5) into (4), we can obtain

$$y(t)=-\sum_{j=1}^{r}\sum_{i=1}^{n_j}a_{ji}x_{j\mathrm{f}}(t-i)+\sum_{j=1}^{r}\sum_{i=1}^{n_j}b_{ji}u_{j\mathrm{f}}(t-i)+\sum_{i=1}^{n_d}d_iy_{\mathrm{f}}(t-i)+v(t). \tag{6}$$

Let the superscript T denote the vector/matrix transpose. Define the filtered information vectors

$$\begin{aligned}\boldsymbol{\varphi}_{\mathrm{f}}(t) &:= \left[\boldsymbol{\phi}_{1\mathrm{f}}^{\mathrm{T}}(t),\boldsymbol{\phi}_{2\mathrm{f}}^{\mathrm{T}}(t),\cdots,\boldsymbol{\phi}_{r\mathrm{f}}^{\mathrm{T}}(t),y_{\mathrm{f}}(t-1),y_{\mathrm{f}}(t-2),\cdots,y_{\mathrm{f}}(t-n_d)\right]^{\mathrm{T}}\in\mathbb{R}^{n},\\ \boldsymbol{\phi}_{j\mathrm{f}}(t) &:= [-x_{j\mathrm{f}}(t-1),\cdots,-x_{j\mathrm{f}}(t-n_j),u_{j\mathrm{f}}(t-1),\cdots,u_{j\mathrm{f}}(t-n_j)]^{\mathrm{T}}\in\mathbb{R}^{2n_j},\end{aligned}$$

and the parameter vectors

$$\begin{aligned}\boldsymbol{\theta} &:= \left[\boldsymbol{\vartheta}_1^{\mathrm{T}},\boldsymbol{\vartheta}_2^{\mathrm{T}},\cdots,\boldsymbol{\vartheta}_r^{\mathrm{T}},\boldsymbol{d}^{\mathrm{T}}\right]^{\mathrm{T}}\in\mathbb{R}^{n},\\ \boldsymbol{\vartheta}_j &:= \left[a_{j1},a_{j2},\cdots,a_{jn_j},b_{j1},b_{j2},\cdots,b_{jn_j}\right]^{\mathrm{T}}\in\mathbb{R}^{2n_j},\\ \boldsymbol{d} &:= [d_1,d_2,\cdots,d_{n_d}]^{\mathrm{T}}\in\mathbb{R}^{n_d}.\end{aligned}$$

Equations (3) and (6) can be rewritten as

$$x_{j\mathrm{f}}(t) = \boldsymbol{\phi}_{j\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{\vartheta}_j(t), \tag{7}$$

$$y(t) = \boldsymbol{\varphi}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{\theta}+v(t). \tag{8}$$

Equation (8) is the identification model of the multi-input OEMA system in (1), and contains all the parameters in $\boldsymbol{\theta}$ to be identified.

3. **The Data Filtering Based Multi-Innovation Stochastic Gradient Identification Algorithm.** The focus of this paper is to combine the multi-innovation identification theory and the filtering technique to improve the parameter estimation accuracy of the stochastic gradient identification algorithm. Note that the filtered variables $u_{j\mathrm{f}}(t)$, $x_{j\mathrm{f}}(t)$ and $y_{\mathrm{f}}(t)$ in the information vector $\boldsymbol{\varphi}_{\mathrm{f}}(t)$ are all unknown, so the stochastic gradient algorithm cannot be applied to obtain the estimate of $\boldsymbol{\theta}$ directly. We use the estimates $\hat{u}_{j\mathrm{f}}(t)$, $\hat{x}_{j\mathrm{f}}(t)$ and $\hat{y}_{\mathrm{f}}(t)$ to construct the estimate of $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$ [32]. The details are as follows.

Let $\hat{\boldsymbol{\theta}}(t)$, $\hat{\boldsymbol{\vartheta}}_j(t)$ and $\hat{\boldsymbol{d}}(t)$ be the estimates of $\boldsymbol{\theta}$, $\boldsymbol{\vartheta}_j$ and $\boldsymbol{d}$ at time t. Replacing $\boldsymbol{\phi}_{j\mathrm{f}}(t)$ and $\boldsymbol{\vartheta}_j$ in (7) with their estimates $\hat{\boldsymbol{\phi}}_{j\mathrm{f}}(t)$ and $\hat{\boldsymbol{\vartheta}}_j(t)$, the estimate of $x_{j\mathrm{f}}(t)$ can be computed by

$$\hat{x}_{j\mathrm{f}}(t) = \hat{\boldsymbol{\phi}}_{j\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}_j(t).$$

Use $\hat{\boldsymbol{d}}(t) := \left[\hat{d}_1(t), \hat{d}_2(t), \cdots, \hat{d}_{n_d}(t)\right]^{\mathrm{T}}$ to construct the estimate $\hat{D}(t,q)$ of the polynomial $D(q)$:

$$\hat{D}(t,q) := 1 + \hat{d}_1(t)q^{-1} + \hat{d}_2(t)q^{-2} + \cdots + \hat{d}_{n_d}(t)q^{-n_d},$$

and the estimates $\hat{y}_{\mathrm{f}}(t)$ and $\hat{u}_{j\mathrm{f}}(t)$ can be obtained by filtering the input-output data with $\hat{D}^{-1}(t,q)$:

$$\hat{y}_{\mathrm{f}}(t) := \hat{D}^{-1}(t,q)y(t) = [-\hat{y}_{\mathrm{f}}(t-1), -\hat{y}_{\mathrm{f}}(t-2), \cdots, -\hat{y}_{\mathrm{f}}(t-n_d)]\hat{\boldsymbol{d}}(t) + y(t),$$
$$\hat{u}_{j\mathrm{f}}(t) := \hat{D}^{-1}(t,q)u_{j\mathrm{f}}(t) = [-\hat{u}_{j\mathrm{f}}(t-1), -\hat{u}_{j\mathrm{f}}(t-2), \cdots, -\hat{u}_{j\mathrm{f}}(t-n_d)]\hat{\boldsymbol{d}}(t) + u_j(t).$$

Then the estimate $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$ can be constructed by the estimates $\hat{x}_{j\mathrm{f}}(t-i)$, $\hat{u}_{j\mathrm{f}}(t-i)$ and $\hat{y}_{\mathrm{f}}(t)$:

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t) := \left[\hat{\boldsymbol{\phi}}_{1\mathrm{f}}^{\mathrm{T}}(t), \hat{\boldsymbol{\phi}}_{2\mathrm{f}}^{\mathrm{T}}(t), \cdots, \hat{\boldsymbol{\phi}}_{r\mathrm{f}}^{\mathrm{T}}(t), \hat{y}_{\mathrm{f}}(t-1), \hat{y}_{\mathrm{f}}(t-2), \cdots, \hat{y}_{\mathrm{f}}(t-n_d)\right]^{\mathrm{T}} \in \mathbb{R}^n,$$
$$\hat{\boldsymbol{\phi}}_{j\mathrm{f}}(t) := [-\hat{x}_{j\mathrm{f}}(t-1), \cdots, -\hat{x}_{j\mathrm{f}}(t-n_j), \hat{u}_{j\mathrm{f}}(t-1), \cdots, \hat{u}_{j\mathrm{f}}(t-n_j)]^{\mathrm{T}} \in \mathbb{R}^{2n_j}.$$

By defining and minimizing the quadratic criterion function

$$J(\boldsymbol{\theta}) := \mathrm{E}\left[\left\|\mathrm{y(t)} - \hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(\mathrm{t})\boldsymbol{\theta}\right\|^2\right],$$

We can obtain the following extended stochastic gradient algorithm for estimating $\boldsymbol{\theta}$ based on the data filtering (the F-ESG algorithm for short):

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \frac{\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)}{r(t)}e(t), \tag{9}$$

$$e(t) = y(t) - \hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1), \tag{10}$$

$$r(t) = r(t-1) + \|\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)\|^2, \quad r(0) = 1. \tag{11}$$

In order to improve the parameter estimation accuracy of the F-ESG algorithm, we expand the scalar innovation $e(t) \in \mathbb{R}$ to an innovation vector $\boldsymbol{E}(p,t) \in \mathbb{R}^p$ by adopting the multi-innovation identification theory. Define the innovation vector

$$\boldsymbol{E}(p,t) := \begin{bmatrix} y(t) - \hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1) \\ y(t-1) - \hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t-1)\hat{\boldsymbol{\theta}}(t-1) \\ \vdots \\ y(t-p+1) - \hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t-p+1)\hat{\boldsymbol{\theta}}(t-1) \end{bmatrix} \in \mathbb{R}^p,$$

where p represents the innovation length. Referring to the method in [33], we can summarize the data filtering based multi-innovation extended stochastic gradient (F-MI-ESG) algorithm as

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \frac{\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)}{r(t)}\boldsymbol{E}(p,t), \tag{12}$$

$$\boldsymbol{E}(p,t) = \hat{\boldsymbol{Y}}(p,t) - \hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)\hat{\boldsymbol{\theta}}(t-1), \tag{13}$$

$$r(t) = r(t-1) + \left\|\hat{\mathbf{\Phi}}_{\mathrm{f}}(p,t)\right\|^2, \quad r(0)=1, \tag{14}$$

$$\boldsymbol{Y}(p,t) = [y(t), y(t-1), \cdots, y(t-p+1)]^{\mathrm{T}}, \tag{15}$$

$$\hat{\mathbf{\Phi}}_{\mathrm{f}}(p,t) = [\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t), \hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1), \cdots, \hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)]^{\mathrm{T}}, \tag{16}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t) = \left[\hat{\boldsymbol{\phi}}_{1\mathrm{f}}^{\mathrm{T}}(t), \hat{\boldsymbol{\phi}}_{2\mathrm{f}}^{\mathrm{T}}(t), \cdots, \hat{\boldsymbol{\phi}}_{r\mathrm{f}}^{\mathrm{T}}(t), \hat{y}_{\mathrm{f}}(t-1), \hat{y}_{\mathrm{f}}(t-2), \cdots, \hat{y}_{\mathrm{f}}(t-n_d)\right]^{\mathrm{T}}, \tag{17}$$

$$\hat{\boldsymbol{\phi}}_{j\mathrm{f}}(t) = [-\hat{x}_{j\mathrm{f}}(t-1), \cdots, -\hat{x}_{j\mathrm{f}}(t-n_j), \hat{u}_{j\mathrm{f}}(t-1), \cdots, \hat{u}_{j\mathrm{f}}(t-n_j)]^{\mathrm{T}}, \tag{18}$$

$$\hat{y}_{\mathrm{f}}(t) = [-\hat{y}_{\mathrm{f}}(t-1), -\hat{y}_{\mathrm{f}}(t-2), \cdots, -\hat{y}_{\mathrm{f}}(t-n_d)]\,\hat{\boldsymbol{d}}(t) + y(t), \tag{19}$$

$$\hat{u}_{j\mathrm{f}}(t) = [-\hat{u}_{j\mathrm{f}}(t-1), -\hat{u}_{j\mathrm{f}}(t-2), \cdots, -\hat{u}_{j\mathrm{f}}(t-n_d)]\,\hat{\boldsymbol{d}}(t) + u_j(t), \tag{20}$$

$$\hat{x}_{j\mathrm{f}}(t) = \hat{\boldsymbol{\phi}}_{j\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}_j(t), \tag{21}$$

$$\hat{\boldsymbol{\theta}}(t) = \left[\hat{\boldsymbol{\vartheta}}_1^{\mathrm{T}}(t), \hat{\boldsymbol{\vartheta}}_2^{\mathrm{T}}(t), \cdots, \hat{\boldsymbol{\vartheta}}_r^{\mathrm{T}}(t), \hat{\boldsymbol{d}}^{\mathrm{T}}(t)\right]^{\mathrm{T}}. \tag{22}$$

When $p = 1$, we can obtain the F-ESG algorithm in (9)-(11). That is, the F-ESG algorithm is a special case of the F-MI-ESG algorithm. The steps of computing the parameter estimation vector $\hat{\boldsymbol{\theta}}(t)$ as t increases are as follows.

1. Initialize and choose p: let $t = 1$, $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_n/p_0$, $\hat{x}_{j\mathrm{f}}(i) = 1/p_0$, $\hat{u}_{j\mathrm{f}}(i) = 1/p_0$ and $\hat{y}_{\mathrm{f}}(i) = 1/p_0$ for $i \leq 0$, $p_0 = 10^6$.
2. Collect the input-output data $\{u_j(t), y(t)\colon j = 1, 2, \cdots, r\}$.
3. Form $\hat{\boldsymbol{\phi}}_{j\mathrm{f}}(t)$ by (18), $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$ by (17), and form $\boldsymbol{Y}(p,t)$ by (15), $\hat{\mathbf{\Phi}}_{\mathrm{f}}(p,t)$ by (16).
4. Compute $\boldsymbol{E}(p,t)$ and $r(t)$ using (13) and (14).
5. Update the parameter estimation vector $\hat{\boldsymbol{\theta}(t)}$ by (12).
6. Read $\hat{\boldsymbol{d}}(t)$ and $\hat{\boldsymbol{\vartheta}}_j(t)$ from $\hat{\boldsymbol{\theta}(t)}$ in (22).
7. Compute $\hat{u}_{j\mathrm{f}}(t)$, $\hat{y}_{\mathrm{f}}(t)$ and $\hat{x}_{j\mathrm{f}}(t)$ by (20), (19) and (21), respectively.
8. Increase t by 1 and go to Step 2.

4. **Example.**

Consider the following multi-input OEMA system:

$$\begin{aligned}
y(t) &= \frac{B_1(q)}{A_1(q)}u_1(t) + \frac{B_2(q)}{A_2(q)}u_2(t) + D(q)v(t),\\
A_1(q) &= 1 + a_{11}q^{-1} + a_{12}q^{-2} = 1 - 0.23q^{-1} - 0.12q^{-2},\\
B_1(q) &= b_{11}q^{-1} + b_{12}q^{-2} = 0.78q^{-1} + 0.45q^{-2},\\
A_2(q) &= 1 + a_{21}q^{-1} + a_{22}q^{-2} = 1 + 0.26q^{-1} + 0.26q^{-2},\\
B_2(q) &= b_{21}q^{-1} + b_{22}q^{-2} = 0.46q^{-1} + 1.00q^{-2},\\
D(q) &= 1 + d_1q^{-1} + d_2q^{-2} = 1 + 0.06q^{-1} + 0.08q^{-2}.
\end{aligned}$$

The parameter vector to be estimated is

$$\boldsymbol{\theta} = [a_{11}, a_{12}, b_{11}, b_{12}, a_{21}, a_{22}, b_{21}, b_{22}, d_1, d_2]^{\mathrm{T}} \tag{23}$$

$$= [-0.23, -0.12, 0.78, 0.45, 0.26, 0.26, 0.46, 1.00, 0.06, 0.08]^{\mathrm{T}}. \tag{24}$$

In simulation, the inputs $\{u_1(t), u_2(t)\}$ are taken as two uncorrelated stochastic signal sequences with zero mean and unit variance, and $\{v(t)\}$ as a white noise sequence with zero mean and variance $\sigma^2 = 0.10^2$. We apply the MI-ESG algorithm and the F-MI-ESG algorithm to estimating the parameters of this system. The parameter estimates and errors are shown in Table 1 with $p = 10$, and the parameter estimation errors δ versus t are shown in Figure 1 with $p = 1, 2, 5, 10$.

From the simulation results in Table 1 and Figure 1, we can draw the conclusions that the parameter estimates given by the F-MI-ESG algorithm have higher accuracy than those given by the MI-ESG algorithm in the same situation. Increasing the innovation

Table 1. The MI-ESG and F-MI-ESG parameter estimates and errors

Algorithms	t	a_{11}	a_{12}	b_{11}	b_{12}	a_{21}	a_{22}	b_{21}	b_{22}	d_1	d_2	δ (%)
MI-ESG	500	−0.2268	−0.1206	0.7822	0.4562	0.2646	0.2601	0.4565	0.9903	0.1259	−0.1828	18.1441
	1000	−0.2251	−0.1212	0.7776	0.4537	0.2598	0.2593	0.4590	0.9898	0.1253	−0.1793	17.9049
	2000	−0.2225	−0.1202	0.7784	0.4520	0.2594	0.2580	0.4599	0.9941	0.1248	−0.1773	17.7628
	3000	−0.2232	−0.1214	0.7789	0.4517	0.2585	0.2583	0.4612	0.9987	0.1245	−0.1761	17.6701
F-MI-ESG	500	−0.2038	−0.0974	0.7685	0.4495	0.2682	0.3076	0.4383	0.9383	0.0280	0.0269	7.2665
	1000	−0.2042	−0.1000	0.7707	0.4496	0.2717	0.3051	0.4483	0.9568	0.0252	0.0327	6.2654
	2000	−0.2019	−0.0987	0.7755	0.4494	0.2725	0.3045	0.4540	0.9709	0.0222	0.0320	5.9898
	3000	−0.2020	−0.0994	0.7772	0.4495	0.2716	0.3042	0.4562	0.9796	0.0233	0.0332	5.7072
True values		−0.2300	−0.1200	0.7800	0.4500	0.2600	0.2600	0.4600	1.0000	0.0600	0.0800	

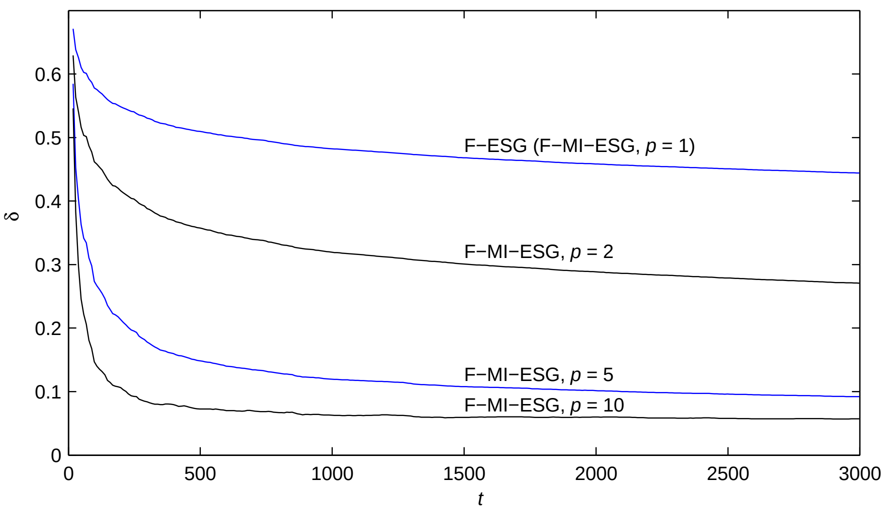

Figure 1. The estimation errors δ versus t

length p can improve the parameter estimation accuracy of the F-MI-ESG algorithm, that is we can obtain highly accurate estimates as the innovation length p increases.

5. **Conclusions.** By means of the filtering technique and the multi-innovation identification theory, a filtering based multi-innovation extended stochastic gradient algorithm is proposed. The simulation results show that the proposed F-MI-ESG algorithm provides more accurate parameter estimates than the MI-ESG algorithm for the same innovation length. In the future, the proposed algorithm in this paper can be extended to nonlinear models with colored noise, state space systems, sensor networks and feed back nonlinear systems.

Acknowledgment. This work was supported by the Visiting Scholar Project for Young Backbone Teachers of Shandong Province, China. This work was completed while the author was a visiting scholar in the School of Internet of Things Engineering at Jiangnan University (Wuxi, China) from September 2015 to July 2016. The author is grateful to her supervisor Professor Feng Ding for his helpful suggestions and the main idea of this work comes from him and his book "System Identification – New Theory and Methods, Beijing: Science Press, 2013".

REFERENCES

[1] F. Ding, *System Identification – New Theory and Methods*, Science Press, Beijing, 2013.
[2] F. Ding, *System Identification – Performances Analysis for Identification Methods*, Science Press, Beijing, 2014.

[3] F. Ding, *System Identification – Multi-Innovation Identification Theory and Methods*, Science Press, Beijing, 2016.

[4] F. Ikhouane and F. Giri, A unified approach for the parametric identification of SISO/MIMO Wiener and Hammerstein systems, *Journal of the Franklin Institute*, vol.351, no.3, pp.1717-1727, 2014.

[5] S. Yerramilli and A. K. Tangirala, Detection and diagnosis of model-plant mismatch in MIMO systems using plant-model ratio, *IFAC-PapersOnLine*, vol.49, pp.266-271, 2016.

[6] L. Xu, A proportional differential control method for a time-delay system using the Taylor expansion approximation, *Applied Mathematics and Computation*, vol.236, pp.391-399, 2014.

[7] S. S. Saab and R. Toukhtarian, A MIMO sampling-rate-dependent controller, *IEE Trans. Industrial Electronics*, vol.62, no.6, pp.3662-3671, 2015.

[8] Y. J. Wang and F. Ding, The filtering based iterative identification for multivariable systems, *IET Control Theory and Applications*, vol.10, no.8, pp.894-902, 2016.

[9] D. Q. Wang and F. Ding, Parameter estimation algorithms for multivariable Hammerstein CARMA systems, *Information Sciences*, vol.355, pp.237-248, 2016.

[10] A. Nasirin, S. K. Nguang and A. Swain, Adaptive sliding mode control for a class of MIMO nonlinear systems with uncertainties, *Journal of the Franklin Institute*, vol.351, pp.2048-2061, 2014.

[11] G. Tao, Multivariable adaptive control: A survey, *Automatica*, vol.50, no.12, pp.2737-2764, 2014.

[12] S. Mobayen, An LMI-based robust tracker for uncertain linear systems with multiple time-varying delays using optimal composite nonlinear feedback technique, *Nonlinear Dynamics*, vol.80, pp.917-927, 2015.

[13] L. Xu, Application of the Newton iteration algorithm to the parameter estimation for dynamical systems, *Journal of Computational and Applied Mathematics*, vol.288, pp.33-43, 2015.

[14] D. Q. Wang, Hierarchical parameter estimation for a class of MIMO Hammerstein systems based on the reframed models, *Applied Mathematics Letters*, vol.57, pp.13-19, 2016.

[15] F. Ding, X. M. Liu and M. M. Liu, The recursive least squares identification algorithm for a class of Wiener nonlinear systems, *Journal of the Franklin Institute*, vol.353, no.7, pp.1518-1526, 2016.

[16] F. Ding and Q. J. Chen, Recursive least squares identification methods for output nonlinear equation-error type systems, *Journal of Nanjing University of Information Science and Technology (Natural Science Edition)*, vol.7, no.3, pp.193-213, 2015.

[17] D. Q. Wang and W. Zhang, Improved least squares identification algorithm for multivariable Hammerstein systems, *Journal of the Franklin Institute*, vol.352, no.11, pp.5292-5307, 2015.

[18] F. Ding, Auxiliary model based identification methods – Part A: Autoregressive output-error systems, *Journal of Nanjing University of Information Science and Technology (Natural Science Edition)*, vol.8, no.1, pp.1-22, 2016.

[19] F. Ding and H. B. Chen, Auxiliary model based identification methods – Part B: Input nonlinear output-error systems, *Journal of Nanjing University of Information Science and Technology (Natural Science Edition)*, vol.8, no.2, pp.97-115, 2016.

[20] F. Ding and H. B. Chen, Auxiliary model based identification methods – Part C: Input nonlinear output-error autoregressive systems, *Journal of Nanjing University of Information Science and Technology (Natural Science Edition)*, vol.8, no.3, pp.193-214, 2016.

[21] J. L. Ding, Data filtering based recursive and iterative least squares algorithms for parameter estimation of multi-input output systems, *Algorithms*, vol.9, 2016.

[22] L. Xu, L. Chen and W. L. Xiong, Parameter estimation and controller design for dynamic systems from the step responses based on the Newton iteration, *Nonlinear Dynamics*, vol.79, no.3, pp.2155-2163, 2015.

[23] L. Xu, The damping iterative parameter identification method for dynamical systems based on the sine signal measurement, *Signal Processing*, vol.120, pp.660-667, 2016.

[24] Y. W. Mao and F. Ding, A novel parameter separation based identification algorithm for Hammerstein systems, *Applied Mathematics Letters*, vol.60, pp.21-27, 2016.

[25] F. Ding and L. J. Guo, Multi-innovation identification methods for linear-parameter systems, *Journal of Nanjing University of Information Science and Technology (Natural Science Edition)*, vol.7, no.4, pp.289-312, 2015.

[26] F. Ding, Multi-innovation identification methods for equation-error systems, *Journal of Nanjing University of Information Science and Technology (Natural Science Edition)*, vol.7, no.5, pp.385-407, 2015.

[27] F. Ding, Multi-innovation identification methods for output-error systems, *Journal of Nanjing University of Information Science and Technology (Natural Science Edition)*, vol.7, no.6, pp.481-503, 2015.

[28] Q. Y. Shen and F. Ding, Hierarchical multi-innovation extended stochastic gradient algorithms for input nonlinear multivariable OEMA systems by the key-term separation principle, *Nonlinear Dynamics*, vol.85, no.1, pp.499-507, 2016.
[29] F. Ding and Y. W. Mao, Multi-innovation identification methods for input nonlinear equation-error autoregressive systems, *Journal of Nanjing University of Information Science and Technology (Natural Science Edition)*, vol.7, no.1, pp.1-23, 2015.
[30] F. Ding and H. B. Chen, Multi-innovation identification methods for input nonlinear equation-error systems, *Journal of Nanjing University of Information Science and Technology (Natural Science Edition)*, vol.7, no.2, pp.97-124, 2015.
[31] Y. J. Wang and F. Ding, Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model, *Automatica*, vol.71, pp.308-313, 2016.
[32] F. Ding, X. M. Liu and Y. Gu, An auxiliary model based least squares algorithm for a dual-rate state space system with time-delay using the data filtering, *Journal of the Franklin Institute*, vol.353, no.2, pp.398-408, 2016.
[33] X. H. Wang and F. Ding, Convergence of the auxiliary model based multi-innovation generalized extended stochastic gradient algorithm for Box-Jenkins systems, *Nonlinear Dynamics*, vol.82, nos.1-2, pp.269-280, 2015.