



ARTICLE

Pattern-Moving-Based Parameter Identification of Output Error Models with Multi-Threshold Quantized Observations

Xiangquan Li^{1,2}, Zhengguang Xu^{1,*}, Cheng Han¹ and Ning Li¹

¹School of Automation and Electrical Engineering, University of Science and Technology Beijing; Key Laboratory of Knowledge Automation for Industrial Processes, Ministry of Education, Beijing, 100083, China

²School of Information Engineering, Jingdezhen University, Jingdezhen, 333000, China

*Corresponding Author: Zhengguang Xu. Email: xzgustb@163.com

Received: 08 June 2021 Accepted: 03 September 2021

ABSTRACT

This paper addresses a modified auxiliary model stochastic gradient recursive parameter identification algorithm (M-AM-SGRPIA) for a class of single input single output (SISO) linear output error models with multi-threshold quantized observations. It proves the convergence of the designed algorithm. A pattern-moving-based system dynamics description method with hybrid metrics is proposed for a kind of practical single input multiple output (SIMO) or SISO nonlinear systems, and a SISO linear output error model with multi-threshold quantized observations is adopted to approximate the unknown system. The system input design is accomplished using the measurement technology of random repeatability test, and the probabilistic characteristic of the explicit metric value is employed to estimate the implicit metric value of the pattern class variable. A modified auxiliary model stochastic gradient recursive algorithm (M-AM-SGRA) is designed to identify the model parameters, and the contraction mapping principle proves its convergence. Two numerical examples are given to demonstrate the feasibility and effectiveness of the achieved identification algorithm.

KEYWORDS

Pattern moving; multi-threshold quantized observations; output error model; auxiliary model; parameter identification

1 Introduction

In the metallurgical, petroleum-chemical and steel industries, there are technologically complicated, highly energy-consuming and polluting large-scale equipments such as electrolytic tank, sintering machine, blast furnace cement rotary kiln and so on. The production process of this kind of equipments presents the following characteristics [1]. 1) The complex system mechanism beyonds the accurate description of mathematical and physical equations; 2) Working conditions and quality parameters are in large quantities, and the system moving mode is full of distributiveness, nonlinearity and parameter perturbations; 3) Some physical and chemical processes are in conformity with statistical law of moving. A feasible method of system modeling and control is



the pattern recognition technology for these considered processes [2] and most researchers' practice is to design the corresponding model and controller according to the different pattern class of the system working condition [3,4]. A novel pattern-moving-based system dynamics description method was proposed in [5]. Its basic idea is to take the pattern class as a moving variable, and it is mapped to a computable space by class centers [6,7], interval numbers [8], and cells [9] due to its lack of arithmetic operation attribute. Furthermore, in view of various metric methods of pattern class, the linear autoregressive model with exogenous input (ARX) or interval ARX (IARX) model was established, and the parameter identification algorithm based on least square [6], minimum-variance-based controller [5], optimal controller [10], state-feedback controller [7] and predictive controller [11] were designed.

Although a series of research achievements have been made in the system modeling, parameter identification and control based on pattern moving, the differences of each pattern sample in one pattern class have never been considered. Since a pattern class is a set of pattern samples with the same or similar characteristics, the hybrid metrics, that is, the combination of implicit metric $\bar{D}(\cdot)$ and class center explicit metric $D(\cdot)$ is proposed. The way of class center explicit metric indicates the statistical attribute of pattern class, and the way of implicit metric denotes the difference of each pattern sample in one pattern class. By using the hybrid metrics, the outputs of pattern-moving-based system dynamics description are in-line with multi-threshold quantized observations. Therefore, this paper will study the pattern-moving-based parameter identification of a kind of linear system models with multi-threshold quantized observations. To our best knowledge, for the model parameter identification based on accurate measured values, a great number of research results were introduced in [12], such as least square method [13], maximum likelihood and prediction error methods [14], unbiased finite impulse response filter algorithm [15], Kalman linear filtering and prediction algorithm [16,17], Bayesian approach [18], and so forth.

The system with multi-threshold quantized observations is considered as a set-valued system which is different from a conventional one with accurate measurement outputs [19,20]. Although the identification of set-valued systems is not easy task, a series of important research results have been achieved in the past decade. For the linear system identification with threshold quantized observations, a full rank input design method (such as repetitive input design) and an empirical distribution function method were proposed in [21–23]. For the identification of Hammerstein and Wiener systems with binary outputs, the methods of proportional full rank signal, joint identifiability and strong full rank periodic signal inputs were proposed in [24,25], which effectively overcame the problems of system nonlinearity and rough binary output information. The parameter identification of Wiener system was also studied in terms of quantized inputs and binary outputs in [26]. Based on the truncated empirical measurement method of probabilistic statistics, a non-truncated empirical measurement method was proposed for the finite impulse response (FIR) system model with binary outputs, and the progressive effectiveness of the algorithm was demonstrated in [27]. A two-segment design method for a class of FIR systems with binary-valued observations was investigated in [28]. The parameter estimation problem of set-valued system was explained comprehensively from the perspective of combining system principle and practical application in [29]. The idea of parameter identification based on an auxiliary model was systematically expounded in [30]. An auxiliary-model-based least squares recursive algorithm was proposed for a quantized control system with communication constraints in [31].

This work investigates a M-AM-SGRPIA for a kind of linear output error models with multi-threshold quantized observations which is established from the perspective of pattern moving and

hybrid metrics. Compared with the existed research results, the main differences and contributions of this paper are summarized as follows:

1) Different from the previous system identification problem of ARX or IARX models based on pattern moving and single metric [6,7], this paper considers hybrid metrics, the model noise distribution and proves the convergence of the designed M-AM-SGRPIA.

2) Compared with the parameter identification algorithms of set-valued system in [21–29], this paper adopts an auxiliary model and designs a M-AM-SGRPIA, which will reduce the estimation error radio to some extern.

3) An auxiliary-model-based least squares recursive algorithm has been designed for a class of linear systems with communication constraints in [31], which is different from the designed algorithm and physical background in this paper.

The outline of this work is organized as follows. Section 2 presents a pattern-moving-based system dynamics description and problem formulation. Section 3 proposes a M-AM-SGRPIA and its convergence is proved by the contraction mapping principle in Section 4. Section 5 demonstrates the feasibility and effectiveness of the M-AM-SGRPIA by two numerical examples. The conclusion comes in Section 6.

Notation: \mathbb{R} stands for the real number field; Z_+ denotes the positive integer field; $\lambda_{max}[A]$ is the maximum eigenvalue of matrix A , and $\lambda_{min}[A]$ is its minimum eigenvalue; $f(t) = o(g(t))$ represents $g(t) > 0$, $\lim_{t \rightarrow \infty} \frac{f(t)}{g(t)} = 0$; $\arg \max_i(\cdot)$ is the value of i when (\cdot) gets the maximum value; $\|\cdot\|$ is the Euclidean norm; the superscript T denotes the transposition; $F(\cdot)$ and $F^{-1}(\cdot)$ represent the probability distribution function and its inverse function respectively; $E(\cdot)$ is the mathematical expectation; $P(\cdot)$ denotes the probability.

2 Preliminary and Problem Formulation

2.1 Pattern-Moving-Based System Dynamics Description

Consider a class of unknown SIMO non-affine nonlinear discrete-time systems as follows:

$$\left\{ \begin{array}{l} y_1(k) = f_1(y_1(k-1), \dots, y_1(k-n_1), u(k-1), \dots, u(k-m_1)), \\ y_2(k) = f_2(y_2(k-1), \dots, y_2(k-n_2), u(k-1), \dots, u(k-m_2)), \\ \vdots \\ y_q(k) = f_q(y_n(k-1), \dots, y_n(k-n_q), u(k-1), \dots, u(k-m_q)), \end{array} \right. \quad (1)$$

where $q > 0$; $y_i(k) \in \mathbb{R}$ and $u(k) \in \mathbb{R}$ denote the output of $f_i(\cdot)$ and the whole system's input respectively; $n_i \in Z_+$ and $m_i \in Z_+$ are the unknown output and input orders; $f_i(\cdot)$ denotes an unknown nonlinear discrete-time function; $i \in \{1, \dots, q\}$.

Assumption 2.1 The input of system (1) is bounded, i.e., a constant M_1 exists and satisfies that $|u(k)| \leq M_1$.

A pattern-moving-based system dynamics description [32] corresponding to system (1) is proposed in the following three steps:

1) Feature extraction ($T(\cdot)$) and output dimension reduction. A large number of input and output data are collected offline, and the input data set $\{u(k)\}$ and q -dimensional output vector set

$\{[y_1(k), \dots, y_q(k)]\}$ are obtained. Drawing on the data-driven modeling ideas [33,34] and through the principal component analysis (PCA) feature extraction [35] of the output data, the first principal component information is obtained, and then one-dimensional principal component information set $\{y(k)\}$ will be obtained.

2) Classification ($M(\cdot)$) and metrics ($D(\cdot), \bar{D}(\cdot)$). Using pattern classification technology to classify the first principal component information, the number of pattern class (N), the class center value (s_i), class radius (r_i) and threshold value (C_i) of each pattern class (dx_i) can be obtained, $i = [1, \dots, N]$. At any time instant k , the system output vector $[y_1(k), \dots, y_q(k)]$ corresponds to a working condition and a pattern class $dx(k)$ at the same time. Since the pattern class does not have the arithmetic operation attribute, that is, pattern class 1 + pattern class 2 \neq pattern class 3, the pattern class variable needs to be measured. Since the pattern class is a set of pattern samples with the same or similar attributes, the method of combining the class center explicit metric $D(\cdot)$ and implicit metric $\bar{D}(\cdot)$ is adopted, that is, $s_i = D(dx_i)$ and $\bar{dx}_i = \bar{D}(dx_i)$. The implicit metric values are unknown, but there is a definite relationship between an implicit metric value and a class center explicit metric value by choosing a reasonable classification method, such as $|s_i - \bar{dx}_i| \leq r_i$.

3) Establishing the pattern-moving-based system dynamics equations. The input sequence $\{u(k)\}$, implicit metric sequence $\{\bar{dx}(k)\}$ and class center explicit metric sequence $\{s(k)\}$ are employed to construct the following dynamics equations of the system:

$$\bar{dx}(k) = f(\bar{dx}(k-1), \dots, \bar{dx}(k-n), u(k-1), \dots, u(k-m)); \quad (2)$$

$$s(k) = D(M(\bar{dx}(k))) = \begin{cases} s_1, \bar{dx}(k) \in [s_1 - r_1, s_1 + r_1]; \\ s_2, \bar{dx}(k) \in [s_2 - r_2, s_2 + r_2]; \\ \vdots \\ s_N, \bar{dx}(k) \in [s_N - r_N, s_N + r_N], \end{cases} \quad (3)$$

where $f(\cdot)$ is a unknown system function; m, n denote the input and output orders of system (2) respectively. It exits a classification-metric deviation $e(k)$ between the $\bar{dx}(k)$ and $s(k)$, and $|e(k)| = |s(k) - \bar{dx}(k)| \leq r_1$ while $s(k) = c_i$.

If the contribution rate of the first principal component information obtained by feature extraction $T(\cdot)$ is greater than 85%, it can be considered that the output information of source system (1) is not lost or loses very little. For many practical SIMO physical systems, the contribution rate of the first principal component information can reach or even exceed 85%. Taking sintering machine as an example, through data analysis, it is found that the contribution rate of first principal component information is far greater than 85% when the temperatures of four positions near the sintering node under the excitation of ignition input are considered and feature extracted [5].

Remark 2.1 In the process of establishing the pattern-moving-based dynamics Eqs. (2) and (3), the condition of parameter configuration of classification method is to ensure that a certain pattern class corresponds to a specific quality index of the product [36]. Furthermore, a physical SISO nonlinear discrete-time system can also be transformed into a pattern-moving-based SISO system, but it does not need the first step of feature extraction process.

2.2 Problem Formulation

Although there inevitably unmodeled dynamics problems, it is common to employ a linear model to approximate the situation that the system (2) is unknown. Choosing a reasonable classification method such as a modified quantized control classification [36], the following linear output error model with multi-threshold quantized observations can be constructed for system (2)–(3).

$$\left\{ \begin{array}{l} \overline{dx}(k) = a_1\overline{dx}(k-1) + \dots + a_n\overline{dx}(k-n) + b_1u(k-d-1) + \dots + b_mu(k-d-m); \\ \hat{\overline{dx}}(k) = \overline{dx}(k) + v(k); \\ s(k) = D(M(\hat{\overline{dx}}(k))) = \begin{cases} c_1, c_1 - r_1 < \hat{\overline{dx}}(k) \leq c_1 + r_1; \\ c_2, c_2 - r_2 < \hat{\overline{dx}}(k) \leq c_2 + r_2; \\ \vdots \\ c_N, c_N - r_N < \hat{\overline{dx}}(k) \leq c_N + r_N, \end{cases} \end{array} \right. \quad (4)$$

where $v(k)$ is the model noise; $s(k) \in \{c_1, c_2, \dots, c_N\}$ is the class center explicit metric value; $D(M(\cdot))$ denotes a classification-metric process; $C_i = c_i + r_i = c_{i+1} - r_{i+1}$ is satisfied under a reasonable classification.

Assumption 2.2 The input and output orders of the model are known and equal, that is, $n = m$.

Assumption 2.3 $v(k)$ is an independent and identically distributed (*i.i.d.*) random signal noise, and its probability distribution function $F(s)$ and inverse function $F^{-1}(s)$ are known and continuously differentiable.

Under the Assumption 2.2, the first expression of model (4) can be written as follows:

$$\overline{dx}(k) = \frac{B(q)}{A(q)}u(k) = \varphi^T(k)\theta, \quad (5)$$

where $A(q) = 1 - a_1q - \dots - a_nq^n$; $B(q) = b_1q + b_2q^2 + \dots + b_nq^n$; $\varphi(k) = [\overline{dx}(k-1), \dots, \overline{dx}(k-n), u(k-1), \dots, u(k-n)]^T$; the parameter vector $\theta = [a_1, \dots, a_n, b_1, \dots, b_n]^T$ is to be identified.

If the input sequence $\{u(k)\}$ and output sequence $\{\overline{dx}(k)\}$ can be obtained directly, it is known to all that there exists a lot of methods to identify the model parameters. However, one can only get the observation sequence $\{s(k)\}$ and $\{u(k)\}$. Therefore, this paper will design a parameter identification algorithm based on these two sequences.

Remark 2.2 The model orders and cumulative distribution function of $v(k)$ are supposed to be known in Assumption 2.2 and Assumption 2.3, respectively. These two problems are not the focus of this paper. If the model orders are unknown, in practice, the order selection may mainly depend on the experiential knowledge of the system. From a theoretical point of view, the model order selection could be addressed by means of the parameter estimation technique with quantized observations and drawing on the classical order selection methods of cross-validation, information criteria, the F -test and the statistical tests on the residuals. In view of the unknown distribution function of model noise, it is not easy task to identify that. However, one can refer to the parameterization technique proposed in [37].

3 Design of Parameter Identification Algorithm

3.1 Estimation of Implicit Metric Value $\overline{dx}(k)$

For the convenience of calculation, $s(k)$ is denoted by $ss(k) = [ss_k^1, ss_k^2, \dots, ss_k^N]^T$, and the corresponding indicator function of each element ss_k^i is depicted as follows:

$$ss_k^i = \begin{cases} 1, & s(k) \leq C_i; \\ 0, & \text{others,} \end{cases} \quad (6)$$

where N is the number of the pattern class; C_i denotes the threshold value and it satisfies that $C_i = c_i + r_i$.

The measurement technology of random repeatability test is employed to design the input $u(k)$. The size of input cycle is set to be $t(t > 2n)$ and the number of cycles is l . Therefore, the noise free output $\overline{dx}(k)$ is an unmeasurable signal with cycle equal to t . Totally $t \cdot l$ rows of vector data $\{ss^T(k)\}$ and a data matrix $Q_{(t \cdot l) \cdot N}$ is produced. It is known from Assumption 2.3 that the probability distribution of $v(k)$ is $F(s)$, and it can be derived from the random repeatability test technology that the occurrence probability of event $\widehat{dx}(k) \leq C_i$ at time instant k is p_i , e.g., the occurrence probability of event $ss_k^i = 1$ is $p_i = P\{\widehat{dx}(k) \leq C_i\} = P\{\overline{dx}(k) + v(k) \leq C_i\} = F(C_i - \overline{dx}(k))$, and then $\overline{dx}(k) = C_i - F^{-1}(p_i)$. Due to the influence of $v(k)$, the corresponding output $s(k)$ obtained from each test may differ under the identical test conditions. Let $\xi_l^i = \frac{1}{l} \sum_{j=0}^{l-1} ss_{k+t \cdot j}^i$. It is known from the probability statistics that $\{ss^i(k+t \cdot l)\}$ is the result of Bernoulli tests and it satisfies that $E(\xi_l^i) = p_i$. Therefore, it follows that ξ_l^i is the estimated value of p_i . Further, the estimated value of $\overline{dx}(k)$ can be gained that $\widetilde{\overline{dx}}(k) = C_i - F^{-1}(\xi_l^i)$ and the estimation error $e(k) = \widetilde{\overline{dx}}(k) - \overline{dx}(k)$. In terms of the estimation error $e(k)$, there exists two Lemmas and one conclusion as follows.

Lemma 3.1 ([31]) For the model (4) and the output sequence $\{ss^i(k+t \cdot l)\}$ generated by repeatability test, the estimation error sequence $\{e(k): e(k) = \widetilde{\overline{dx}}(k) - \overline{dx}(k)\}$ is a time-varying sequence with mean 0 and bounded variance, and the estimation error $e(k)$ converges to 0 according to probability 1.

Lemma 3.2 ([31]) Under the full order input sequence $\{u(k+t \cdot j)\}, 1 \leq k \leq t, 0 \leq j \leq l-1$ with cycle equal to $t(t > 2n)$, the signals to be identified generated by model (4) have the property of persistent excitation, e.g.,

$$\frac{1}{k} \sum_{i=1}^k \varphi(k_0 - i + 1) \varphi^T(k_0 - i + 1) \geq \alpha I, \quad (7)$$

where $\alpha > 0$; k_0 denotes the starting time; I is an identity matrix.

According to Lemma 3.1 and Lemma 3.2, the estimation error sequence $\{e(k), \Gamma_k\}$ is regarded as a martingale difference sequence defined in a probability space $\{\Omega, \Gamma, P\}$, and it satisfies the following conditions:

$$A_1) E[e(k)|\Gamma_{k-1}] = 0, a.s.;$$

$$A_2) E[e^2(k)|\Gamma_{k-1}] = \sigma_e^2(k) \leq M\sigma_\zeta^2 < \infty,$$

where $\{\Gamma_k\}$ is an σ algebra sequence generated by $\{e(k)\}$ and $\bar{\sigma}_e^2 = M\sigma_\zeta^2$.

3.2 Design of M-AM-SGRA

To accomplish the identification task, it is necessary to determine the probability estimate ξ_l^i and find the corresponding threshold C_i firstly. Considering the presence of $\nu(k)$, it is required to determine the corresponding threshold segmentation point C_i of $s(k)$ at the occurrence time of event $ss_k^i = 1$. For the matrix $Q_{(t,l) \cdot N}$, the corresponding row k is the demonstrative vector $ss^T(k)$, and it is known from its definition that if $ss^T(k)$, then $ss_k^j = 1, j \geq i$. The ξ_l^i and threshold value C_i are determined by the following formula:

$$\xi_l^i = \begin{cases} \arg \max_i \left\{ \xi_l^i : \xi_l^i = \frac{1}{l} \sum_{j=0}^{l-1} ss_{k+tj}^i \right\}, & 0 < \xi_l^i < 1, i = 1, 2, \dots, N; \\ \text{Num}; \xi_l^i = 0 \text{ or } \xi_l^i = 1, i, j \in [1, N], \end{cases} \quad (8)$$

1) If $\xi_l^i \in (0, 1)$, the maximum value of ξ_l^i and its corresponding interval number i can be found. Further, the corresponding threshold C_i is determined.

2) If ξ_l^i is equal to 0 or 1, then corresponding interval number i can be determined when the first $\xi_l^i = 1$. Further, the corresponding threshold C_i is determined. Furthermore, e.g., let $\xi_l^i = \text{Num} = 0.99$, which is a truncation data.

The $\{e(k), \Gamma_k\}$ is considered as a non-stationary martingale difference sequence with mean 0 in Section 3.1. Therefore, one can establish a new identification model by using the estimated value $\tilde{dx}(k)$ such as:

$$\tilde{dx}(k) = \overline{dx}(k) + e(k) = \varphi^T(k)\theta + e(k). \quad (9)$$

Due to the unknown variables $\overline{dx}(k-1), \dots, \overline{dx}(k-n)$ exist in (9), an auxiliary model is established as follows:

$$\overline{dx}_a(k) = \tilde{\varphi}^T(k)\tilde{\theta}(k), \quad (10)$$

where $\tilde{\theta}(k)$ is an estimate of θ , and $\tilde{\varphi}(k) = [\overline{dx}_a(k-1), \dots, \overline{dx}_a(k-n), u(k-1), \dots, u(k-n)]^T$ is an estimate of $\varphi(k)$ at time instant k .

By introducing a convergence index $\varepsilon(\frac{1}{2} < \varepsilon \leq 1)$ and using the auxiliary model (10), a modified auxiliary model stochastic gradient recursive algorithm (M-AM-SGRA) for the parameter vector θ is designed as follows:

$$\tilde{\theta}(k) = \tilde{\theta}(k-1) + \frac{\tilde{\varphi}(k)}{r^\varepsilon(k)}\pi(k), \tilde{\theta}(0) = \frac{\mathbf{1}_{2n}}{p_0}, p_0 > 0; \quad (11)$$

$$\pi(k) = \tilde{dx}(k) - \tilde{\varphi}^T(k)\tilde{\theta}(k-1); \quad (12)$$

$$r(k) = r(k-1) + \|\tilde{\varphi}(k)\|^2, r(0) = 1; \quad (13)$$

$$\overline{dx}_a(k) = \tilde{\varphi}^T(k)\tilde{\theta}(k), \overline{dx}_a(-i) = \frac{1}{p_0}, i = 0, 1, \dots, n; \quad (14)$$

$$\tilde{d}x(k) = C_i - F^{-1}(\xi_l^i), \xi_l^i = \underset{i}{\operatorname{argmax}} \left(\xi_l^i : \xi_l^i = \frac{1}{l} \sum_{j=0}^{l-1} ss_{k+t-j}^i \right) \text{ or Num.} \tag{15}$$

Remark 3.1 A truncation method is adopted to get the value ξ_l^i . The determination of truncation rules depends on prior information. The truncated empirical measure method was proposed in [25], and the asymptotic efficiency of the algorithm was proved. This paper chooses a reasonable truncation value of ξ_l^i by referring to [25].

4 Main Results

For the designed M-AM-SGRA, the following Lemma is given and its convergence is to be proved.

Lemma 4.1 ([31]) For the given M-AM-SGRA (11)–(15), the following inequalities hold.

$$\sum_{i=1}^k \frac{\|\tilde{\varphi}(i)\|^2}{r(i)} \leq \ln r(k), a.s.; \tag{16}$$

$$\sum_{i=1}^{\infty} \frac{\|\tilde{\varphi}(i)\|^2}{r^\eta(i)} < \infty, a.s. \eta > 1 \tag{17}$$

where *a.s.* means almost surely.

Theorem 4.1 For the model (9) and the corresponding M-AM-SGRA (11)–(15), under the conditions that A_1) and A_2) are satisfied, if $A(q)$ is strictly positive real, $r(k) \rightarrow \infty$, letting

$$R(k) = \sum_{i=1}^{\infty} \tilde{\varphi}(i)\tilde{\varphi}^T(i) \in R^{2n \cdot 2n} \tag{18}$$

and it is assumed that the following inequality holds

$$r^\varepsilon(k) \geq \frac{1}{2} \|\tilde{\varphi}(k)\|^2, \tag{19}$$

thus, the estimation error vector $\check{\theta}(k)$ ($\check{\theta}(k) = (\tilde{\theta}(k) - \theta)$) satisfies that

$$\|\tilde{\theta}(k) - \theta\|^2 = o \left\{ \frac{r^{2-\varepsilon}(k)}{\lambda_{\min}[R(k)]} \right\}, a.s.. \tag{20}$$

Proof.

It is concluded from (11) that

$$\begin{aligned} \check{\theta}(k) &= \check{\theta}(k-1) + \frac{\tilde{\varphi}(k)}{r^\varepsilon(k)} \pi(k) = \check{\theta}(k-1) + \frac{\tilde{\varphi}(k)}{r^\varepsilon(k)} \left[\tilde{\varphi}^T(k)\theta + e(k) - \tilde{\varphi}^T(k)\check{\theta}(k-1) \right] \\ &= \check{\theta}(k-1) + \frac{\tilde{\varphi}(k)}{r^\varepsilon(k)} \left[-\tilde{\varphi}^T(k)\check{\theta}(k-1) + e(k) \right] = \check{\theta}(k-1) + \frac{\tilde{\varphi}(k)}{r^\varepsilon(k)} \left[-\check{y}(k) + e(k) \right], \end{aligned} \tag{21}$$

where $\check{y}(k) = \tilde{\varphi}^T(k)\check{\theta}(k-1)$.

Taking the norm of both sides of (21), one has

$$\begin{aligned}
 \|\check{\theta}(k)\|^2 &= \left\| \check{\theta}(k-1) + \frac{\tilde{\varphi}(k)}{r^\varepsilon(k)} [-\check{y}(k) + e(k)] \right\|^2 \\
 &= \|\check{\theta}(k-1)\|^2 + \frac{2}{r^\varepsilon(k)} \check{\theta}^T(k-1) \tilde{\varphi}(k) [-\check{y}(k) + e(k)] + \frac{\|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} [-\check{y}(k) + e(k)]^2 \\
 &= \|\check{\theta}(k-1)\|^2 - \frac{2r^\varepsilon(k) - \|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} \check{y}^2(k) + \frac{2[r^\varepsilon(k) - \|\tilde{\varphi}(k)\|^2]}{r^{2\varepsilon}(k)} \check{y}(k)e(k) + \frac{\|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} e^2(k) \\
 &\leq \|\check{\theta}(k-1)\|^2 - \frac{2r^\varepsilon(k) - \|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} \check{y}^2(k) + \frac{\|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} e^2(k) + \frac{2}{r^\varepsilon(k)} \check{y}(k)e(k). \tag{22}
 \end{aligned}$$

Since $e(k)$ is uncorrelated with $\check{\theta}(k)$, $\check{y}(k)$, $r(k)$, and $\tilde{\varphi}(k)$, and it is Γ_{k-1} measurable, according to (16)-(17) and taking the conditional expectation of (22) with respect to Γ_{k-1} , one gets

$$E \left[\|\check{\theta}(k)\|^2 \mid \Gamma_{k-1} \right] \leq \|\check{\theta}(k-1)\|^2 - \frac{2r^\varepsilon(k) - \|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} \check{y}^2(k) + \frac{\|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} M\sigma_\zeta^2 \tag{23}$$

According to Lemma 4.1, the sum of the third term on the right side of (23) from $t = 1$ to $t = \infty$ is finite. Based on the condition (19), it can be concluded from the martingale convergence Theorem (Lemma d.5.3 in [39]) that $\|\check{\theta}(k)\|^2$ uniformly converges to a finite random variable C and it is noted as

$$\lim_{k \rightarrow \infty} \|\check{\theta}(k)\|^2 = C < \infty, a.s.. \tag{24}$$

It can be also derived from the martingale convergence Theorem that

$$\sum_{k=1}^{\infty} \frac{2r^\varepsilon(k) - \|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} \check{y}^2(k) < \infty, a.s.. \tag{25}$$

Letting $\varepsilon = 1$, it is easy to conclude that (19) is tenable to get and one has

$$\begin{aligned}
 \sum_{k=1}^{\infty} \frac{2r^\varepsilon(k) - \|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} \check{y}^2(k) &= \left[\sum_{k=1}^{\infty} \frac{2r(k) - \|\tilde{\varphi}(k)\|^2}{r^2(k)} \check{y}^2(k) \right]_{\varepsilon=1} = \sum_{k=1}^{\infty} \frac{r(k) + r(k-1)}{r^2(k)} \check{y}^2(k) \\
 &\geq \sum_{k=1}^{\infty} \frac{1}{r(k)} \check{y}^2(k) \geq \left[\sum_{k=1}^{\infty} \frac{1}{r^{2-\kappa}(k)} \check{y}^2(k) \right]_{\frac{1}{2} < \kappa \leq 1} \tag{26}
 \end{aligned}$$

One can sequentially get

$$\sum_{k=1}^{\infty} \left[\frac{1}{r(k)} \check{y}^2(k) \right] \leq \sum_{k=1}^{\infty} \left[\frac{2r^\varepsilon(k) - \|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} \check{y}^2(k) \right], \quad (27)$$

$$\left[\sum_{k=1}^{\infty} \frac{1}{r^{2-k}(k)} \check{y}^2(k) \right]_{\frac{1}{2} < \kappa \leq 1} \leq \sum_{k=1}^{\infty} \left[\frac{2r^\varepsilon(k) - \|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} \check{y}^2(k) \right]. \quad (28)$$

Further, using the Kronecker Lemma (Lemma d.5.3 in [38]) for inequalities (27), (28) to yield

$$\lim_{k \rightarrow \infty} \frac{1}{r(k)} \sum_{i=1}^k \check{y}^2(i) = 0, \quad (29)$$

and

$$\lim_{k \rightarrow \infty} \frac{1}{r^{2-\varepsilon}(k)} \sum_{i=1}^k \check{y}^2(i) = 0. \quad (30)$$

It is known from the above proof process that the parameter estimation error is continuous and bounded. Using (21), one obtains

$$\check{\theta}(k) = \check{\theta}(k-i) + \sum_{j=0}^{i-1} \frac{\tilde{\varphi}(k-j)}{r^\varepsilon(k-j)} [-\check{y}(k-j) + e(k-j)], i \geq 1.$$

Replacing k with $k-i$ in $\check{y} = \tilde{\varphi}^T(k)\check{\theta}(k-1)$ yields

$$\tilde{\varphi}^T(k-i)\check{\theta}(k) = \check{y}(k-i) + \tilde{\varphi}^T(k-i) \sum_{j=0}^i \frac{\tilde{\varphi}(k-j)}{r^\varepsilon(k-j)} [-\check{y}(k-j) + e(k-j)]. \quad (31)$$

Taking the square on both sides of (31) and adopting the inequality $(a+b)^2 \leq 2(a^2+b^2)$, one gets

$$\begin{aligned} \left[\tilde{\varphi}^T(k-i)\check{\theta}(k) \right]^2 &= \left\{ \check{y}(k-i) + \tilde{\varphi}^T(k-i) \sum_{j=0}^i \frac{\tilde{\varphi}(k-j)}{r^\varepsilon(k-j)} [-\check{y}(k-j) + e(k-j)] \right\}^2 \\ &\leq 2\check{y}^2(k-i) + 2\|\tilde{\varphi}(k-i)\|^2 \left\| \sum_{j=0}^i \frac{\tilde{\varphi}(k-j)}{r^\varepsilon(k-j)} [-\check{y}(k-j) + e(k-j)] \right\|^2. \end{aligned} \quad (32)$$

Since $e(k-j)$ is uncorrelated with $\check{y}(k-j)$, $r(k-j)$, and $\tilde{\varphi}(k-j)$, and it is Γ_{k-1} measurable, according to (16), (17) and taking the conditional expectation on both sides of (32) with respect to Γ_{k-1} , one has

$$E\left\{\left[\tilde{\varphi}^T(k-i)\check{\theta}(k)\right]^2 \mid \Gamma_{k-1}\right\} \leq 2\check{y}^2(k-i) + 2\|\tilde{\varphi}(k-i)\|^2 \sum_{j=0}^i \left\{ \frac{\|\tilde{\varphi}(k-i)\|^2}{r^{2\varepsilon}(k-j)} \left[\check{y}^2(k-i) + M\sigma_\zeta^2\right] \right\}. \quad (33)$$

Taking the sum from $i=0$ to $i=k-1$ on both sides of (33) and dividing by $r^{2-\varepsilon}(k)$ yield

$$E\left\{\sum_{i=0}^{k-1} \left[\tilde{\varphi}^T(k-i)\check{\theta}(k)\right]^2 \mid \Gamma_{k-1}\right\} \frac{1}{r^{2-\varepsilon}(k)} \leq \frac{2}{r^{2-\varepsilon}(k)} \sum_{i=0}^{k-1} \check{y}^2(k-i) + 2 \sum_{i=1}^{k-1} \frac{\|\tilde{\varphi}(k-i)\|^2}{r^{2-\varepsilon}(k)} \sum_{j=0}^i \left[\frac{\|\tilde{\varphi}(k-j)\|^2}{r^{2\varepsilon}(k-j)} \check{y}^2(k-j) \right] + 2 \sum_{i=1}^{k-1} \frac{\|\tilde{\varphi}(k-i)\|^2}{r^{2-\varepsilon}(k)} \sum_{j=0}^i \left[\frac{\|\tilde{\varphi}(k-j)\|^2}{r^{2\varepsilon}(k-j)} M\sigma_\zeta^2 \right],$$

where

$$\psi_1(k) = 2 \sum_{i=1}^{k-1} \frac{\|\tilde{\varphi}(k-i)\|^2}{r^{2-\varepsilon}(k)} \sum_{j=0}^i \left[\frac{\|\tilde{\varphi}(k-j)\|^2}{r^{2\varepsilon}(k-j)} M\sigma_\zeta^2 \right],$$

and

$$\psi_2(k) = 2 \sum_{i=1}^{k-1} \frac{\|\tilde{\varphi}(k-i)\|^2}{r^{2-\varepsilon}(k)} \sum_{j=0}^i \left[\frac{\|\tilde{\varphi}(k-j)\|^2}{r^{2\varepsilon}(k-j)} \check{y}^2(k-j) \right].$$

According to Lemma 4.1, the following two inequalities can be obtained.

$$\begin{aligned} \psi_1(k) &= 2 \sum_{i=1}^{k-1} \frac{\|\tilde{\varphi}(k-i)\|^2}{r^{2-\varepsilon}(k)} \sum_{j=0}^i \left[\frac{\|\tilde{\varphi}(k-j)\|^2}{r^{2\varepsilon}(k-j)} M\sigma_\zeta^2 \right] \\ &= 2 \frac{2}{r^{2-\varepsilon}(k)} \left\{ \frac{[r(k-1) - r(0)] \|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} M\sigma_\zeta^2 + \sum_{i=1}^{k-1} \left[\frac{[r(k-1) - r(0)] \|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} M\sigma_\zeta^2 \right] \right\} \\ &\leq \frac{2}{r^{2-\varepsilon}(k)} \frac{\|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon-1}(k)} M\sigma_\zeta^2 + \frac{2}{r^{2-\varepsilon}(k)} \sum_{i=1}^{k-1} \left[\frac{\|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon-1}(i)} M\sigma_\zeta^2 \right] \\ &= \frac{2M\sigma_\zeta^2}{r^{2-\varepsilon}(k)} \sum_{i=1}^k \left[\frac{r^{2-2\varepsilon} \|\tilde{\varphi}(i)\|^2}{r(i)} \right] \leq \frac{2M\sigma_\zeta^2 \ln(r(k))}{r^\varepsilon(k)} \rightarrow 0, a.s., k \rightarrow \infty, \frac{1}{2} < \varepsilon \leq 1. \end{aligned}$$

$$\psi_2(k) = 2 \sum_{i=1}^{k-1} \frac{\|\tilde{\varphi}(k-i)\|^2}{r^{2-\varepsilon}(k)} \sum_{j=0}^i \left[\frac{\|\tilde{\varphi}(k-j)\|^2}{r^{2\varepsilon}(k-j)} \check{y}^2(k-j) \right]$$

$$\begin{aligned}
&= 2 \frac{\|\tilde{\varphi}(k-1)\|^2}{r^{2-\varepsilon}(k)} \left[\frac{\|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} \check{y}^2(k) + \frac{\|\tilde{\varphi}(k-1)\|^2}{r^{2\varepsilon}(k-1)} \check{y}^2(k-1) \right] \\
&\quad + 2 \frac{\|\tilde{\varphi}(k-2)\|^2}{r^{2-\varepsilon}(k)} \left[\frac{\|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} \check{y}^2(k) + \frac{\|\tilde{\varphi}(k-1)\|^2}{r^{2\varepsilon}(k-1)} \check{y}^2(k-1) + \frac{\|\tilde{\varphi}(k-2)\|^2}{r^{2\varepsilon}(k-2)} \check{y}^2(k-2) \right] \\
&+ \dots \\
&\quad + 2 \frac{\|\tilde{\varphi}(1)\|^2}{r^{2-\varepsilon}(k)} \left[\frac{\|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} \check{y}^2(k) + \frac{\|\tilde{\varphi}(k-1)\|^2}{r^{2\varepsilon}(k-1)} \check{y}^2(k-1) + \dots + \frac{\|\tilde{\varphi}(1)\|^2}{r^{2\varepsilon}(1)} \check{y}^2(1) \right] \\
&= \frac{2}{r^{2-\varepsilon}(k)} \frac{[r(k-1) - r(0)] \|\tilde{\varphi}(k)\|^2}{r^{2\varepsilon}(k)} \check{y}^2(k) + \frac{2}{r^{2-\varepsilon}(k)} \sum_{i=1}^{k-1} \left[\frac{[r(i) - r(0)] \|\tilde{\varphi}(i)\|^2}{r^{2\varepsilon}(i)} \check{y}^2(i) \right] \\
&\leq \frac{2}{r^{2-\varepsilon}(k)} \frac{2r^\varepsilon(k)}{r^{2\varepsilon-1}(k)} \check{y}^2(k) + \frac{2}{r^{2-\varepsilon}(k)} \sum_{i=1}^{k-1} \left[\frac{2r^\varepsilon(k)}{r^{2\varepsilon-1}(i)} \check{y}^2(i) \right] \\
&= \frac{4}{r^{2-\varepsilon}(k)} \sum_{i=1}^k \left[\frac{1}{r^{\varepsilon-1}(i)} \check{y}^2(i) \right] \leq \frac{4}{r(k)} \sum_{i=1}^k \check{y}^2(i) \rightarrow 0, a.s., k \rightarrow \infty, \frac{1}{2} < \varepsilon \leq 1.
\end{aligned}$$

Using the above two inequalities, one has

$$\begin{aligned}
E \left\{ \sum_{i=0}^{k-1} \left[\tilde{\varphi}^T(k-i) \check{\theta}(k) \right]^2 \middle| \Gamma_{k-1} \right\} \frac{1}{r^{2-\varepsilon}(k)} &= \frac{E \left\{ \check{\theta}^T(k) R(k) \check{\theta}(k) \middle| \Gamma_{k-1} \right\}}{r^{2-\varepsilon}(k)} \\
&\leq \frac{2}{r^{2-\varepsilon}(k)} \sum_{i=0}^{k-1} \check{y}^2(k-i) + \psi_1(k) + \psi_2(k) \rightarrow 0, a.s. \quad k \rightarrow \infty, \frac{1}{2} < \varepsilon \leq 1.
\end{aligned}$$

Therefore,

$$\frac{E \left\{ \check{\theta}^T(k) R(k) \check{\theta}(k) \middle| \Gamma_{k-1} \right\}}{r^{2-\varepsilon}(k)} \rightarrow 0, a.s., k \rightarrow \infty.$$

and

$$\left\| \tilde{\theta}(k) - \theta \right\|^2 = o \left\{ \frac{r^{2-\varepsilon}(k)}{\lambda_{\min}[R(k)]} \right\}, a.s..$$

The proof of this Theorem is completed.

5 Numerical Examples

Example 1: Consider a SISO linear output error model with multi-threshold quantized observations which has been established based on pattern-moving and hybrid metrics as follows.

$$\left\{ \begin{array}{l} \hat{dx}(k) = 0.5\overline{dx}(k-1) + 0.3\overline{dx}(k-2) + u(k-1) + 0.7u(k-2) + v(k) \\ s(k) = D(M(\hat{dx}(k))) = \begin{cases} c_1, c_1 - r_1 \leq \hat{dx}(k) \leq c_1 + r_1; \\ c_2, c_2 - r_2 < \hat{dx}(k) \leq c_2 + r_2; \\ \vdots \\ c_{13}, c_{13} - r_{13} < \hat{dx}(k) \leq c_{13} + r_{13}, \end{cases} \end{array} \right.$$

where $v(k) \sim \mathcal{N}(0, 0.25^2)$ denotes the model noise; $u(k) \sim \mathcal{N}(0, 2^2)$ represents the system input; $\theta = [a_1, a_2, b_1, b_2]^T = [0.5, 0.3, 1.0, 0.7]^T$ is the model parameter vector to be identified; the detail property values of pattern classes are known and shown in [Table 1](#).

Table 1: Property values of pattern classes for Example 1

Class No.	Class center c_i	Class radius r_i	Threshold C_i
1	-7.8750	2.6250	-5.2500
2	-3.9375	1.3125	-2.6250
3	-1.9688	0.6563	-1.3125
4	-0.9844	0.3281	-0.6563
5	-0.4922	0.1641	-0.3281
6	-0.2461	0.0820	-0.1641
7	0.0000	0.1641	0.1641
8	0.2461	0.0820	0.3281
9	0.4922	0.1641	0.6563
10	0.9844	0.3281	1.3125
11	1.9688	0.6563	2.6250
12	3.9375	1.3125	5.2500
13	7.8750	2.6250	10.5000

The size of input cycle is initially set as $t = 3000$, and the number of cycles is $l = 300$. The initial conditions are $\overline{dx}(1 + i \cdot l) = \overline{dx}(2 + i \cdot l) = 0; i = 0, 1, \dots, l - 1$. It can be collected 900000 groups of data with the corresponding input sequence $\{u(k)\}$ and output sequence $\{s(k)\}$. Since the distribution of model noise and pattern class thresholds are known, $\{\tilde{dx}(k)\}$ can be estimated by the sequence $\{ss^T(k)\}$. Using the input series $\{u(k)\}$ and $\{\tilde{dx}(k)\}$, the model parameters can be identified by the designed M-AM-SGRA, and the identification processes with different convergence indexes such as $\varepsilon = 0.99, 0.98, 0.97, 0.96$ respectively are shown in [Fig. 1](#). The parameter estimates and the estimation error ratios ($\delta = \|\tilde{\theta}(k) - \theta\| / \|\theta\|$) with different convergence indexes at time instant $k = 3000$ are shown in [Table 2](#).

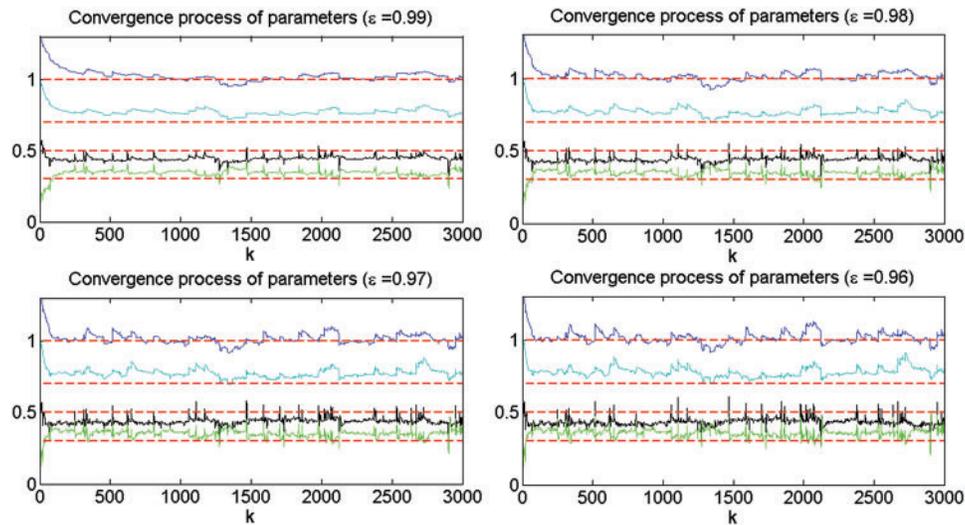


Figure 1: Convergence process of the M-AM-SGRA for Example 1

Table 2: The parameter estimates and their errors ratios for Example 1

ε	a_1	a_2	b_1	b_2	Δ (%)
0.99	0.4544	0.3672	1.0162	0.7609	7.92
0.98	0.4525	0.3752	1.0223	0.7677	8.00
0.97	0.4429	0.3791	1.0337	0.7804	8.39
0.96	0.4312	0.3858	1.0481	0.7957	9.03
The values	0.5	0.3	1.0	0.7	

It is shown from Fig. 1 and Table 2 that the convergence index ε is smaller, the parameter estimate converges faster, and the estimation error ratio becomes larger. One can get a good strategy which is to choose a small convergence index at the beginning, and then let it gradually increase with time instant k , and finally approach unity. Therefore, the system identification accuracy can be satisfied by choosing and adjusting the appropriate convergence index ε .

Example 2: Consider a SIMO unknown nonlinear system

$$\begin{cases} y_1(k) = \frac{u(k-1)}{1+u^2(k-1)} + u(k-2) + \vartheta(k) \\ y_2(k) = 0.1y_2(k-1) + 0.5y_2(k-2) + \frac{u(k-1)}{1+u^2(k-1)} + u(k-2) + \vartheta(k) \\ y_3(k) = 0.2y_3(k-1) + \frac{u(k-1)}{1+u^2(k-1)} + u(k-2) + \vartheta(k) \end{cases}$$

where $\vartheta(k) \sim \mathcal{N}(0, 0.1^2)$ denotes the system noise; $u(k) \sim \mathcal{N}(0, 1)$ is the system input.

1) Design of System input. The size of input cycle and the number of cycle are set the same as Example 1. The initial condition is set as $y_j(1+i \cdot l) = y_j(2+i \cdot l) = 0; i = 0, 1, \dots, l-1; j = 1, 2, 3$. It can be collected 900000 groups of data with the corresponding input sequence $\{u(k)\}$ and output sequences $\{y_j(k)\}$. The outputs are normalized and a PCA method [32] is used to handle the normalized data to get the first principal component information $\{y(k)\}$.

2) Classification and determining the actual output $s(k)$, class quantity N , class center c_i , class radius r_i , and class threshold C_i . A modified quantized control classification method ($M(\cdot)$) [39] is adopted and described in the following formula:

$$s(k+1) = D(M(y(k+1))) = \begin{cases} \check{y}(k+1), & \text{if } \frac{1}{1+\Delta}\kappa_i < y(k+1) \leq \frac{1}{1-\Delta}\kappa_i \\ 0, & \text{if } -\frac{1}{1+\Delta}\kappa_N < y(k+1) \leq \frac{1}{1+\Delta}\kappa_N \\ -\check{y}(k+1), & \text{if } -\frac{1}{1-\Delta}\kappa_i < y(k+1) \leq -\frac{1}{1+\Delta}\kappa_i \end{cases} \quad (34)$$

where $\check{y}(k+1) = \frac{1+\rho_0}{4}\kappa_i(\rho_0^{i-1} + \rho_0^i)$; $\rho_0 \in (0, 1)$; $\kappa_i = \rho_0^i \kappa_0$; $\Delta = \frac{1-\rho_0}{1+\rho_0}$; κ_0 is the maximum working range of $y(k)$ ($\kappa_0 \geq \max(|y(k)|)$); $i = 1, 2, \dots, N$.

Given the upper limit of the initial class radius r_0 at the working point 0 and the other parameters such as ρ_0, κ_0 , one can obtain $L \geq \frac{\ln(r_0 \frac{1+\Delta}{\kappa_0})}{\ln \rho_0}$, and the first principal component information $\{y(k)\}$ is divided into $2L + 1$ segments. Furthermore, the number of pattern class ($N = 2L + 1$), the class center c_i , class radius $r_i = \frac{1+\rho^2}{4\rho}$ and threshold value C_i can be obtained respectively, $i = 1, \dots, N$. The initial parameters are set as $\rho_0 = 0.5, \kappa_0 = 5, r_0 = 0.4, L = \frac{\ln(r_0 \frac{1+\Delta}{\kappa_0})}{\ln \rho_0}$. The inputs ($u(k)$), first principal component information ($y(k)$) and class center metric values ($s(k)$) in the first cycle are shown in Fig. 2, and the detail property values of pattern classes are shown in Table 3.

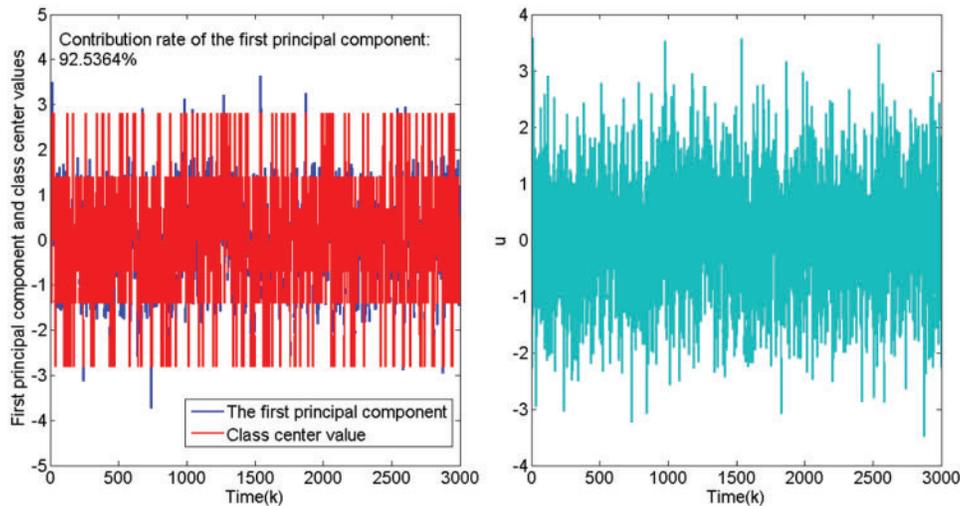


Figure 2: The distribution of $u(k)$, $y(k)$ and $s(k)$ for Example 2

Table 3: Property values of pattern classes for Example 2

Class No.	Class center c_i	Class radius r_i	Threshold C_i
1	-2.8125	0.9375	-1.8750
2	-1.4063	0.4688	-0.9375
3	-0.7031	0.2344	-0.4688
4	-0.3516	0.1172	-0.2344
5	0.0000	0.2344	0.2344
6	0.3516	0.1172	0.4688
7	0.7031	0.2344	0.9375
8	1.4063	0.4688	1.8750
9	2.8125	0.9375	3.7500

3) Establishing a pattern-moving-based linear output error model with multi-threshold quantized observations. According to the information obtained from the above steps, the pattern-moving-based system dynamics description Eqs. (2) and (3) is to be obtained. The input and output orders are assumed to be known ($m = n = 2$), and the distribution of model noise is assumed to be known ($\vartheta(k) \sim \mathcal{N}(0, 0.1^2)$). Therefore, a following linear output error model with multi-threshold quantized observations is established to approximate the system (2)–(3).

$$\left\{ \begin{array}{l} \hat{dx}(k) = a_1 \overline{dx}(k-1) + a_2 \overline{dx}(k-2) + b_1 u(k-1) + b_2 u(k-2) + v(k); \\ s(k) = D(M(\hat{dx}(k))) = \begin{cases} c_1, c_1 - r_1 \leq \hat{dx}(k) \leq c_1 + r_1; \\ c_2, c_2 - r_2 < \hat{dx}(k) \leq c_2 + r_2; \\ \vdots \\ c_9, c_9 - r_9 < \hat{dx}(k) \leq c_9 + r_9, \end{cases} \end{array} \right.$$

where $\theta = [a_1, a_2, b_1, b_2]^T$ is the model parameter vector to be identified.

4) Model parameters identification. The designed M-AM-SGRA is employed to identify the model parameters. Referring to the conclusion of Example 1, the convergence factor ε is set by segments in this example. Fig. 3 shows the curves of parameters identification with a fixed value $\varepsilon = 0.995$ and some interval variable values such as $\varepsilon = 0.9, k \leq 100$; $\varepsilon = 0.96, k \leq 200$; $\varepsilon = 0.99, k \leq 500$; $\varepsilon = 0.995, k \leq 1000$; $\varepsilon = 0.999, k > 1000$. At time instant $k = 3000$, the estimation vector of θ is $[0.0673, 0.1655, 0.3315, 0.9200]^T$ while $\varepsilon = 0.995$, and it is $[0.0815, 0.1577, 0.3270, 0.9278]^T$ while ε is set as the interval variable value. Fig. 3 shows that using the time-varying convergence indexes can not only obtain a fast convergence speed at the beginning, but also gradually achieve a stable convergence effect.

Remark 5.1 It is known to all that there exists many classifications and clustering methods in pattern recognition technology, such as C-means, ISODATA, and so on. Here a modified quantized control classification and class center explicit metric method is utilized. The initial parameters setting principle is based on the quality parameters of the product. In Example 2, it is assumed that the good quality is obtained with the output $y(k) \in (-0.2344, 0.2344]$, so the initial parameters are set to ensure that the working condition data $y(k) \in (-0.2344, 0.2344]$ belongs to one pattern class.

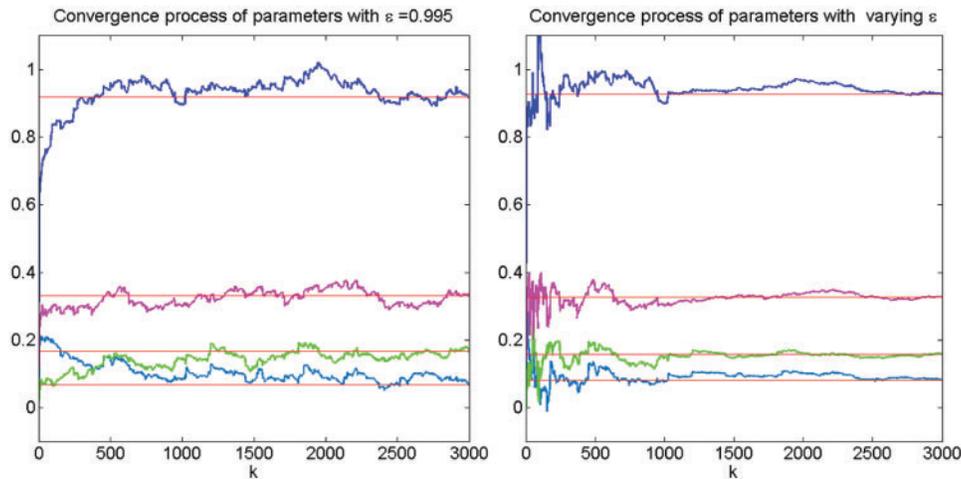


Figure 3: Convergence process of the M-AM-SGRA for Example 2

6 Conclusions

According to the characteristics of classification and hybrid metrics, mappings are in line with the set-valued system, a linear output error model with multi-threshold quantized observations is adopted to approximate the unknown system, and an M-AM-SGRA is designed with its convergence proved. Finally, the validity and desired effect of the parameter identification algorithm is demonstrated by two numerical examples. Future works will focus on pattern-moving-based set-valued system modeling and control methods for achieving optimal robustness.

Funding Statement: This work was supported by the National Natural Science Foundation of China (62076025).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Qu, S. D. (1998). Pattern recognition approach to intelligent automation for complex industrial processes. *Journal of University of Science and Technology Beijing*, 20(4), 385–389. DOI 10.13374/j.issn1001-053x.1998.04.018.
2. Saridis, G. N. (1981). Application of pattern recognition methods to control systems. *IEEE Transactions on Automatic Control*, 26(3), 638–645. DOI 10.1109/TAC.1981.1102685.
3. Li, X., Zhang, L. (2011). Intelligent controller based on pattern recognition. *ICECC*, 1512–1515. DOI 10.1109/ICECC.2011.6066661.
4. Yu, B., Zhang, X., Wu, L., Chen, X. (2020). A novel postprocessing method for robust myoelectric pattern-recognition control through movement pattern transition detection. *IEEE Transactions on Human-Machine Systems*, 50(1), 32–41. DOI 10.1109/THMS.2019.2953262.
5. Xu, Z. G. (2001). *Pattern recognition method of intelligent automation and its implementation in engineering (Ph.D. Thesis)*. University of Science and Technology Beijing
6. Wang, M. S., Xu, Z. G., Guo, L. L. (2017). Pattern recognition based dynamics description of production processes in metric spaces. *Asian Journal of Control*, 19(4), 1–13. DOI 10.1002/asjc.1471.
7. Wang, M. S., Xu, Z. G., Guo, L. L. (2019). Stability and stabilization for a class of complex production processes via LMIs. *Optimal Control Applications and Methods*, 40(3), 460–478. DOI 10.1002/oca.2488.

8. Sun, C. P., Xu, Z. G. (2016). Multi-dimensional moving pattern prediction based on multi-dimensional interval T-S fuzzy model. *Control and Decision*, 31(9), 1569–1576. DOI 10.13195/j.kzyjc.2015.0944.
9. Guo, L. L., Xu, Z. G., Wang, Y. (2014). Dynamic modeling and optimal control for complex systems with statistical trajectory. *Discrete Dynamics in Nature & Society*, 2014(1), 1–8. DOI 10.1155/2015/245685.
10. Xu, Z. G., Wu, J. X., Guo, L. L. (2013). Modeling and optimal control based on moving pattern. *Chinese Control Conference*, pp. 7894–7899. Xi'an, China.
11. Xu, Z. G., Wu, J. X. (2012). Data-driven pattern moving and generalized predictive control. *IEEE International Conference on Systems, Man and Cybernetics*, 1604–1609. DOI 10.1109/ICSMC.2012.6377966.
12. Ljung, L. (1999). *System identification: Theory for the user*. 2nd ed. Englewood Cliffs: Prentice-Hall.
13. Clarke, D. W. (1967). Generalized least squares estimation of the parameters of a dynamic model. *Proceeding of IFAC Symposium on Identification in Automatic Control Systems*. Prague.
14. Astrom, K. J. (1980). Maximum likelihood and prediction error methods. *Automatica*, 16(5), 551–574. DOI 10.1016/0005-1098(80)90078-3.
15. Zhao, S. Y., Huang, B. (2020). Trial-and-error or avoiding a guess? Initialization of the Kalman filter. *Automatica*, 121(21), 109184. DOI 10.1016/j.automatica.2020.109184.
16. Zhao, S. Y., Yuriy, S. S., Ahn, C., Liu, F. (2018). Adaptive-horizon iterative UFIR filtering algorithm with applications. *IEEE Transactions on Industrial Electronics*, 65(8), 6393–6402. DOI 10.1109/TIE.2017.2784405.
17. Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1), 35–45. DOI 10.1115/1.3662552.
18. Ho, Y., Lee, R. (1964). A Bayesian approach to problems in stochastic estimation and control. *IEEE Transactions on Automatic Control*, 9(4), 333–339. DOI 10.1109/TAC.1964.1105763.
19. Zhao, Y. L., Zhang, J. F., Guo, J. (2012). System identification and adaptive control of set-valued systems. *Journal of Systems Science and Mathematical Sciences*, 32(10), 1257–1265. DOI 10.12341/jssms11944.
20. Wang, T., Bi, W., Zhao, Y. L., Xue, W. (2016). Radar target recognition algorithm based on RCS observation sequence set-valued identification method. *Journal of Systems Science and Complexity*, 29(3), 573–588. DOI 10.1007/s11424-015-4151-8.
21. Wang, L. Y., Zhang, J. F., Yin, G. G. (2003). System identification using binary sensors. *IEEE Transactions on Automatic Control*, 48(11), 1892–1907. DOI 10.1109/TAC.2003.819073.
22. Wang, L. Y., Yin, G. G., Zhang, J. F., Zhao, Y. L. (2008). Space and time complexities and sensor threshold selection in quantized identification. *Automatica*, 44(12), 3014–3024. DOI 10.1016/j.automatica.2008.04.022.
23. Wang, L. Y., Yin, G., Zhang, J. F., Zhao, Y. L. (2010). *System identification with quantized observations, theory and applications*. Birkhäuser, Boston: Springer.
24. Zhao, Y. L., Wang, L. Y., Yin, G. G., Zhang, J. F. (2007). Identification of Wiener systems with binary-valued output observations. *Automatica*, 43(10), 1752–1765. DOI 10.1016/j.automatica.2007.03.006.
25. Zhao, Y. L., Zhang, J. F., Wang, L. Y., Yin, G. G. (2010). Identification of Hammerstein systems with quantized observations. *SIAM Journal on Control and Optimization*, 48(7), 4352–4376. DOI 10.1137/070707877.
26. Guo, J., Wang, L. Y., Yin, G. G., Zhao, Y. L. (2017). Identification of Wiener systems with quantized inputs and binary-valued output observations. *Automatica*, 78(12), 280–286. DOI 10.1016/j.automatica.2016.12.034.
27. Wang, T., Tan, J. W., Zhao, Y. L. (2018). Asymptotically efficient nontruncated identification for FIR systems with binary-valued outputs. *Science China Information Sciences*, 61(12), 052201. DOI 10.1007/s11432-018-9646-7.
28. Li, X. Q., Xu, Z. G., Cui, J. R., Zhang, L. X. (2021). Suboptimal adaptive tracking control for FIR systems with binary-valued observations. *Science China Information Sciences*, 64(7), 172202: 1–172202: 11. DOI 10.1007/s11432-020-2914-2.
29. Wang, T., Zhang, H., Zhao, Y. L. (2019). Parameter estimation based on set-valued signals: Theory and application. *Acta Mathematicae Applicatae Sinica, English Series*, 35(2), 255–263. DOI 10.1007/s10255-019-0822-x.

30. Ding, F. (2017). *System identification: Auxiliary model identification idea and methods*. Beijing, Science-Press.
31. Xie, L. B., Ding, F., Wang, Y. (2009). Auxiliary model-based identification method for quantized control systems. *Control Theory and Applications*, 26(3), 277–282.
32. Li, X. Q., Xu, Z. G., Lu, Y. L., Cui, J. R., Zhang, L. X. (2021). Modified model free adaptive control for a class of nonlinear systems with multi-threshold quantized observations. *International Journal of Control, Automation and Systems*. DOI 10.1007/s12555-020-0289-9.
33. Yang, X., Gao, J. J., Li, L. L., Luo, H., Ding, S. X. et al. (2020). Data-driven design of fault-tolerant control systems based on recursive stable image representation. *Automatica*, 122(6), 109246. DOI 10.1016/j.automatica.2020.109246.
34. Yang, X., Gao, J. J., Huang, B. (2021). Data-driven design of fault detection and isolation method for distributed homogeneous systems. *Journal of the Franklin Institute*, 358(9), 4929. DOI 10.1016/j.jfranklin.2021.04.016.
35. Yang, J., Zhang, D., Frangi, A. F., Yang, J. Y. (2004). Two-dimensional PCA: A new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1), 131–137. DOI 10.1109/TPAMI.2004.1261097.
36. Xu, Z. G., Qu, S. D., Yu, L. (2003). Self-learning pattern recognition method based on statistical space mapping. *Journal of University of Science and Technology Beijing*, 25(5), 480–482. DOI 10.13374/j.issn1001-053x.2003.05.050.
37. Guo, J., Yin, G., Zhao, Y., Zhang, J. F. (2015). Asymptotically efficient identification of FIR systems with quantized observations and general quantized inputs. *Automatica*, 57(5), 113–122. DOI 10.1016/j.automatica.2015.04.009.
38. Goodwin, G. C., Sin, K. S. (1984). *Adaptive filtering, prediction and control*. New Jersey: Prentice-Hall.
39. Guo, G. J., Chen, T. W. (2008). A new approach to quantized feedback control systems. *Automatica*, 44(2), 534–542. DOI 10.1016/j.automatica.2007.06.015.