International Journal of Control, Automation, and Systems 21(2) (2023) 553-562
http://dx.doi.org/10.1007/s12555-021-0572-4

ISSN:1598-6446 eISSN:2005-4092
http://www.springer.com/12555

Robust Gradient Estimation Algorithm for a Stochastic System with Colored Noise

Wentao Liu and Weili Xiong*

Abstract: This paper studies the parameter estimation algorithms of a finite impulse response system with colored noise. To suppress the negative effects of the colored noises, a novel gradient-based algorithm is developed by means of the cost function of the continuous mixed p-norm (CMPN). It combines the p-norms for $1 \leqslant p \leqslant 2$, which control the proportions of the error norms and generate an adjustable gain to adapt the data quality. Moreover, to improve the convergence rate, a CMPN multi-innovation gradient recursive algorithm is derived through expanding the innovation scalar to the innovation vector. Finally, two examples are given to demonstrate the validity of the proposed algorithms.

Keywords: Continuous mixed p-norm, gradient search, multi-innovation identification, parameter estimation.

1. INTRODUCTION

The dynamic behaviors of many physical plants are modeled as stochastic systems in the vicinity of a specific operating point. Thus the identification methods for stochastic systems have received much attention for decades [1-5]. Typical stochastic systems include finite impulse response system, controlled autoregressive moving average system [6] and output error moving average system to name a few. Among them, the finite impulse response system is crucial in signal processing, filter designing and dynamical system modeling. Many identification methods and parameter estimation algorithms have been developed for stochastic systems [7-10].

The stochastic gradient (SG) algorithm is an effective identification method owing to its computational convenience therein. Recently, many gradient-based algorithms have been applied for parameter estimation. For instance, Chen modified the traditional SG algorithms by combining the polynomial transformation technique and thus proposed a polynomial transformation SG algorithm for dual-rate sampled systems [11]. However, these contributions are developed for the stochastic control systems in the presence of white noise and cannot be applied directly for the systems contaminated by colored noises.

For the stochastic control systems, noise interference, contaminates the measurement outputs of the systems resulting in low identification accuracy [12-15]. To weaken the colored noises interference and improve parameter estimation accuracy, an effective method is to apply adaptive filter algorithms against noise interference such as the least mean square algorithm and the normalized algorithm [16,17]. Zhang and Ding presented an optimal adaptive filtering algorithm by using the fractional-order derivative [18]. Navia-Vazquez *et al*. applied the combinations of two recursive least p-norm algorithms for the impulse response systems with non-Gaussian noise by adaptively minimizing the l_p norm [19]. Zheng *et al*. presented a variable step-size method for the finite-variance impulsive systems by using the least mean p-th norm algorithm, where the cost function was selected as the p-th order moment of errors [20]. Zayyani generalized the mixed norm defined in the robust mixed norm algorithm to a continuous mixed p-norm (CMPN) and thus developed a CMPN adaptive filtering algorithm for system identification [21]. Another way is introducing the multi-innovation to enhance the performance of the convergence rates. The innovation is the valuable information. The key of multi-innovation theory is to extend the innovation vector to the innovation matrix [22-24]. Ding *et al*. presented a filtering based multi-innovation identification method for multivariate output-error systems [25]. Other identification methods can be found in [26-29].

Inspired by the above researches, this paper proposes the parameter estimation algorithms for a finite impulse response system with colored noise through the CMPN algorithm and the multi-innovation identification theory. For one thing, the proposed algorithms control the proportions

Manuscript received July 8, 2021; revised March 8, 2022; accepted April 7, 2022. Recommended by Associate Editor Jun Moon under the direction of Editor Jay H. Lee. This work was supported by the National Natural Science Foundation of China (No. 61773182) and the 111 Project (B12018).

Wentao Liu and Weili Xiong are with the Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), the School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China (e-mails: wtliu12@126.com, weili_xiong@jiangnan.edu.cn).
* Corresponding author.

of the error norms and offer an extra degree of freedom within the adaptation. By taking into consideration each p-norm of errors for $1 \leqslant p \leqslant 2$, the proposed algorithms combine the benefits of the variable error norms and thus are more robust against colored noise interference. For another thing, because the multi-innovation algorithms use not only the current data but also the past data and thus can make full use of input-output data, the parameter estimation accuracy can be improved. The main contributions of this paper are as follows;

1) The mixed p-norm-based extended stochastic gradient algorithm is presented for a stochastic system to improve performance in the presence of colored noise by minimizing a continuous mixed p-norm cost function.
2) The CMPN multi-innovation gradient-based recursive algorithm is derived for a stochastic system with colored noise to improve the estimation accuracy by means of the multi-innovation identification theory.

The rest of this paper is organized as follows: Sections 2 and 3 derive the continuous mixed p-norm multi-innovation extended stochastic gradient (ESG) algorithm and the continuous mixed p-norm multi-innovation extended stochastic gradient (MIESG) algorithm respectively. Section 4 gives the simulation examples to illustrate the effectiveness of the proposed algorithms. Section 5 shows some concluding remarks.

2. CONTINUOUS MIXED P-NORM ESG ALGORITHM

Consider the following stochastic control system with colored noise disturbance described by finite impulse response moving average (FIR-MA) model,

$$y(t)=B(z)u(t)+D(z)v(t), \tag{1}$$

where $\{u(t)\}$ and $\{y(t)\}$ are the input and output of the system respectively, $\{v(t)\}$ is independent and identically distributed random noise with zero mean and variance σ^2, the polynomials $B(z)$ and $D(z)$ are the functions in the unit backward shift operator z^{-1}

$$B(z):=b_1z^{-1}+b_2z^{-2}+\cdots+b_{n_b}z^{-n_b},$$
$$D(z):=1+d_1z^{-1}+d_2z^{-2}+\cdots+d_{n_d}z^{-n_d}.$$

The information vector $\varphi(t)$ and the parameter vector ϑ are defined as

$$\varphi(t):=\begin{bmatrix}\kappa(t)\\ \psi(t)\end{bmatrix}\in\mathbb{R}^n,\quad n:=n_b+n_d,$$
$$\kappa(t):=[u(t-1),\ u(t-2),\ \cdots,\ u(t-n_b)]^{\mathrm{T}}\in\mathbb{R}^{n_b},$$
$$\psi(t):=[v(t-1),\ v(t-2),\ \cdots,\ v(t-n_d)]^{\mathrm{T}}\in\mathbb{R}^{n_d},$$
$$\vartheta:=\begin{bmatrix}b\\ d\end{bmatrix}\in\mathbb{R}^n,$$
$$b:=[b_1,\ b_2,\ \cdots,\ b_{n_b}]^{\mathrm{T}}\in\mathbb{R}^{n_b},$$
$$d:=[d_1,\ d_2,\ \cdots,\ d_{n_d}]^{\mathrm{T}}\in\mathbb{R}^{n_d}.$$

Then equation (1) can be rewritten as

$$y(t)=\varphi^{\mathrm{T}}(t)\vartheta+v(t). \tag{2}$$

The proposed parameter estimation algorithms in this paper are based on the identification model in (2) of the FIR-MA system in (1). Many identification methods are derived based on the identification models of the systems [30-34] and these methods can be used to estimate the parameters of other linear systems and nonlinear systems [35-39] and can be applied to other fields [40-45] such as chemical process control systems.

Here, $\{u(t),y(t)\}$ is the available observation data and $\{v(t)\}$ is the unmeasurable noise. Assume that the orders n_b and n_d are known. The objective is to present efficient identification algorithms with robustness for estimating the parameter vector ϑ from measurements $\{u(t),y(t)\}$.

To suppress the effect of the colored noises and to provide the robust parameter approach, define the continuous mixed p-norm cost function

$$J_1(\vartheta):=\int_1^2\lambda_t(p)|v(t)|^p\mathrm{d}p, \tag{3}$$

where $\lambda_t(p)$ is the probability density-like weighting function which is constrained by $\int_1^2\lambda_t(p)\mathrm{d}p=1$, and $v(t)=y(t)-\varphi^{\mathrm{T}}(t)\vartheta$. Taking the gradient of $J_1(\vartheta)$ with respect to b and d gives

$$\mathrm{grad}[J_1(\vartheta)]=\frac{\partial J_1(\vartheta)}{\partial\vartheta}=\begin{bmatrix}\dfrac{\partial J_1(\vartheta)}{\partial b}\\ \dfrac{\partial J_1(\vartheta)}{\partial d}\end{bmatrix}, \tag{4}$$

$$\begin{aligned}\frac{\partial J_1(\vartheta)}{\partial b}&=\int_1^2 p\lambda_t(p)|v(t)|^{p-1}\mathrm{sgn}(v(t))\\&\quad\times\frac{\partial(y(t)-\kappa^{\mathrm{T}}(t)b-\psi^{\mathrm{T}}(t)d)}{\partial b}\mathrm{d}p\\&=-\int_1^2 p\lambda_t(p)|v(t)|^{p-1}\mathrm{sgn}(v(t))\kappa(t)\mathrm{d}p,\\ \frac{\partial J_1(\vartheta)}{\partial d}&=\int_1^2 p\lambda_t(p)|v(t)|^{p-1}\mathrm{sgn}(v(t))\\&\quad\times\frac{\partial(y(t)-\kappa^{\mathrm{T}}(t)b-\psi^{\mathrm{T}}(t)d)}{\partial d}\mathrm{d}p\\&=-\int_1^2 p\lambda_t(p)|v(t)|^{p-1}\mathrm{sgn}(v(t))\psi(t)\mathrm{d}p.\end{aligned}$$

Therefore, (4) can be expressed as

$$\begin{aligned}\mathrm{grad}[J_1(\vartheta)]&=-\int_1^2 p\lambda_t(p)|v(t)|^{p-1}\mathrm{sgn}(v(t))\varphi(t)\mathrm{d}p\\&=-\gamma(t)\mathrm{sgn}(v(t))\varphi(t),\end{aligned}$$
$$\gamma(t):=\int_1^2 p\lambda_t(p)|v(t)|^{p-1}\mathrm{d}p. \tag{5}$$

Obviously, the information vector $\varphi(t)$ contains the unmeasurable noise $v(t-i)$. The solution is to replace the unmeasurable noise $v(t-i)$ with the identified innovation $e(t-i)$. The estimate of $\varphi(t)$ defined by the measurable input $u(t-i)$ and the innovation $e(t-i)$ is

$$\begin{aligned}\hat{\varphi}(t) :=& [u(t-1),\ u(t-2),\ \cdots,\ u(t-n_b),\\ & e(t-1),\ e(t-2),\ \cdots,\ e(t-n_d)]^{\mathrm{T}} \in \mathbb{R}^n,\end{aligned}$$

where the innovation $e(t)$ is defined as

$$e(t) := y(t) - \hat{\varphi}^{\mathrm{T}}(t)\hat{\vartheta}(t-1).$$

Using the negative gradient search, the cost function $J_1(\vartheta)$ is minimized to get

$$\begin{aligned}\hat{\vartheta}(t) &= \hat{\vartheta}(t-1) - \mu(t)\mathrm{grad}[J_1(\hat{\vartheta})]\\ &= \hat{\vartheta}(t-1) + \mu(t)\hat{\gamma}(t)\mathrm{sgn}(e(t))\hat{\varphi}(t), \qquad (6)\end{aligned}$$

where $\mu(t) \geqslant 0$ is the step size, and the estimate of $\gamma(t)$ is

$$\hat{\gamma}(t) := \int_1^2 p\lambda_t(p)|e(t)|^{p-1}\mathrm{d}p. \qquad (7)$$

$\lambda_t(p)$ is a probability density-like weighting function with the constraint $\int_1^2 \lambda_t(p) = 1$. To obtain a closed form formula for $\hat{\gamma}(t)$, a uniform weighting function $\lambda_t(p) = 1$ is assumed. While $\lambda_t(p) = 1$, the constraint is met and $\hat{\gamma}(t)$ can be computed by

$$\begin{aligned}\hat{\gamma}(t) &= \int_1^2 p\lambda_t(p)|e(t)|^{p-1}\mathrm{d}p\\ &= \frac{1}{ln|e(t)|}\int_1^2 p\mathrm{d}|e(t)|^{p-1}\\ &= \frac{p|e(t)|^{p-1}|_1^2}{ln|e(t)|} - \frac{\int_1^2 |e(t)|^{p-1}\mathrm{d}p}{ln|e(t)|}\\ &= \frac{(2|e(t)|-1)}{ln|e(t)|} - \frac{(|e(t)|-1)}{\ln^2(|e(t)|)}\\ &= \frac{(2|e(t)|-1)\ln(|e(t)|) - |e(t)| + 1}{\ln^2(|e(t)|)}. \qquad (8)\end{aligned}$$

Generally speaking, the estimate $\hat{\vartheta}(t)$ approaches the true value of ϑ as t increases, and the innovation $e(t)$ may be close to zero. To avoid division by zero, Equation (8) can be modified as

$$\hat{\gamma}(t) = \frac{(2|e(t)|-1)\ln(|e(t)|) - |e(t)| + 1}{\ln^2(|e(t)|) + 1}. \qquad (9)$$

Substituting $\vartheta = \hat{\vartheta}(t)$ into (3) gives

$$\begin{aligned}J_1(\hat{\vartheta}(t)) &= \int_1^2 \lambda_t(p)|y(t) - \hat{\varphi}^{\mathrm{T}}(t)\hat{\vartheta}(t)|^p\mathrm{d}p\\ &= \int_1^2 \lambda_t(p)|y(t) - \hat{\varphi}^{\mathrm{T}}(t)[\hat{\vartheta}(t-1)\\ &\quad + \mu(t)\hat{\gamma}(t)\mathrm{sgn}(e(t))\hat{\varphi}(t)]|^p\mathrm{d}p\\ &= \int_1^2 \lambda_t(p)|e(t) - \mu(t)\hat{\gamma}(t)\\ &\quad \times \mathrm{sgn}(e(t))\|\hat{\varphi}(t)\|^2|^p\mathrm{d}p. \qquad (10)\end{aligned}$$

Inserting (7) into (10) gives

$$\begin{aligned}J_1(\hat{\vartheta}(t)) &= \int_1^2 \lambda_t(p)|e(t) - \mu(t)\int_1^2 p\lambda_t(p)|e(t)|^{p-1}\mathrm{d}p\\ &\quad \times \mathrm{sgn}(e(t))\|\hat{\varphi}(t)\|^2|^p\mathrm{d}p\\ &= \int_1^2 |e(t) - \mu(t)\int_1^2 p|e(t)|^{p-1}\mathrm{d}p\\ &\quad \times \mathrm{sgn}(e(t))\|\hat{\varphi}(t)\|^2|^p\mathrm{d}p\\ &= \int_1^2 |e(t)[1 - \mu(t)\int_1^2 p|e(t)|^{p-2}\mathrm{d}p\\ &\quad \times \|\hat{\varphi}(t)\|^2]|^p\mathrm{d}p\\ &= \int_1^2 |e(t)[1 - \mu(t)\rho(t)\|\hat{\varphi}(t)\|^2]|^p\mathrm{d}p. \qquad (11)\end{aligned}$$

To guarantee the convergence of the algorithm, the optimal step-size can be modified as

$$\mu(t) := \frac{1}{r(t)},\ r(t) = r(t-1) + \|\hat{\varphi}(t)\|^2. \qquad (12)$$

For the recursive identification algorithm, the parameter estimates are expected to approach their true values of the identified system continuously with the recursive variable t increasing. This behavior is called the convergence of the identification algorithms. In this paper, the modified step size $\mu(t) > 0$ can ensure the stability and convergence of the proposed algorithm. Thus the continuous mixed p-norm extended stochastic gradient (CMPN-ESG) algorithm (13)-(18) for the FIR-MA system is summarized in the following:

$$\hat{\vartheta}(t) = \hat{\vartheta}(t-1) + \frac{1}{r(t)}\hat{\gamma}(t)\mathrm{sgn}(e(t))\hat{\varphi}(t), \qquad (13)$$

$$\begin{aligned}\hat{\varphi}(t) =& [u(t-1),\ u(t-2),\ \cdots,\ u(t-n_b),\\ & e(t-1),\ e(t-2),\ \cdots,\ e(t-n_d)]^{\mathrm{T}}, \qquad (14)\end{aligned}$$

$$r(t) = r(t-1) + \|\hat{\varphi}(t)\|^2, \qquad (15)$$

$$\hat{\gamma}(t) = \frac{(2|e(t)|-1)\ln(|e(t)|) - |e(t)| + 1}{\ln^2(|e(t)|) + 1}, \qquad (16)$$

$$\hat{\vartheta}(t) = [\hat{b}_1(t),\ \cdots,\ \hat{b}_{n_b}(t),\ \hat{d}_1(t),\ \cdots,\ \hat{d}_{n_d}(t)]^{\mathrm{T}}, \qquad (17)$$

$$e(t) = y(t) - \hat{\varphi}^{\mathrm{T}}(t)\hat{\vartheta}(t-1). \qquad (18)$$

The steps of the CMPN-ESG algorithm for the FIR-MA system to compute $\hat{\vartheta}(t)$ are as follows:

1) Initialization: Let $t = 1$, set the data length L and some initial values be $\hat{\vartheta}_0 = \mathbf{1}_n/p_0$, $r_0 = 1$, $u(t-i) = 1/p_0$, $y(t-i) = 1/p_0$, $e(t-i) = 1/p_0$, $i = 1, 2, \cdots, n_d$, $p_0 = 10^6$.
2) Collect the input and output data $u(t)$ and $y(t)$.
3) Form the information vectors $\hat{\varphi}(t)$ using (14).

4) Compute $r(t)$ and $\hat{\gamma}(t)$ using (15) and (16), respectively.
5) Update the parameter estimation vector $\hat{\vartheta}(t)$ using (13).
6) Compute the innovation $e(t)$ using (18).
7) If $t < L$, increase t by 1 and go to Step 2; else obtain the parameter estimate $\hat{\vartheta}(L)$, and finish the computation process.

Remark 1: The CMPN-ESG algorithm integrates the noise regression terms and the parameters d into the information matrix and the parameter vectors ϑ, respectively.

Remark 2: The cost function $J_1(\vartheta)$ is a continuous mixed p-norm function of the error. Since it combines p-norms for $1 \leqslant p \leqslant 2$, and it combines the advantages of different error norms.

Remark 3: The CMPN-ESG algorithm controls the proportions of the error norms and offers an extra degree of freedom within the adaptation and thereby it has higher parameter estimation accuracy compared with the ESG algorithms.

Remark 4: The CMPN-ESG algorithm is a recursive algorithm which updates the estimates along the negative gradient optimization. However, each recursion of the CMPN-ESG algorithm only contains the data at the current moment. Generally speaking, the more observation data the algorithm uses, the higher parameter estimation accuracy the algorithm will get. Therefore, the multi-innovation principle will be introduced to make full use of input-output data in Section 3.

3. CONTINUOUS MIXED P-NORM MIESG ALGORITHM

Based on the CMPN-ESG algorithm in (13)-(18), we introduce the multi-innovation theory to enhance the performance by making full use of data information in this section. Note that the innovation $e(t)$ is the valuable information that can improve the parameter estimation accuracy for identification algorithms.

Extend the innovation scalar $e(t)$ to a large innovation vector

$$E(l,t) = Y(l,t) - \hat{\Phi}^{\mathrm{T}}(l,t)\hat{\vartheta}(t-1),$$

where

$$\hat{\Phi}(l,t) := [\hat{\varphi}(t),\ \hat{\varphi}(t-1),\ \cdots,\ \hat{\varphi}(t-l+1)] \in \mathbb{R}^{n\times l},$$
$$Y(l,t) := [y(t),\ y(t-1),\ \cdots,\ y(t-l+1)]^{\mathrm{T}} \in \mathbb{R}^{l}.$$

Consider the measurements from $t-l+1$ to t and define the cost function

$$J_2(\vartheta) := \sum_{j=0}^{l-1}\int_1^2 \lambda_t(p)|v(t-j)|^p \mathrm{d}p. \tag{19}$$

Taking the gradient of $J_2(\vartheta)$ gives

$$\mathrm{grad}[J_2(\vartheta)] = \frac{\partial J_2(\vartheta)}{\partial \vartheta} = \begin{bmatrix} \dfrac{\partial J_2(\vartheta)}{\partial b} \\ \dfrac{\partial J_2(\vartheta)}{\partial d} \end{bmatrix}, \tag{20}$$

$$\frac{\partial J_2(\vartheta)}{\partial b} = -\sum_{j=0}^{l-1}\int_1^2 p\lambda_t(p)|v(t-j)|^{p-1} \times \mathrm{sgn}(v(t-j))\kappa(t-j)\mathrm{d}p,$$

$$\frac{\partial J_2(\vartheta)}{\partial d} = -\sum_{j=0}^{l-1}\int_1^2 p\lambda_t(p)|v(t-j)|^{p-1} \times \mathrm{sgn}(v(t-j))\psi(t-j)\mathrm{d}p.$$

Therefore, (20) can be expressed as

$$\begin{aligned}\mathrm{grad}[J_2(\vartheta(t))] &= -\sum_{j=0}^{l-1}\int_1^2 p\lambda_t(p)|v(t-j)|^{p-1} \\ &\quad \times \mathrm{sgn}(v(t-j))\varphi(t-j)\mathrm{d}p, \\ &= -\sum_{j=0}^{l-1}\gamma(t-j)\mathrm{sgn}(v(t-j))\varphi(t-j),\end{aligned}$$

$$\gamma(t-j) := \int_1^2 p\lambda_t(p)|v(t-j)|^{p-1}\mathrm{d}p.$$

Similar to the derivation of the CMPN-ESG algorithm, using the negative search and minimizing $J_2(\vartheta)$ yield

$$\begin{aligned}\hat{\vartheta}(t) &= \hat{\vartheta}(t-1) - \frac{1}{r(t)}\mathrm{grad}[J_2(\hat{\vartheta}(t))] \\ &= \hat{\vartheta}(t-1) + \frac{1}{r(t)}\hat{\Phi}(l,t)\Gamma(l,t),\end{aligned} \tag{21}$$

$$\Gamma(l,t) := \begin{bmatrix} \mathrm{sgn}(e(t))\hat{\gamma}(t) \\ \mathrm{sgn}(e(t-1))\hat{\gamma}(t-1) \\ \vdots \\ \mathrm{sgn}(e(t-l-1))\hat{\gamma}(t-l+1) \end{bmatrix} \in \mathbb{R}^{l}.$$

Thus, the continuous mixed p-norm multi-innovation extended stochastic gradient (CMPN-MIESG) algorithm for the FIR-MA system is summarized as follows:

$$\hat{\vartheta}(t) = \hat{\vartheta}(t-1) + \frac{1}{r(t)}\hat{\Phi}(l,t)\Gamma(l,t), \tag{22}$$

$$\hat{\gamma}(t) = \frac{(2|e(t)|-1)\ln(|e(t)|) - |e(t)| + 1}{\ln^2(|e(t)|)+1}, \tag{23}$$

$$e(t) = y(t) - \hat{\varphi}^{\mathrm{T}}(t)\hat{\vartheta}(t-1), \tag{24}$$

$$r(t) = r(t-1) + \|\hat{\varphi}(t)\|^2, \tag{25}$$

$$\begin{aligned}\hat{\varphi}(t) = [&u(t-1),\ u(t-2),\ \cdots,\ u(t-n_b), \\ &e(t-1),\ e(t-2),\ \cdots,\ e(t-n_d)]^{\mathrm{T}},\end{aligned} \tag{26}$$

$$\hat{\Phi}(l,t) = [\hat{\varphi}(t),\ \hat{\varphi}(t-1),\ \cdots,\ \hat{\varphi}(t-l+1)], \tag{27}$$

$$Y(l,t) = [y(t),\ y(t-1),\ \cdots,\ y(t-l+1)]^{\mathrm{T}}, \tag{28}$$

$$E(l,t) = Y(l,t) - \hat{\Phi}^{\mathrm{T}}(l,t)\hat{\vartheta}(t-1), \tag{29}$$

$$\Gamma(l,t) = [\mathrm{sgn}(e(t))\hat{\gamma}(t),\ \mathrm{sgn}(e(t-1))\hat{\gamma}(t-1),\ \cdots,\ \mathrm{sgn}(e(t-l-1))\hat{\gamma}(t-l+1)]^{\mathrm{T}}. \quad (30)$$

The proposed algorithms in this paper can joint other identification methods [46-49] to new parameter estimation approaches of linear and nonlinear systems [50-54] and can be applied to other literature [55-58] such as engineering application systems.

Remark 5: Not only does the CMPN-MIESG algorithm utilize the current data $\{y(t), \hat{\varphi}(t)\}$ and the innovation $e(t)$, but also utilizes the past data $\{y(t-j), \hat{\varphi}(t-j), j = 1, 2, \cdots, l-1\}$ and the innovation $e(t-j)$, which improves the estimation accuracy by using the observations repeatedly. When $l = 1$, the CMPN-MIESG algorithm reduces to the the CMPN-ESG algorithm.

4. EXAMPLES

Example 1: Consider the following FIR-MA system,

$$\begin{aligned} y(t) &= B(z)u(t) + D(z)v(t), \\ B(z) &= b_1 z^{-1} + b_2 z^{-2} = 1.68z^{-1} + 2.32z^{-2}, \\ D(z) &= 1 + d_1 z^{-1} + d_2 z^{-2} = 1 - 0.75z^{-1} + 0.75z^{-2}. \end{aligned}$$

The parameter vector to be estimated is

$$\begin{aligned} \vartheta &:= [b_1,\ b_2,\ d_1,\ d_2]^{\mathrm{T}} \\ &= [1.68,\ 2.32,\ -0.75,\ 0.75]^{\mathrm{T}}. \end{aligned}$$

In simulation, the input $\{u(t)\}$ is taken as a persistent excitation signal sequence, $\{v(t)\}$ is taken as a normal distribution white noise sequence with zero mean and variance σ^2. Take the data length $L = 3000$ and noise variances $\sigma^2 = 1.00^2$, $\sigma^2 = 2.00^2$, $\sigma^2 = 3.00^2$, $\sigma^2 = 4.00^2$ and $\sigma^2 = 5.00^2$ respectively. The ESG algorithm and the CMPN-ESG algorithm are applied to estimate the parameters of the FIR-MA system under the different noise variances. The parameter estimates and errors of the ESG algorithm and the CMPN-ESG algorithm are illustrated in Tables 1 and 2. The parameter estimation errors $\delta := \|\hat{\vartheta}(t) - \vartheta\|/\|\vartheta\|$ versus t under different σ^2 of the ESG algorithm and the CMPN-ESG algorithm are depicted in Figs. 1 and 2.

From Tables 1-2 and Figs. 1-2, the following conclusions can be drawn.

1) Although the estimation results for a few of parameters are not quite good, the total estimation errors δ decay as t increases, and both the ESG algorithm and the CMPN-ESG algorithm are effective for identifying the example system.
2) The parameter estimation errors decrease with the noise-to-signal ratio decreases.

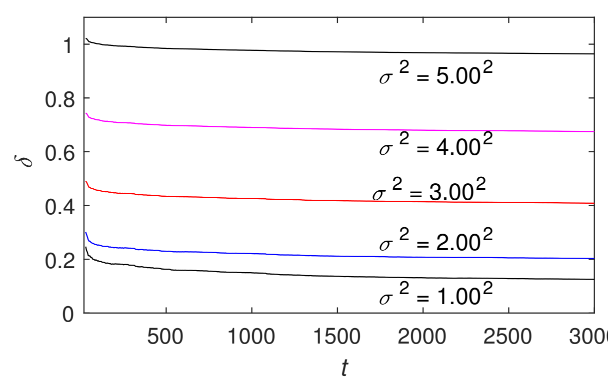

Fig. 1. The ESG estimation errors δ versus t with different σ^2 for Example 1.

Table 1. The ESG estimates and their errors with different σ^2 for Example 1.

σ^2	t	b_1	b_2	d_1	d_2	δ (%)
5.00^2	100	1.99842	5.34353	-0.52347	0.49176	100.16781
	200	1.99410	5.31753	-0.51759	0.45406	99.43924
	500	1.99238	5.28463	-0.51165	0.44000	98.43195
	2000	1.98830	5.23800	-0.53122	0.44583	96.84372
	3000	1.98695	5.22394	-0.53091	0.44052	96.40406
3.00^2	100	1.51106	3.67966	-0.62393	0.51990	45.67095
	200	1.51235	3.64655	-0.62242	0.50235	44.71509
	500	1.52179	3.60670	-0.61735	0.50149	43.43261
	2000	1.53114	3.55100	-0.63748	0.52989	41.39400
	3000	1.53346	3.53418	-0.63721	0.52743	40.86429
1.00^2	100	1.36784	2.45719	-0.61496	0.28641	19.35245
	200	1.39858	2.45700	-0.63433	0.30314	18.25852
	500	1.46033	2.46708	-0.63882	0.34305	16.29922
	2000	1.51613	2.46379	-0.66056	0.42638	13.10616
	3000	1.53049	2.45984	-0.66214	0.43889	12.52748
True values		1.68000	2.32000	-0.75000	0.75000	

Table 2. The CMPN-ESG estimates and their errors with different σ^2 for Example 1.

σ^2	t	b_1	b_2	d_1	d_2	δ (%)
5.00^2	100	1.32497	3.47429	-0.65817	0.47185	40.68378
	200	1.32545	3.46685	-0.67387	0.48311	40.33665
	500	1.33200	3.45890	-0.67823	0.50855	39.85073
	2000	1.33793	3.44580	-0.69489	0.56684	39.02691
	3000	1.33966	3.44142	-0.69425	0.57068	38.85698
3.00^2	100	1.12655	2.73123	-0.69929	0.49221	24.15684
	200	1.13471	2.72809	-0.71065	0.49828	23.80706
	500	1.15939	2.73201	-0.70517	0.52446	23.00267
	2000	1.18533	2.73150	-0.71826	0.59324	21.70708
	3000	1.19308	2.73025	-0.71560	0.59665	21.47065
1.00^2	100	1.23308	2.20961	-0.54105	0.27084	22.80429
	200	1.28335	2.25088	-0.56516	0.28296	21.07542
	500	1.38028	2.31402	-0.56911	0.32244	18.09240
	2000	1.46645	2.36305	-0.61503	0.41731	13.74876
	3000	1.48941	2.37087	-0.62138	0.43051	12.99417
True values		1.68000	2.32000	-0.75000	0.75000	

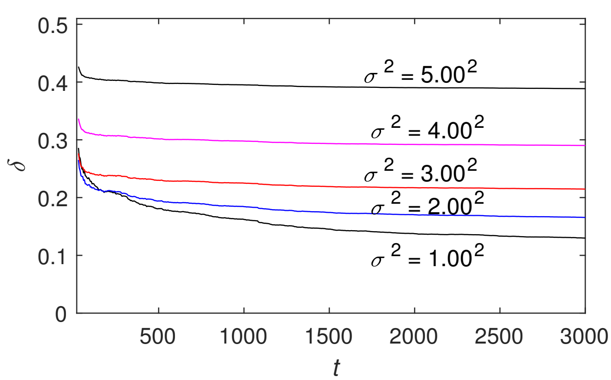

Fig. 2. The CMPN-ESG estimation errors δ versus t with different σ^2 for Example 1.

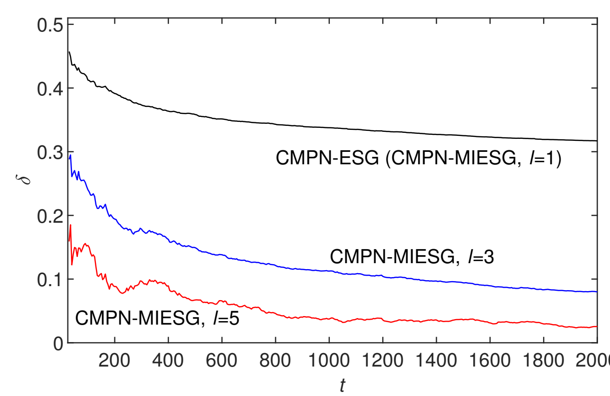

Fig. 3. The CMPN-MIESG estimation errors δ versus t under different l with $\sigma^2 = 1.00^2$ for Example 2.

3) The CMPN-ESG algorithm possesses higher parameter estimation accuracy at the same noise variances compared with the ESG algorithm.

Example 2: Consider the following FIR-MA system

$$\begin{aligned}
y(t) &= B(z)u(t) + D(z)v(t),\\
B(z) &= b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}\\
&= 1.25z^{-1} + 2.02z^{-2} + 0.45z^{-2},\\
D(z) &= 1 + d_1 z^{-1} + d_2 z^{-2} + d_3 z^{-3}\\
&= 1 + 0.65z^{-1} + 0.55z^{-2} + 0.35z^{-2}.
\end{aligned}$$

The parameter vector to be estimated is

$$\begin{aligned}
\vartheta &:= [b_1,\ b_2,\ b_3,\ d_1,\ d_2,\ d_3]^{\mathrm{T}}\\
&= [1.25,\ 2.02,\ 0.45,\ 0.65,\ 0.55,\ 0.35]^{\mathrm{T}}.
\end{aligned}$$

The simulation circumstance is similar in Example 1. The noise variance is set as $\sigma^2 = 1.00^2$. Take the data length $L = 2000$. Under the different innovation length l, the parameter estimates and errors of the CMPN-ESG algorithm (That is, the innovation length $l = 1$) and the CMPN-MIESG algorithm (the innovation length $l = 3, 5$) are depicted in Tables 3-5. Estimation errors δ versus t is shown in Fig. 3. The CMPN-MIESG estimates $\hat{\vartheta}$ versus t are depicted in Figs. 4-5. It can be seen from Fig. 3 that three estimation errors decrease as t increases, and a larger innovation length l results in the higher estimation accuracy. Furthermore, Figs. 4-5 show the CMPN-MIESG ($l = 5$) estimates can rapidly reach the vicinity of the true values with t increasing. This indicates that the CMPN-MIESG algorithm is effective for the FIR-MA system.

5. CONCLUSIONS

By exploiting continuous mixed p-norm and the multi-innovation theory, this paper derives the CMPN-ESG al-

Table 3. The CMPN-ESG estimates and their errors for Example 2.

t	b_1	b_2	b_3	d_1	d_2	d_3	δ (%)
100	0.81642	1.29989	-0.13783	0.66099	0.29422	0.20475	41.24811
200	0.86585	1.33491	-0.11971	0.66905	0.30523	0.19691	39.13768
500	0.94173	1.40151	-0.09661	0.69732	0.31971	0.18258	35.83848
1000	0.97603	1.44707	-0.07979	0.70490	0.33721	0.17467	33.76446
2000	1.00546	1.49523	-0.06275	0.70227	0.35440	0.16748	31.69614
True values	1.25000	2.02000	0.45000	0.65000	0.55000	0.35000	

Table 4. The CMPN-MIESG estimates and their errors with $l = 3$ for Example 2.

t	b_1	b_2	b_3	d_1	d_2	d_3	δ (%)
100	1.13034	1.72256	-0.06464	0.71904	0.61138	0.24898	24.02723
200	1.19841	1.75021	0.03219	0.66675	0.58068	0.31584	19.41954
500	1.26605	1.85418	0.10527	0.66612	0.52278	0.30947	14.93260
1000	1.25982	1.91851	0.18199	0.66235	0.51867	0.31237	11.25450
2000	1.25153	1.97019	0.25762	0.64864	0.52367	0.29812	8.00407
True values	1.25000	2.02000	0.45000	0.65000	0.55000	0.35000	

Table 5. The CMPN-MIESG estimates and their errors with $l = 5$ for Example 2.

t	b_1	b_2	b_3	d_1	d_2	d_3	δ (%)
100	1.33359	1.97392	0.07967	0.71540	0.59606	0.28880	15.28654
200	1.34179	1.93718	0.25879	0.64022	0.55688	0.35978	8.82167
500	1.33720	2.00664	0.30311	0.64694	0.52588	0.33473	6.71532
1000	1.28510	2.03758	0.36488	0.65395	0.53523	0.33410	3.72217
2000	1.25880	2.05831	0.42196	0.63835	0.55574	0.30624	2.56816
True values	1.25000	2.02000	0.45000	0.65000	0.55000	0.35000	

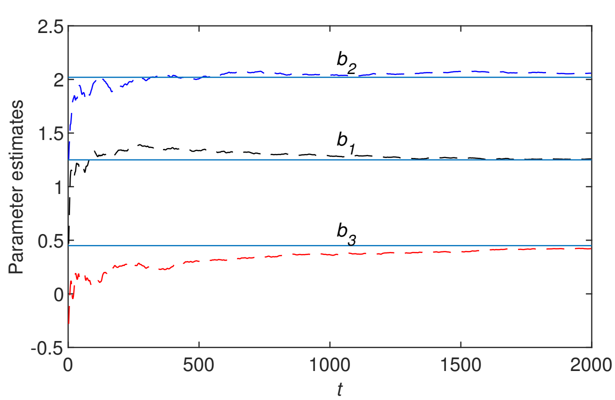

Fig. 4. The CMPN-MIESG ($l = 5$) estimates $\hat{b}_1(t)$, $\hat{b}_2(t)$ and $\hat{b}_3(t)$ versus t for Example 2.

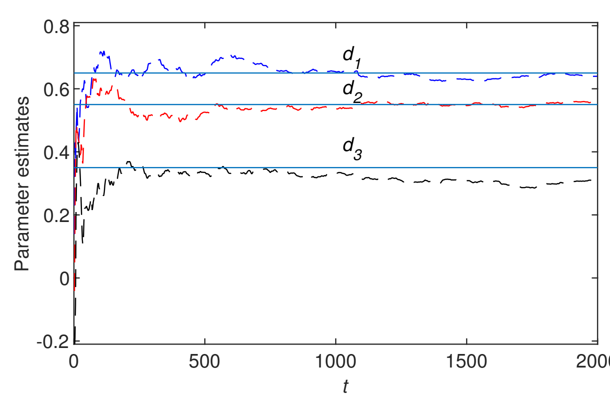

Fig. 5. The CMPN-MIESG ($l = 5$) estimates $\hat{d}_1(t)$, $\hat{d}_2(t)$ and $\hat{d}_3(t)$ versus t Example 2.

gorithm and the CMPN-MIESG algorithm for the FIR-MA systems. The continuous mixed p-norm theory is introduced to generate an adjustable gain in the algorithms and yield robustness to noise interference. The multi-innovation identification theory is introduced to improve the estimation accuracy. It is clear from the simulation results that good parameter estimates can be acquired by applying the proposed CMPN-ESG approach and CMPN-MIESG approach.

Although the CMPN-ESG algorithm and the CMPN-MIESG algorithm are effective for identifying the finite impulse response system with colored noise, they have some limitations. For example, the estimation results for a few of parameters are not quite good; if the systems have hidden variables, the algorithm may be invalid. Thus developing some algorithms with wide applicability to remedy these problems is a more challenging and interesting topic in the future. The basic idea of this paper can be combined the hierarchical identification and filtering theory to treat other linear and nonlinear identification problems.

REFERENCES

[1] M. H. Li and X. M. Liu, "Maximum likelihood least squares based iterative estimation for a class of bilinear systems using the data filtering technique," *International Journal of Control, Automation, and Systems*, vol. 18, no. 6, pp. 1581-1592, December 2020.

[2] L. Xu, "Separable Newton recursive estimation method through system responses based on dynamically discrete measurements with increasing data length," *International Journal of Control, Automation, and Systems*, vol. 20, no. 2, pp. 432-443, February 2022.

[3] L. Xu, "Separable multi-innovation Newton iterative modeling algorithm for multi-frequency signals based on the sliding measurement window," *Circuits Systems and Signal Processing*, vol. 41, no. 2, pp. 805-830, February 2022.

[4] F. Ding and T. Chen, "Combined parameter and output estimation of dual-rate systems using an auxiliary model," *Automatica*, vol. 40, no. 10, pp. 1739-1748, 2004.

[5] F. Ding and T. Chen, "Parameter estimation of dual-rate stochastic systems by using an output error method," *IEEE Transactions on Automatic Control*, vol. 50, no. 9, pp. 1436-1441, September 2005.

[6] J. Pan, X. Jiang, and W. Ding, "A filtering based multi-innovation extended stochastic gradient algorithm for multivariable control systems," *International Journal of Control, Automation, and Systems*, vol. 15, no. 3, pp. 1189-1197, June 2017.

[7] M. Li and X. Liu, "Maximum likelihood hierarchical least squares-based iterative identification for dual-rate stochastic systems," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 2, pp. 240-261, February 2021.

[8] M. Li and X. Liu, "Iterative identification methods for a class of bilinear systems by using the particle filtering technique," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 10, pp. 2056-2074, October 2021.

[9] P. Ma and L. Wang, "Filtering-based recursive least squares estimation approaches for multivariate equation-error systems by using the multiinnovation theory," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 9, pp. 1898-1915, September 2021.

[10] J. Chen, B. Huang, M. Gan, and C.L.P. Chen, "A novel reduced-order algorithm for rational models based on Arnoldi process and Krylov subspace," *Automatica*, vol. 129, Article Number: 109663, July 2021.

[11] J. Chen, "Several gradient parameter estimation algorithms for dual-rate sampled systems," *Journal of the Franklin Institute*, vol. 351, no. 1, pp. 543-554, January 2014.

[12] F. Ding, G. Liu, and X. P. Liu, "Partially coupled stochastic gradient identification methods for non-uniformly sampled systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1976-1981, August 2010.

[13] F. Ding, Y. Shi, and T. Chen, "Performance analysis of estimation algorithms of non-stationary ARMA processes," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1041-1053, 2006.

[14] L. Xu and G. Song, "A recursive parameter estimation algorithm for modeling signals with multi-frequencies," *Circuits Systems and Signal Processing*, vol. 39, no. 8, pp. 4198-4224, August 2020.

[15] F. Ding, X. P. Liu, and H. Z. Yang, "Parameter identification and intersample output estimation for dual-rate systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, vol. 38, no. 4, pp. 966-975, 2008.

[16] L. Shi, H. Zhao, and Y. Zakharov, "Generalized variable step size continuous mixed p-norm adaptive filtering algorithm," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 66, no. 6, pp. 1078-1082, June 2019.

[17] N. K. R. Rosalin and D. P. Das, "Filter proportionate normalized least mean square algorithm for a sparse system," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 11, pp. 1695-1705, October 2019.

[18] X. Zhang, "Optimal adaptive filtering algorithm by using the fractional-order derivative," *IEEE Signal Processing Letters*, vol. 29, pp. 399-403, 2022.

[19] A. Navia-Vazquez and J. Arenas-Garcia, "Combination of recursive least p-norm algorithms for robust adaptive filtering in alpha-stable noise," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1478-1482, March 2012.

[20] Y. R. Zheng and H. Nascimento, "Two variable step-size adaptive algorithms for non-Gaussian interference environment using fractionally lower-order moment minimization," *Digital Signal Processing*, vol. 23, no. 3, pp. 831-844, May 2013.

[21] H. Zayyani, "Continuous mixed p-norm adaptive algorithm for system identification," *IEEE Signal Processing Letters*, vol. 21, no. 9, pp. 1108-1110, September 2014.

[22] F. Ding, X. P. Liu, and G. Liu, "Multi-innovation least squares identification for linear and pseudo-linear regression models," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 40, no. 3, pp. 767-778, June 2010.

[23] L. Xu and E. F. Yang, "Auxiliary model multiinnovation stochastic gradient parameter estimation methods for nonlinear sandwich systems," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 1, pp. 148-165, January 2021.

[24] L. Xu, F. Ding, and Q. M. Zhu, Separable synchronous multi-innovation gradient-based iterative signal modeling from on-line measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, p. 6501313, 2022.

[25] F. Ding, L. Wan, Y. Guo, and F. Y. Chen, "The filtering based auxiliary model generalized extended stochastic gradient identification for a multivariate output-error system with autoregressive moving average noise using the multi-innovation theory," *Journal of the Franklin Institute*, vol. 357, no. 9, pp. 5591-5609, June 2020.

[26] J. Wang, Y. Ji, and C. Zhang, "Iterative parameter and order identification for fractional-order nonlinear finite impulse response systems using the key term separation," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 8, pp. 1562-1577, August 2021.

[27] Y. J. Wang, "Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model," *Automatica*, vol. 71, pp. 308-313, September 2016.

[28] Z. Kang, Y. Ji, and X. M. Liu, "Hierarchical recursive least squares algorithms for Hammerstein nonlinear autoregressive output-error systems," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 11, pp. 2276-2295, November 2021.

[29] Y. J. Liu and Y. Shi, "An efficient hierarchical identification method for general dual-rate sampled-data systems," *Automatica*, vol. 50, no. 3, pp. 962-970, March 2014.

[30] Y. Ji, Z. Kang, X. Zhang, and L. Xu, "Model recovery for multi-input signal-output nonlinear systems based on the compressed sensing recovery theory," *Journal of the Franklin Institute*, vol. 359, no. 5, pp. 2317-2339, March 2022.

[31] Y. Ji and Z. Kang, "Three-stage forgetting factor stochastic gradient parameter estimation methods for a class of nonlinear systems," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 3, pp. 971-987, February 2021.

[32] Y. Ji, Z. Kang, and X.M. Liu, "The data filtering based multiple-stage Levenberg-Marquardt algorithm for Hammerstein nonlinear systems," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 15, pp. 7007-7025, October 2021.

[33] Y. Zhou and X. Zhang, "Hierarchical estimation approach for RBF-AR models with regression weights based on the increasing data length," *IEEE Transactions on Circuits and Systems-II: Express Briefs*, vol. 68, no, 12, pp. 3597-3601, December 2021.

[34] Y. Ji, Z. Kang, and C. Zhang, "Two-stage gradient-based recursive estimation for nonlinear models by using the data filtering," *International Journal of Control, Automation, and Systems*, vol. 19, no. 8, pp. 2706-2715, August 2021.

[35] L. Xu, F.Y. Chen, and T. Hayat, "Hierarchical recursive signal modeling for multi-frequency signals based on discrete measured data," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 5, pp. 676-693, May 2021.

[36] Y. Wang, S. Tang, and X. Gu, "Parameter estimation for nonlinear Volterra systems by using the multi-innovation identification theory and tensor decomposition," *Journal of the Franklin Institute*, vol. 359, no. 2, pp. 1782-1802, January 2022.

[37] Y. Wang and L. Yang, "An efficient recursive identification algorithm for multilinear systems based on tensor decomposition," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 16, pp. 7920-7936, November 2021.

[38] J. Ding and G. Liu, "Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data," *IEEE Transactions on Automatic Control*, vol. 56, no. 11, pp. 2677-2683, November 2011.

[39] Y. Gu, Q.M. Zhu, and H. Nouri, "Identification and U-control of a state-space system with time-delay," *International Journal of Adaptive Control and Signal Processing*, vol. 36, no. 1, pp. 138-154, January 2022.

[40] Y. H. Zhou and X. Zhang, "Partially-coupled nonlinear parameter optimization algorithm for a class of multivariate hybrid models," *Applied Mathematics and Computation*, vol. 414, p. 126663, February 2022.

[41] J. L. Ding and W. H. Zhang, "Finite-time adaptive control for nonlinear systems with uncertain parameters based on the command filters," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 9, pp. 1754-1767, September 2021.

[42] J. Pan, W. Li, and H. P. Zhang, "Control algorithms of magnetic suspension systems based on the improved double exponential reaching law of sliding mode control," *International Journal of Control, Automation, and Systems*, vol. 16, no. 6, pp. 2878-2887, December 2018.

[43] J. Xiong, J. Pan, and G. Y. Chen, "Sliding mode dual-channel disturbance rejection attitude control for a quadrotor," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 10, pp. 10489-10499, 2022.

[44] J. Ding, Z. Cao, J. Chen, and G. Jiang, "Weighted parameter estimation for Hammerstein nonlinear ARX systems," *Circuits Systems and Signal Processing*, vol. 39, no. 4, pp. 2178-2192, April 2020.

[45] Y. H. Zhou, "Modeling nonlinear processes using the radial basis function-based state-dependent autoregressive models," *IEEE Signal Processing Letters*, vol. 27, pp. 1600-1604, 2020.

[46] Y. Fan and X. Liu, "Two-stage auxiliary model gradient-based iterative algorithm for the input nonlinear controlled autoregressive system with variable-gain nonlinearity," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 14, pp. 5492-5509, September 2020.

[47] X. Liu and Y. Fan, "Maximum likelihood extended gradient-based estimation algorithms for the input nonlinear controlled autoregressive moving average system with variable-gain nonlinearity," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 9, pp. 4017-4036, June 2021.

[48] X. Zhang and E. Yang, "Highly computationally efficient state filter based on the delta operator," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 6, pp. 875-889, June 2019.

[49] X. Zhang and E. Yang, "State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 7, pp. 1157-1173, July 2019.

[50] H. Ma, J. Pan, and W. Ding, "Partially-coupled least squares based iterative parameter estimation for multivariable output-error-like autoregressive moving average systems," *IET Control Theory and Applications*, vol. 13, no. 18, pp. 3040-3051, December 2019.

[51] J. Pan, H. Ma, and J. Sheng, "Recursive coupled projection algorithms for multivariable output-error-like systems with coloured noises," *IET Signal Processing*, vol. 14, no. 7, pp. 455-466, September 2020.

[52] L. Xu and E. Yang, "Separable recursive gradient algorithm for dynamical systems based on the impulse response signals," *International Journal of Control, Automation, and Systems*, vol. 18, no. 12, pp. 3167-3177, December 2020.

[53] L. Xu and Q. Zhu, "Decomposition strategy-based hierarchical least mean square algorithm for control systems from the impulse responses," *International Journal of Systems Science*, vol. 52, no. 9, pp. 1806-1821, 2021.

[54] Y. Ji, C. Zhang, Z. Kang, and T. Yu, "Parameter estimation for block-oriented nonlinear systems using the key term separation," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 9, pp. 3727-3752, June 2020.

[55] X. Zhang, "Adaptive parameter estimation for a general dynamical system with unknown states," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 4, pp. 1351-1372, March 2020.

[56] X. Zhang, "Recursive parameter estimation methods and convergence analysis for a special class of nonlinear systems," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 4, pp. 1373-1393, March 2020.

[57] Y. Ji, X. Jiang, and L. Wan, "Hierarchical least squares parameter estimation algorithm for two-input Hammerstein finite impulse response systems," *Journal of the Franklin Institute*, vol. 357, no. 8, pp. 5019-5032, May 2020.

[58] X. Zhang, "Hierarchical parameter and state estimation for bilinear systems,' *International Journal of Systems Science*, vol. 51, no. 2, 275-290, 2020.

Wentao Liu received his B.Sc. and M.Sc. degrees in automatic control from Qingdao University of Science and Technology, Qingdao, China, in 2016 and 2020, respectively. He is currently working toward a Ph.D. degree in control science and engineering from Jiangnan University, Wuxi, China. His research interests include nonlinear control, intelligent soft sensing technology, and system identification.

Weili Xiong received her Ph.D. degree in control science and engineering from Jiangnan University, Wuxi, China in 2007. She is currently a Full Professor at Jiangnan University, Wuxi, China. Her current interests include modeling and optimization of complex industrial processes, intelligent soft sensing technology and application and fault detection.