Joint Identification of System Parameter and Noise Parameters in Quantized Systems ⋆

Jieming Ke [a,b], Yanlong Zhao [a,b], Ji-Feng Zhang [a,b],

[a] *Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, P. R. China.*

[b] *School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, P. R. China.*

Abstract

This paper investigates the joint identification problem of unknown system parameter and noise parameters in quantized systems when the noises involved are Gaussian with unknown variance and mean value. Under such noises, previous investigations show that the unknown system parameter and noise parameters are not jointly identifiable in the single-threshold quantizer case. The joint identifiability in the multi-threshold quantizer case still remains an open problem. This paper proves that the unknown system parameter, the noise variance and the mean value are jointly identifiable if and only if there are at least two thresholds. Then, a decomposition-recombination identification algorithm is proposed to jointly identify the unknown system parameter and noise parameters. Firstly, a technique is designed to convert the identification problem with unknown noise parameters into an extended parameter identification problem with standard Gaussian noises. Secondly, the extended parameter is identified by a stochastic approximation method for quantized systems. For the effectiveness, this paper obtains the strong consistency and the L^p convergence for the algorithm under non-persistently exciting inputs and without any *a priori* knowledge on the range of the unknown system parameter. The almost sure convergence rate is also obtained. Furthermore, when the mean value is known, the unknown system parameter and noise variance can be jointly identified under weaker conditions on the inputs and the quantizer. Finally, the effectiveness of the proposed algorithm is demonstrated by simulation.

Key words: Quantized systems, stochastic systems, recursive identification, stochastic approximation, non-persistent excitations.

1 Introduction

In identification problems, quantized systems have been extensively studied over the past two decades due to the practical importance. Such systems widely exist in practical fields.

i) *Genome-Wide Association Studies* [1]: Quantized phenotypes (e.g., survival or relapse, and the scoring of health behaviour-related preferences) are common in genome-wide association studies. Therefore, the association between phenotypes and genotypes can be modelled as quantized systems.

ii) *Lithology Classification* [2]: New well-logging lithology classification methods based on inaccurate quantized observations receive increasing attention in recent years. Compared with traditional methods based on core recovery or well-log analysis, the quantized observation-based methods have advantages of low labour cost and low time cost.

iii) *Air-to-Fuel Ratio Control* [3]: In heated exhaust gas oxygen sensors, the voltage outputs shapely change their value when the air-to-fuel ratio exceeds a certain threshold. Therefore, heated exhaust gas oxygen sensors can be treated as binary sensors.

More practical examples of the quantized systems are referred to [4]. To analyze and control such systems well, it is necessary to investigate the identification problem of quantized systems.

However, the quantized system identification problem is difficult because of the little information contained in the quantized observations. There are mainly three tech-

⋆ The work is supported by National Natural Science Foundation of China under Grants 62025306 and T2293770, CAS Project for Young Scientists in Basic Research under Grant YSBR-008. This paper was not presented at any conference. Corresponding author: Yanlong Zhao.

Email addresses: `kejieming@amss.ac.cn` (Jieming Ke), `ylzhao@amss.ac.cn` (Yanlong Zhao), `jif@iss.ac.cn` (Ji-Feng Zhang).

niques to overcome the difficulty. The first technique is to design adaptive quantizers. For example, by properly adjusting thresholds in quantizers, [5–7] propose sign-error type identification algorithms. The second technique is based on the statistical characteristic of noises. For example, by using the full knowledge on noise distribution, [4, 8] propose empirical measurement methods, [9, 10] propose maximum likelihood methods, and [11–13] propose recursive projection methods. The third technique is to use strongly exciting inputs. For example, [14, 15] consider independent and identically distributed (i.i.d.) and sufficiently rich inputs, and proposes a stochastic approximation (SA) algorithm with expanding truncations.

It is worth mentioning that, for practical use, these three techniques have their own limitations. For the first technique, adaptive quantizers are not common in practical scenarios. Instead, many practical quantizers are fixed [1–3]. For the second technique, full knowledge of noise distribution is hard to be known in practice, since, for example, the variance of noises is often unavailable. For the third technique, strong input assumptions limit the application of the algorithms. Therefore, this paper considers fixed quantizers, and tries to investigate the identification problem under weaker assumptions on inputs, noises, and the unknown system parameter.

The identification problem of quantized systems has been widely investigated under different excitation conditions on inputs [4–6, 11–13, 16–21]. Excitation conditions can be classified as persistently exciting (PE) ones and non-persistently exciting (non-PE) ones. For example, when the inputs are deterministic, [17, 19, 20] assume the inputs to be periodic and full rank. [11, 16, 18] extend the results to the uniform PE input case. Furthermore, [9, 12] consider general PE conditions. When the inputs are stochastic, [21] assumes the inputs to be strong-mixing and sufficiently rich. [5] simultaneously considers deterministic and stochastic inputs, and give corresponding PE condition. Excitation conditions in all these works are PE conditions. On the other hand, [13] considers a non-PE condition, because the non-PE condition is more general and easier to be guaranteed in feedback control systems.

There are a series of important achievement in identifying quantized systems without the full knowledge of noise distribution [14, 19, 20, 22]. Under periodic inputs, [19, 20] figure out that a parameterizable model of noise distribution is necessary to obtain the unknown system parameter in quantized systems. Then, they propose empirical measurement methods to identify the system parameter and the noise parameters. Specially, [20] shows that when the noise is Gaussian with unknown variance and mean value, the unknown system parameter and noise parameters are not jointly identifiable. [22] considers the Gaussian noises with unknown variance and zero mean, and proposes an offline maximum likelihood algorithm to jointly identify the system parameter and the variance. [14] also considers the Gaussian noises with unknown variance and zero mean, and proposes an SA algorithm with expanding truncations to jointly identify the system parameter and the threshold in quantizers.

A prior knowledge on system parameter is required in some identification methods for quantized systems [9, 13, 21]. For example, [21] points out that to accurately identify the quantized system, the norm of the system parameter should be known *a priori* in the noise-free and zero threshold case. [9, 13] propose an online maximum likelihood algorithm and an adaptive projection-based algorithm, respectively. To prove the strong consistency of the algorithms, they assumes that the unknown parameter is in a known compact set. There are important attempts in identifying the quantized systems without *a priori* knowledge on parameter range [5, 14, 18, 23]. For example, [18, 23] propose identification algorithms without truncations or projections. They prove the convergence of the algorithms by estimating distribution tails of the algorithms. Besides, [5, 14] introduce expanding truncations. The design of such truncations does not rely on *a priori* knowledge for the range of unknown parameter.

Despite the remarkable progress in the identification methods for quantized systems, there are still important issues deserving to be investigated. Firstly, when the noise is Gaussian with unknown noise variance and mean value, the joint identifiability of the unknown system parameter and noise parameters in the multi-threshold quantizer case still remains an open problem. Secondly, [13] gives a non-PE condition for the identification problem of quantized systems, but the condition is not sufficient to obtain the unknown parameter when the noise variance and mean value are both unknown. Therefore, it seems that a new non-PE condition is needed on inputs when there are unknown noise parameters. Thirdly, although [5, 14, 18, 23] do not require *a prior* knowledge on unknown parameter, their techniques are difficult to extend to the non-PE input case. Therefore, we should develop new techniques for the problem.

The paper proposes a decomposition-recombination identification algorithm to jointly identify the unknown system parameter and the Gaussian noise parameters in the quantized system. The main contributions of this paper can be summarized as follows:

i) We prove that the unknown system parameter and noise parameters are jointly identifiable if and only if there are at least two thresholds in quantizers when the noises are Gaussian with unknown variance and mean value. Therefore, for the identification problems of quantized systems, the zero-mean assumption [14, 22] can be removed by increasing threshold numbers.

ii) The excitation condition for the inputs is non-PE, which is weaker than the PE condition commonly adopted in existing literature [9, 14, 19, 20]. Under the non-PE condition, the strong consistency and the L^p convergence are achieved. Besides, the almost sure convergence rates is obtained under PE condition and a class of non-PE conditions. The almost sure convergence rates are consistent with the best ones for the classical SA-based identification algorithms [24–26].

iii) Any *a priori* knowledge on the range of unknown system parameter is not required for the convergence of the proposed algorithm. A new almost sure boundedness analysis technique based on an auxiliary stochastic process is proposed to achieve this goal. The new technique has wider applicability compared with existing ones [18, 23], because it does not rely on any excitation condition on inputs.

The rest of the paper is organized as follows. Section 2 formulates the identification problems. Section 3 gives the equivalent condition of the joint identifiability for the unknown system parameter and noise parameters. Section 4 proposes the joint identification algorithm. Section 5 analyzes the convergence properties of the proposed algorithm. Section 6 uses a numerical example to demonstrate the main results. Section 7 gives concluding remarks and future works.

Notation. In the rest of the paper, $\mathbb{R}$ and $\mathbb{R}^n$ are the sets of real numbers and n-dimensional real vectors, respectively. $\mathbb{I}_{\{\cdot\}}$ denotes the indicator function, whose value is 1 if its argument (a formula) is true, and 0, otherwise. Given positive integer p, $\|x\|_p$ is the L^p norm for vector x. Specially, $\|x\|_2$ is abbreviated as $\|x\|$. $\|A\|$ is the matrix Euclidean norm for matrix A. I_n is an $n \times n$ identity matrix. $\lfloor x \rfloor$ is the largest integer that is smaller than or equal to $x \in \mathbb{R}$. $\operatorname{diag}\{\cdot\}$ is the diagonal matrix of the corresponding numbers. $\operatorname{col}\{\cdot\}$ is the column matrix stacked by the corresponding matrices. $\operatorname{rank}(\cdot)$ is the rank of the corresponding matrix. For square matrices $A_l, \dots, A_k$, denote $\prod_{i=l}^{k} A_i = A_k \cdots A_l$ for $k \geq l$. Given two sequences $\{a_k\}$ and $\{b_k\}$, denote $a_k = O(b_k)$ if there is a bounded c_k such that $a_k = c_k b_k$ for all k.

2 Problem formulation

This section will formulate the joint identification problem for quantized systems. Consider the following system:

$$y_k = \phi_k^\top \theta + d_k, \quad k \geq 1, \tag{1}$$

where $\theta \in \mathbb{R}^n$ is the unknown constant system parameter. $\{y_k\}$ and $\{d_k\}$ are the sequences of the unobserved system outputs and the system noises, respectively. The regressor ϕ_k is a function of system input sequence $\{u_k\}$. One of the examples of the system (1) is the finite impulse response system, where $\phi_k = [u_k \ \dots \ u_{k-n+1}]^\top$.

The system output y_k cannot be obtained accurately, but can be measured by a quantized sensor with q thresholds $-\infty < C_1 < \cdots < C_q < \infty$. Therefore, the measurement of y_k can be represented as

$$z_k = \sum_{i=1}^{q} \mathbb{I}_{\{y_k > C_i\}} = \begin{cases} 0, & \text{if } y_k \leq C_1; \\ 1, & \text{if } C_1 < y_k \leq C_2; \\ \vdots & \vdots \\ q, & \text{if } C_q < y_k. \end{cases} \tag{2}$$

Remark 2.1 (2) *is an alternative representation of the finite-level quantizer [8, 27, 28]. The output set of the quantizer* $\{0, \dots, q\}$ *can be replaced by any finite set with* q *elements [7, 22].*

Remark 2.2 (2) *gives a general form of finite-level fixed quantizers. Examples of the quantizers include binary-valued quantizers [4], finite-level uniform quantizers [29], and the finite-level logarithmic quantizers [30].*

The noise sequence $\{d_k\}$ is assumed to be i.i.d. Gaussian with unknown variance σ^2 and mean value μ. When $q = 1$, [20] shows that θ, σ^2 and μ is not jointly identifiable. Therefore, existing literature [14, 22] always assumes that the mean value μ is known to be zero. However, the mean value μ is not always available. Then, is it possible to jointly identify θ, σ^2 and μ when $q \geq 2$? If so, how to design the joint identification algorithm?

3 Joint identifiability

The section will discuss the joint identifiability for the system parameter and the noise parameter.

Suppose that the cumulative distribution function $F(x)$ for the noise d_k can be parameterized by a vector α, and thereby can be written as $F(x; \alpha)$. Then, we extend the definition of joint identifiability for the system parameter θ and the noise parameter α in [20] to the multi-threshold quantizer case.

Definition 1 (Joint identifiability) *For given thresholds* $C_1 < C_2 < \cdots < C_q$*, the system parameter* θ *and the noise parameter* α *are said to be jointly identifiable if there exists a positive integer* k *and a sequence* $\{\phi_1, \dots, \phi_k\}$ *such that*

$$\mathbb{F}(\theta; \alpha) = \Big[F(C_1 - \phi_1^\top \theta; \alpha) \ \cdots \ F(C_1 - \phi_k^\top \theta; \alpha) \\ \cdots \ F(C_q - \phi_1^\top \theta; \alpha) \ \cdots \ F(C_q - \phi_k^\top \theta; \alpha)\Big]^\top$$

is injective as a function of θ *and* α*.*

Then, consider Gaussian noise distribution with zero mean and variance σ^2. The following theorem gives an equivalent condition of the joint identifiability for the system parameter θ and the noise parameter $\alpha = [\mu\ \sigma^2]^\top$.

Theorem 3.1 *For Gaussian cumulative distribution function $F(x; [\mu\ \sigma^2])$, assume that $\sigma^2 > 0$. Then, the system parameter θ and the noise parameter $[\mu\ \sigma^2]$ are jointly identifiable if and only if $q \geq 2$.*

Proof. Since [20] has proved that θ and $[\mu\ \sigma^2]$ are not jointly identifiable when $q = 1$, it suffices to prove the joint identifiability when $q \geq 2$.

Now, we prove that if

$$\sum_{t=1}^{k} \begin{bmatrix} \phi_t \phi_t^\top & \phi_t \\ \phi_t^\top & 1 \end{bmatrix} \tag{3}$$

is positive definite, then $\mathbb{F}(\theta; [\mu\ \sigma^2])$ is injective.

Denote $F_0(\cdot)$ and $f_0(\cdot)$ as the cumulative distribution function and density function of the standard Gaussian distribution. Note that $F(x; [\mu\ \sigma^2]) = F_0\left(\frac{x-\mu}{\sigma}\right)$. Then, one can get

$$\begin{aligned}&\frac{\partial}{\partial[\theta^\top, \mu, \sigma]^\top} \mathbb{F}(\theta; [\mu\ \sigma^2]) \\ =&\operatorname{diag}\{f_{1,1}, \ldots, f_{1,k}, \ldots, f_{q,1}, \ldots, f_{q,k}\} \operatorname{col}\{\Xi_1, \ldots, \Xi_q\}.\end{aligned}$$

where $f_{i,k} = f_0\left(\frac{C_i - \phi_k^\top \theta - \mu}{\sigma}\right)$ and

$$\Xi_i = -\frac{1}{\sigma} \begin{bmatrix} \phi_1^\top & 1 & \frac{C_1 - \phi_1^\top \theta - \mu}{\sigma} \\ \vdots & \vdots & \vdots \\ \phi_k^\top & 1 & \frac{C_1 - \phi_k^\top \theta - \mu}{\sigma} \end{bmatrix}.$$

Note that $-\sigma \Xi_i \begin{bmatrix} I_n & & \theta \\ & 1 & \mu \\ & & \sigma \end{bmatrix} = \begin{bmatrix} \phi_1^\top & 1 & C_i \\ \vdots & \vdots & \vdots \\ \phi_k^\top & 1 & C_i \end{bmatrix}$. Then, by the positive definiteness of (3), $\operatorname{rank}(\operatorname{col}\{\Xi_1, \ldots, \Xi_q\}) = n+1$.

Hence, by the inverse function theorem and the rank theorem (Theorems 8.6.1 and 8.6.2 of [31]), one can get $\mathbb{F}(\theta, [\mu\ \sigma^2])$ is injective, which implies that θ and $[\mu\ \sigma^2]$ are jointly identifiable when $q \geq 2$. □

Remark 3.1 *Although θ, σ^2 and μ are not jointly identifiable if $q = 1$, Theorem 3.1 shows that they can become jointly identifiable when increasing threshold number q. Similar phenomena also occur in other types of cumulative distribution functions. For example, if the noise is uniformly distributed with unknown upper bound $\bar{b}$ and lower bound $\underline{b}$, then θ, $\bar{b}$ and $\underline{b}$ are jointly identifiable if and only if $q \geq 2$.*

4 Joint identification algorithm

This section will propose a decomposition-recombination algorithm to jointly identify the system parameter θ, the variance σ^2 and the mean value μ.

1) Decomposition. We decompose the q-threshold quantized system with unknown noise variance and mean value into several binary-valued sub-systems with known noise distribution. By (2), we have $z_k = \sum_{i=1}^{q} s_{i,k}$, where

$$\begin{aligned} s_{i,k} =& \mathbb{I}_{\{y_k > C_i\}} = \mathbb{I}_{\left\{\phi_k^\top \frac{\theta}{\sigma} - \frac{C_i}{\sigma} + \frac{\mu}{\sigma} + \frac{d_k - \mu}{\sigma} > 0\right\}} \\ =& \mathbb{I}_{\{\varphi_{i,k}^\top \vartheta + \mathscr{d}_k > 0\}} = \mathbb{I}_{\{\mathscr{y}_{i,k} > 0\}}, \end{aligned}$$

where

$$\begin{aligned} \vartheta &= \begin{bmatrix} \frac{\theta}{\sigma} & \frac{1}{\sigma} & \frac{\mu}{\sigma} \end{bmatrix}^\top, \qquad \mathscr{d}_k = \frac{d_k - \mu}{\sigma}, \\ \varphi_{i,k} &= \begin{bmatrix} \phi_k^\top & -C_i & 1 \end{bmatrix}^\top, \ \mathscr{y}_{i,k} = \varphi_{i,k}^\top \vartheta + \mathscr{d}_k. \end{aligned} \tag{4}$$

Then, the quantized system (1)-(2) can be decomposed into q binary-valued sub-systems

$$\begin{cases} \mathscr{y}_{i,k} = \varphi_{i,k}^\top \vartheta + \mathscr{d}_k, \\ s_{i,k} = \mathbb{I}_{\{\mathscr{y}_{i,k} > 0\}}, \end{cases} \quad i = 1, \ldots, q. \tag{5}$$

In binary-valued sub-systems (5), $\{\mathscr{d}_k\}$ is a standard Gaussian noise sequence. Besides, the unknown system parameters θ and the variance of noise σ^2 can be uniquely determined by ϑ. Therefore, the decomposition transforms the identification problem into one with known noise distribution.

Remark 4.1 *Similar decomposition technique can be applied in other type of noises, such as the Laplacian noise [32].*

2) Recombination. We use the innovations of binary-valued sub-systems (5) to update the estimate of ϑ based on SA method.

For the i-th sub-system of (5), consider the instantaneous quadratic error $\hat{e}_{i,k}(\hat{\vartheta}) = (s_{i,k} - F_0(\varphi_{i,k}^\top \hat{\vartheta}))^2$. The gradient of $\hat{e}_{i,k}$ at $\hat{\vartheta} = \hat{\vartheta}_{k-1}$ is in the opposite direction of $\varphi_{i,k}(s_{i,k} - F_0(\varphi_{i,k}^\top \hat{\vartheta}_{k-1}))$. By combining the gradients of $\hat{e}_{i,k}$ for all sub-systems, we update the estimate of ϑ as follows:

$$\begin{cases} \hat{\vartheta}_k = \hat{\vartheta}_{k-1} + \frac{1}{k} \left(\sum_{i=1}^{q} \beta_i \varphi_{i,k} \hat{s}_{i,k}\right), \\ \hat{s}_{i,k} = s_{i,k} - F_0(\varphi_{i,k}^\top \hat{\vartheta}_{k-1}), \end{cases} \tag{6}$$

where $\beta_1, \ldots, \beta_q$ are all positive coefficients.

Remark 4.2 *For the traditional systems with observable output sequence* $\{y_k\}$, *[24–26] adopt the step-size* $1/r_k$, *where* $r_k = 1 + \sum_{t=1}^{k} \|\phi_t\|^2$. *This is because when* $\|\phi_k\|$ *is small, there is little information on* θ *contained in* y_k. *At this point, we should reduce the impact of relevant data* ϕ_k *and* y_k *on the algorithm estimate. But in the quantized output case, even when* $\|\phi_k\|$ *is small,* z_k *contains the information on the noise variance* σ^2 *and the mean value* μ, *which is an important knowledge to identify* θ. *Therefore, we replace the step-size* $1/r_k$ *with* $1/k$ *in* (6).

Remark 4.3 *Since* $\sum_{k=1}^{\infty} \frac{1}{k} = \infty$ *and* $\sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$, $\frac{1}{k}$ *is a stochastic approximation step-size [33], and is commonly adopted in identification algorithms [5, 7, 18]. By* $\sum_{k=1}^{\infty} \frac{1}{k} = \infty$, *our estimate* $\hat{\vartheta}_k$ *has the ability to reach any point in the whole space. Therefore, the algorithm* (6) *can be used identify any* $\vartheta \in \mathbb{R}^{n+2}$. *Besides, the step-size should converge to* 0, *because by Cauchy criterion [31], the convergence of* $\hat{\vartheta}_k$ *necessarily implies that* $\hat{\vartheta}_k - \hat{\vartheta}_{k-1}$ *converges to* 0.

5 Main results

This section will prove the convergence properties of the decomposition-recombination algorithm (4)-(6). For convenience, denote $\tilde{\vartheta}_k = \hat{\vartheta}_k - \vartheta$.

5.1 Assumption

This subsection will give assumptions for the noises, quantizers, and the regressors.

Assumption 1 *The noise sequence* $\{d_k\}$ *is assumed to be i.i.d. Gaussian with positive variance.*

Assumption 2 *There exist at least two thresholds in the quantized observation* (2).

Remark 5.1 *By Theorem 3.1, Assumption 2 is necessary for the joint identifiability.*

Assumption 3 *There exists* $M > 0$ *such that for all* $k \geq 1$, $\|\phi_k\| \leq M$ *almost surely, and* $\mathbb{E}\|\phi_k\|^p \leq M^p$ *for any positive integer* p. *Besides,* ϕ_k *is independent of* d_t *for all* $t \geq k$.

Remark 5.2 *The boundedness assumption on* $\{\phi_k\}$ *is commonly required for the online identification algorithms of quantized systems [5, 6, 9, 13].*

Assumption 4 *The regressor sequence* $\{\phi_k\}$ *satisfies that there exist* $k_0 \in \mathbb{N}$ *and* $\delta > 0$ *such that for all* $k \geq k_0$,

$$\frac{1}{k}\sum_{t=1}^{k}\begin{bmatrix}\phi_t\phi_t^\top & \phi_t\\ \phi_t^\top & 1\end{bmatrix} \geq \delta\,(\ln k)^{-1/3}\, I_{n+1},\ \ a.s. \tag{7}$$

Remark 5.3 *Assumption 4 is weak among existing literature. Many existing works [9, 12, 16] consider the PE condition, i.e., for any* $\theta, \theta^\star \in \mathbb{R}^n$, $\lim\limits_{k\to\infty} \frac{1}{k}\sum_{t=1}^{k}(\phi_t^\top\theta - \phi_t^\top\theta^\star)^2 = 0$ *implies* $\theta = \theta^\star$ *[9]. The condition is rather weak for the open-loop systems. But in closed-loop systems, it is hard to guarantee PE conditions and meanwhile to enable* y_k *to track non-persistently exciting references [13, 34]. This limits the application of the identification algorithm on, for example, the adaptive tracking control problems. Therefore, it is important to identify the unknown parameters when* $\{\phi_k\}$ *is non-PE. The condition* (7) *is non-PE. This is because under* (7), $\frac{1}{k}\sum_{t=1}^{k}\phi_t\phi_t^\top$ *is allowed to converge to zero at the rate of* $O\left((\ln k)^{-1/3}\right)$. *In this case,* $\lim\limits_{k\to\infty} \frac{1}{k}\sum_{t=1}^{k}(\phi_t^\top\theta - \phi_t^\top\theta^\star)^2 = 0$ *holds for any* θ *and* $\theta^\star$.

Remark 5.4 *The non-PE condition* (7) *is inspired by the condition for the unquantized case in [26, 35]*

$$\limsup_{k\to\infty} \frac{\lambda_{\max}\left(\sum_{t=1}^{k}\phi_t\phi_t^\top\right)}{(\ln r_k)^{1/3}\,\lambda_{\min}\left(\sum_{t=1}^{k}\phi_t\phi_t^\top\right)} < \infty, \tag{8}$$

where r_k *is given in Remark 4.2, and* $\lambda_{\max}(\cdot)$ *and* $\lambda_{\min}(\cdot)$ *are the maximum and minimum eigenvalues of the corresponding matrix, respectively. This seems the weakest excitation condition for the convergence of SA algorithms when* y_k *can be accurately observed. However, when the mean value* μ *is unknown, the condition* (8) *is not sufficient to identify* θ. *Hence, we change the condition* (8) *accordingly, and give the non-PE condition* (7).

Remark 5.5 *Assumption 4 is easy to be satisfied for open-loop and closed-loop system case. Especially, for closed-loop systems, one can use the attenuating excitation technique [36] such that Assumption 4 holds.*

5.2 Convergence analysis

This section will analyze the strong consistency and the almost sure convergence rate of the decomposition-recombination algorithm (4)-(6).

The proof of the strong consistency consists of three steps.

Firstly, we give the almost sure boundedness of $\hat{\vartheta}_k$.

Lemma 5.1 *Under Assumptions 1-3, the estimate* $\hat{\vartheta}_k$ *given by Algorithm* (4)-(6) *is bounded almost surely.*

Proof. Define

$$W_k = \frac{\sum_{t=1}^{k}\sum_{i=1}^{q}\beta_i\varphi_{i,t}\left(s_{i,k} - F_0(\varphi_{i,k}^\top\vartheta)\right)}{k}. \tag{9}$$

Then, by the law of iterated logarithm [37], it holds that $W_k = O\left(\sqrt{\ln\ln k/k}\right)$ almost surely.

Consider the stochastic process $\psi_k = \hat{\vartheta}_k - W_k$. Then, by (6) and the continuity of $F_0(\cdot)$, one can get

$$\begin{aligned}\psi_k =& \psi_{k-1} + \frac{1}{k}\left(\sum_{i=1}^{q}\beta_i\varphi_{i,k}\left(F_0(\varphi_{i,k}^\top\vartheta) - F_0(\varphi_{i,k}^\top\hat{\vartheta}_{k-1})\right)\right)\\ &+ O\left(\frac{W_{k-1}}{k}\right)\\ =& \psi_{k-1} + \frac{1}{k}\left(\sum_{i=1}^{q}\beta_i\varphi_{i,k}\left(F_0(\varphi_{i,k}^\top\vartheta) - F_0(\varphi_{i,k}^\top\psi_{k-1})\right)\right)\\ &+ O\left(\sqrt{\frac{\ln\ln k}{k^3}}\right), \quad \text{a.s.} \qquad (10)\end{aligned}$$

By the monotonicity of $F_0(\cdot)$, it holds that

$$\left(\varphi_{i,k}^\top(\psi_{k-1} - \vartheta)\right)\left(F_0(\varphi_{i,k}^\top\vartheta) - F_0(\varphi_{i,k}^\top\psi_{k-1})\right) \le 0,$$

which together with (10) implies

$$\|\psi_k - \vartheta\|^2 \le \|\psi_{k-1} - \vartheta\|^2 + O\left(\sqrt{\frac{\ln\ln k}{k^3}}\right), \text{ a.s.}$$

Then, ψ_k is bounded almost surely, which together with $\|\hat{\vartheta}_k\| \le \|\psi_k\| + \|W_k\|$ implies the lemma. □

Remark 5.6 *Because the step-size $\frac{1}{k}$ satisfies $\sum_{k=1}^{\infty}\frac{1}{k} = \infty$, Algorithm (4)-(6) is not uniformly bounded. Therefore, it is difficult to analyze the almost sure boundedness of the algorithm. To overcome the difficulty, [11, 13, 38] introduce projections to restrict the search region of the algorithms in a known compact set. The projected algorithms are thereby uniformly bounded. In these works, the unknown system parameter is assumed to be located in the compact set for the convergence analysis. Such a priori knowledge is not always available. To remove the assumption, [18, 23] estimate the distribution tail of the algorithms. By using this technique, they prove that the probability of the estimates falling outside any neighborhood of the true value converging to 0. But, the technique requires periodic inputs or uniform PE inputs. In Lemma 5.1, a new technique is developed. An auxiliary stochastic process ψ_k is constructed to obtain the almost sure boundedness. This technique has wider applicability. Neither any a priori knowledge on the range of the unknown parameter nor any excitation conditions on inputs are required.*

Secondly, we estimate the matrix

$$\Phi(k,t) = \prod_{s=t+1}^{k}\left(I_{n+2} - \sum_{i=1}^{q}\frac{\beta_i\check{f}_{i,s}\varphi_{i,s}\varphi_{i,s}^\top}{s}\right),$$

where

$$\check{f}_{i,k} = \begin{cases} -\frac{F_0(\varphi_{i,k}^\top\vartheta) - F_0(\varphi_{i,k}^\top\hat{\vartheta}_{k-1})}{\varphi_{i,k}^\top\tilde{\vartheta}_{k-1}}, & \text{if } \varphi_{i,k}^\top\tilde{\vartheta}_{k-1} \ne 0;\\ f_0(\varphi_{i,k}^\top\vartheta), & \text{if } \varphi_{i,k}^\top\tilde{\vartheta}_{k-1} = 0.\end{cases}$$

Lemma 5.2 *For the quantized system (1)-(2) with Assumptions 1-3, suppose there exist $k_0 \in \mathbb{N}$, $\delta > 0$ and $\eta \in [0, \frac{1}{3}]$ such that for all $k \ge k_0$,*

$$\frac{1}{k}\sum_{t=1}^{k}\begin{bmatrix}\phi_t\phi_t^\top & \phi_t\\ \phi_t^\top & 1\end{bmatrix} \ge \delta(\ln k)^{-\eta} I_{n+1}, \ a.s. \qquad (11)$$

Then, there is a positive constant m such that

$$\|\Phi(k,k_0)\| = \begin{cases} O\left(\exp\left(-m(\ln k)^{1-3\eta}\right)\right), & \text{if } 0 \le \eta < \frac{1}{3};\\ O\left(\frac{1}{(\ln k)^m}\right), & \text{if } \eta = \frac{1}{3},\end{cases} \quad a.s.$$

Proof. The proof is inspired by Theorem 2.3.1 and Corollary 2.3.1 in [35].

The proof consists of four parts. Firstly, we show that there is a positive lower bound for $\check{f}_{i,k}$. Secondly, we give an increasing sequence $\{t_l\}$ such that

$$\sum_{t=t_{l-1}}^{t_l-1}\sum_{i=1}^{q}\frac{\beta_i\check{f}_{i,k}\varphi_{i,t}\varphi_{i,t}^\top}{t} > I_{n+2}, \text{ a.s.} \qquad (12)$$

Thirdly, we estimate the upper bound of the matrix $\sum_{t=t_{l-1}}^{t_l-1}\sum_{i=1}^{q}\frac{\beta_i\check{f}_{i,k}\varphi_{i,t}\varphi_{i,t}^\top}{t}$. Finally, we estimate $\|\Phi(t_l-1, t_{l-1}-1)\|$ and further $\|\Phi(k,k_0)\|$.

Part 1. By the definition of $\check{f}_{i,k}$ and the Lagrange's finite-increment theorem (Theorem 5.3.1 in [31]), there exists an x between $\varphi_{i,k}^\top\vartheta$ and $\varphi_{i,k}^\top\hat{\vartheta}_{k-1}$ such that $\check{f}_{i,k} = f_0(x)$, which further implies

$$\check{f}_{i,k} \ge \min\left\{f_0(\varphi_{i,k}^\top\vartheta), f_0(\varphi_{i,k}^\top\hat{\vartheta}_{k-1})\right\}.$$

By Lemma 5.1, $\hat{\vartheta}_{k-1}$ is bounded almost surely. And by Assumption 3, $\varphi_{i,k}$ is bounded almost surely. Therefore, there exists $\underline{f} > 0$ such that $\check{f}_{i,k} > \underline{f}$ for all i and k almost surely.

Part 2. Due to the concavity of $(\ln t)^{1-\eta}$, given any $\varepsilon \in (0,1)$, there exists a sufficiently large N such that $\varepsilon t > 1$ and $(\ln(t-1))^{1-\eta} \ge (\ln t)^{1-\eta} - 1$ for all $t \ge t_0$, $t_0 = \lfloor\exp(N^{\frac{1}{1-\eta}})\rfloor > k' \ge k_0 + 2$.

Let $t_l = \lfloor \exp((N+\alpha l)^{\frac{1}{1-\eta}}) \rfloor$, where

$$\alpha > 1 + \frac{1-\eta}{\delta(1-\varepsilon)}\left(\frac{1}{\min\limits_{i=1,\dots,q} \beta_i \underline{f}} + q\right).$$

Then, we have

$$\begin{aligned}
&(\ln t_l)^{1-\eta} - (\ln t_{l-1})^{1-\eta} \\
=&\left(\ln\lfloor \exp((N+\alpha l)^{\frac{1}{1-\eta}})\rfloor\right)^{1-\eta} \\
&-\left(\ln\lfloor \exp((N+(l-1)\alpha)^{\frac{1}{1-\eta}})\rfloor\right)^{1-\eta} \\
\geq&\left(\ln\left(\exp((N+\alpha l)^{\frac{1}{1-\eta}})-1\right)\right)^{1-\eta} \\
&-\left(\ln\left(\exp((N+(l-1)\alpha)^{\frac{1}{1-\eta}})\right)\right)^{1-\eta} \\
\geq&\alpha - 1,
\end{aligned} \tag{13}$$

and

$$\begin{aligned}
&\sum_{t=t_{l-1}}^{t_l-1}\sum_{i=1}^{q}\frac{\varphi_{i,t}\varphi_{i,t}^\top}{t} \\
=&\sum_{t=t_{l-1}}^{t_l-1}\sum_{i=1}^{q}\frac{1}{t}\left(\sum_{s=1}^{t}\varphi_{i,s}\varphi_{i,s}^\top - \sum_{s=1}^{t-1}\varphi_{i,s}\varphi_{i,s}^\top\right) \\
=&\sum_{t=t_{l-1}+1}^{t_l}\sum_{s=1}^{t-1}\sum_{i=1}^{q}\frac{\varphi_{i,s}\varphi_{i,s}^\top}{t-1} - \sum_{t=t_{l-1}}^{t_l-1}\sum_{s=1}^{t-1}\sum_{i=1}^{q}\frac{\varphi_{i,s}\varphi_{i,s}^\top}{t} \\
\geq&\sum_{t=t_{l-1}+1}^{t_l}\frac{1}{(t-1)t}\sum_{s=1}^{t-1}\sum_{i=1}^{q}\varphi_{i,s}\varphi_{i,s}^\top - qI_{n+2}.
\end{aligned} \tag{14}$$

By (11) and Lemma A.1 in Appendix A, for all $k \geq k_0$,

$$\frac{1}{k}\sum_{t=1}^{k}\sum_{i=1}^{q}\varphi_{i,t}\varphi_{i,t}^\top \geq \delta(\ln k)^{-\eta}I_{n+2}, \text{ a.s.}, \tag{15}$$

which together with (13) and $\varepsilon t > 1$ implies

$$\begin{aligned}
&\sum_{t=t_{l-1}}^{t_l-1}\sum_{i=1}^{q}\frac{\varphi_{i,t}\varphi_{i,t}^\top}{t} \\
\geq&\left(\sum_{t=t_{l-1}+1}^{t_l}\delta\frac{t-1}{t}\cdot\frac{(\ln(t-1))^{-\eta}}{t-1} - q\right)I_{n+2} \\
\geq&\left(\delta(1-\varepsilon)\sum_{t=t_{l-1}+1}^{t_l}\left(\frac{(\ln t)^{1-\eta}}{1-\eta} - \frac{(\ln(t-1))^{-\eta}}{1-\eta}\right) - q\right)I_{n+2} \\
=&\left(\delta(1-\varepsilon)\left(\frac{(\ln t_l)^{1-\eta}}{1-\eta} - \frac{(\ln t_{l-1})^{1-\eta}}{1-\eta}\right) - q\right)I_{n+2}.
\end{aligned}$$

Hence, by (13), we have

$$\begin{aligned}
\sum_{t=t_{l-1}}^{t_l-1}\sum_{i=1}^{q}\frac{\varphi_{i,t}\varphi_{i,t}^\top}{t} \geq& \left(\frac{\delta(1-\varepsilon)}{1-\eta}(\alpha-1) - q\right)I_{n+2} \\
>&\frac{1}{\min\limits_{i=1,\dots,q}\beta_i\underline{f}}I_{n+2}, \text{ a.s.},
\end{aligned}$$

which further implies (12).

Part 3. Note that $\check{f}_{i,k} \leq f(0)$ and $\|\varphi_{i,k}\|^2 \leq M^2 + \max\limits_{i=1,\dots,q} C_i^2$. Then, it holds that

$$\begin{aligned}
&\sum_{t=t_{l-1}}^{t_l-1}\sum_{i=1}^{q}\frac{\beta_i\check{f}_{i,k}\varphi_{i,t}\varphi_{i,t}^\top}{t} \\
<&\mathrm{tr}\left[\sum_{t=t_{l-1}}^{t_l-1}\sum_{i=1}^{q}\frac{\beta_i\check{f}_{i,k}\varphi_{i,t}\varphi_{i,t}^\top}{t}\right]I_{n+2} \\
<&\sum_{t=t_{l-1}}^{t_l-1}\frac{B}{t}I_{n+2} < B\left(\ln(t_l-1) - \ln(t_{l-1}+1) + 2\right)I_{n+2} \\
=&B\Big(\ln(\lfloor\exp((N+\alpha l)^{\frac{1}{1-\eta}})\rfloor - 1) \\
&\quad - \ln(\lfloor\exp((N+(l-1)\alpha)^{\frac{1}{1-\eta}})\rfloor + 1) + 2\Big)I_{n+2} \\
<&B\left(2 + (N+\alpha l)^{\frac{1}{1-\eta}} - (N+(l-1)\alpha)^{\frac{1}{1-\eta}}\right)I_{n+2} \\
<&\frac{2\alpha}{1-\eta}B\left(N+\alpha l\right)^{\frac{\eta}{1-\eta}}I_{n+2}, \text{ a.s.},
\end{aligned} \tag{16}$$

where $B = \max\limits_{i=1,\dots,q}\beta_i f_0(0)\left(M^2 + \max\limits_{i=1,\dots,q}C_i^2\right)$.

Part 4. By (12), (16), and Lemma 2.3.1 in [35], we have

$$\begin{aligned}
&\|\Phi(t_l-1, t_{l-1}-1)\| \\
\leq&\left(1 - \frac{1}{2 + \frac{8\alpha^2}{(1-\eta)^2}(n+1)^2B^2\left(N+\alpha l\right)^{\frac{2\eta}{1-\eta}}}\right)^{\frac{1}{2}}, \text{ a.s.},
\end{aligned}$$

which together with Lemma A.2 in Appendix A implies that there exists an $m > 0$ such that

$$\begin{aligned}
&\|\Phi(k, k_0)\| \\
\leq&\|\Phi(t_0-1, k_0)\| \cdot \prod_{l=1}^{\lfloor\frac{(\ln k)^{1-\eta}-N}{\alpha}\rfloor}\|\Phi(t_l-1, t_{l-1}-1)\| \\
\leq&\prod_{l=1}^{\lfloor\frac{(\ln k)^{1-\eta}-N}{\alpha}\rfloor}\left(1 - \frac{1}{2 + \frac{8\alpha^2}{(1-\eta)^2}(n+1)^2B^2\left(N+\alpha l\right)^{\frac{2\eta}{1-\eta}}}\right)^{\frac{1}{2}} \\
=&\begin{cases} O\left(\exp\left(-m(\ln k)^{1-3\eta}\right)\right), & \text{if } 0 \leq \eta < \frac{1}{3}; \\ O\left(\frac{1}{(\ln k)^m}\right), & \text{if } \eta = \frac{1}{3}, \end{cases} \quad \text{a.s.}
\end{aligned}$$

The proof is completed. □

Remark 5.7 *To achieve the convergence of* $\Phi(k,k_0)$, *[18] requires the uniform PE condition, i.e., there exist* $k_0, \kappa \in \mathbb{N}$ *and* $\delta > 0$ *such that for all* $k \geq k_0$,

$$\frac{1}{\kappa}\sum_{t=k-\kappa+1}^{k} \phi_t\phi_t^\top \geq \delta. \tag{17}$$

The idea is to approximately treat the matrix $\Phi(k, k-\kappa)$ *as a positive definite matrix whose eigenvalues are all less that* $1-\beta^*/k$ *for some constant* $\beta^* > 0$. *However, the technique relies on the condition* (17), *and is thereby hard be to applied in the non-PE condition* (7) *case. The difficulty is solved by adopting the exponential segmentation* $t_l = \lfloor \exp((N+\alpha l)^{\frac{1}{1-\eta}})\rfloor$, *which is inspired by [35].*

Thirdly, we give the strong consistency of the algorithm using Lemmas 5.1 and 5.2.

Theorem 5.1 *For the quantized system* (1)-(2), *assume that Assumptions 1-4 hold. Then, the estimate* $\hat{\vartheta}_k$ *given by the decomposition-recombination algorithm* (4)-(6) *converges to the true parameters* ϑ *almost surely.*

Proof. Since $\tilde{\vartheta}_k = \hat{\vartheta}_k - \vartheta$, we have

$$\begin{aligned}
\tilde{\vartheta}_k =& \tilde{\vartheta}_{k-1} + \frac{1}{k}\left(\sum_{i=1}^{q}\beta_i\varphi_{i,k}\left(s_{i,k} - F_0(\varphi_{i,k}^\top\hat{\vartheta}_{k-1})\right)\right)\\
=& \left(I_{n+2} - \sum_{i=1}^{q}\frac{\beta_i\check{f}_{i,k}\varphi_{i,k}\varphi_{i,k}^\top}{k}\right)\tilde{\vartheta}_{k-1}\\
&+ \sum_{i=1}^{q}\frac{\beta_i\varphi_{i,k}}{k}\left(s_{i,k} - F_0(\varphi_{i,k}^\top\vartheta)\right)\\
=&\Phi(k,k_0)\tilde{\vartheta}_{k_0}\\
&+ \sum_{t=k_0+1}^{k}\Phi(k,t)\sum_{i=1}^{q}\frac{\beta_i\varphi_{i,t}}{t}\left(s_{i,t} - F_0(\varphi_{i,t}^\top\vartheta)\right). \quad (18)
\end{aligned}$$

By Lemma 4 in [23], it holds that

$$\begin{aligned}
&\|\Phi(k,t)\| \leq \|\Phi(k,k_0)\|\left\|\Phi(t,k_0)^{-1}\right\|\\
\leq& \|\Phi(k,k_0)\| \cdot \prod_{s=k_0+1}^{t}\left\|\left(I_{n+2} - \sum_{i=1}^{q}\frac{\beta_i\check{f}_{i,s}\varphi_{i,s}\varphi_{i,s}^\top}{s}\right)^{-1}\right\|\\
\leq& \|\Phi(k,k_0)\| \cdot \prod_{s=k_0+1}^{t}\left(1-\frac{qB}{s}\right)^{-1} = O\left(\|\Phi(k,k_0)\|\, t^{qB}\right),\\
&\text{a.s.} \quad (19)
\end{aligned}$$

where B is defined in (16).

Denote

$$S_k = \sum_{t=k_0+1}^{k}\sum_{i=1}^{q}\frac{\beta_i\varphi_{i,t}}{t}\left(s_{i,t} - F_0(\varphi_{i,t}^\top\vartheta)\right).$$

By Lemma A.3 in Appendix A, there exists a random variable S such that $\lim_{k\to\infty} S_k = S$ and $\lim_{k\to\infty} S_k^- = O\left(\sqrt{\ln\ln k/k}\right)$ almost surely, where $S_k^- = S_k - S$. Then, by (19), we have

$$\begin{aligned}
&\left\|\sum_{t=k_0+1}^{k}\Phi(k,t)\sum_{i=1}^{q}\frac{\beta_i\varphi_{i,t}}{t}\left(s_{i,t} - F_0(\varphi_{i,t}^\top\vartheta)\right)\right\|\\
=&\left\|\sum_{t=k_0+1}^{k}\Phi(k,t)\left(S_t^- - S_{t-1}^-\right)\right\|\\
=&\left\|S_k^- - \Phi(k,k_0)S_{k_0}^- + \sum_{t=k_0}^{k-1}(\Phi(k,t) - \Phi(k,t+1))S_t^-\right\|,
\end{aligned}$$

which implies

$$\begin{aligned}
&\left\|\sum_{t=k_0+1}^{k}\Phi(k,t)\sum_{i=1}^{q}\frac{\beta_i\varphi_{i,t}}{t}\left(s_{i,t} - F_0(\varphi_{i,t}^\top\vartheta)\right)\right\|\\
=&O\left(\sqrt{\frac{\ln\ln k}{k}} + \|\Phi(k,k_0)\|\right)\\
&+\left\|\sum_{t=k_0}^{k-1}\Phi(k,t+1)\left(\sum_{i=1}^{q}\frac{\beta_i\check{f}_{i,t+1}\varphi_{i,t}\varphi_{i,t}^\top}{t+1}\right)S_{t+1}^-\right\|\\
=&O\left(\sqrt{\frac{\ln\ln k}{k}} + \|\Phi(k,k_0)\|\right)\\
&+O\left(\sum_{t=k_0}^{k-1}\|\Phi(k,t+1)\|\sqrt{\frac{\ln\ln t}{t^3}}\right), \text{ a.s.}
\end{aligned} \tag{20}$$

Define

$$\tau(k) = \max\left[t : \sqrt{\frac{\ln\ln t}{t^{1+2qB}}} > \|\Phi(k,k_0)\|\right].$$

Then, by Lemma 5.2, we have $\lim_{k\to\infty}\|\Phi(k,k_0)\| = 0$ almost surely, which implies $\lim\lim_{k\to\infty}\tau(k) = \infty$ almost surely. Besides, one can get

$$\begin{aligned}
\|\Phi(k,k_0)\| \geq& \prod_{s=k_0+1}^{k}\left(1 - \sum_{i=1}^{q}\frac{\beta_i\check{f}_{i,s}\|\varphi_{i,s}\|^2}{s}\right)\\
\geq& \prod_{s=k_0+1}^{k}\left(1-\frac{qB}{s}\right), \text{ a.s.}
\end{aligned}$$

which together with Lemma 4 in [23] implies that $\Phi(k,k_0)$ is of the same order as $\frac{1}{k^{qB}}$. Therefore, $\sqrt{\frac{\ln\ln k}{k^{1+2qB}}} < \|\Phi(k,k_0)\|$ for sufficiently large k. Then, by the definition of $\tau(k)$, we have $\tau(k) < k$ for sufficiently large k.

By (19), we have

$$
\begin{aligned}
&\sum_{t=k_0}^{k-1}\|\Phi(k,t+1)\|\sqrt{\frac{\ln\ln t}{t^3}}\\
\leq&\sum_{t=k_0}^{\tau(k)-1}\|\Phi(k,\tau(k))\|\sqrt{\frac{\ln\ln t}{t^3}}+\sum_{t=\tau(k)}^{k-1}\sqrt{\frac{\ln\ln t}{t^3}}\\
=&O\left(\|\Phi(k,\tau(k))\|+\sqrt{\frac{\ln\ln\tau(k)}{\tau(k)}}\right)\\
=&O\left(\|\Phi(k,k_0)\|\,\tau(k)^{qB}+\sqrt{\frac{\ln\ln\tau(k)}{\tau(k)}}\right),\ \text{a.s.} \qquad (21)
\end{aligned}
$$

By the definition of $\tau(k)$, for sufficiently large k,

$$
\begin{aligned}
&\tau(k)^{-\frac{1}{4}-qB}>\sqrt{\frac{\ln\ln\tau(k)}{\tau(k)^{1+2qB}}}\\
&>\|\Phi(k,k_0)\|\geq\sqrt{\frac{\ln\ln(\tau(k)+1)}{(\tau(k)+1)^{1+2qB}}},\ \text{a.s.},
\end{aligned}
$$

which together with (21) implies for sufficiently large k,

$$
\sum_{t=k_0}^{k-1}\|\Phi(k,t+1)\|\sqrt{\frac{\ln\ln t}{t^3}}=O\left(\|\Phi(k,k_0)\|^{\frac{1}{4qB+1}}\right),\ \text{a.s.} \qquad (22)
$$

By (18), (20) and (22), one can get

$$
\|\tilde{\vartheta}_k\|=O\left(\sqrt{\frac{\ln\ln k}{k}}+\|\Phi(k,k_0)\|^{\min\left\{1,\frac{1}{4qB+1}\right\}}\right),\ \text{a.s.} \qquad (23)
$$

Then, the theorem can be proved by Lemma 5.2. ☐

The following theorem gives the almost sure convergence rates of the decomposition-recombination algorithm (4)-(6) under a class of excitation conditions.

Theorem 5.2 *For the quantized system* (1)-(2)*, assume that Assumptions 1-3 and the excitation condition* (11) *hold. Then, there exists a positive constant m such that*

$$
\|\tilde{\vartheta}_k\|=\begin{cases}O\left(\exp\left(-m(\ln k)^{1-3\eta}\right)\right), & \text{if } 0\leq\eta<\frac{1}{3};\\ O\left(\frac{1}{(\ln k)^m}\right), & \text{if } \eta=\frac{1}{3},\end{cases}\quad \text{a.s.}
$$

Proof. By Theorem 5.1, we have

$$
\lim_{k\to\infty}\hat{f}_{i,k}=f_0(\varphi_{i,k}^{\top}\vartheta)\geq f_0\left(\sqrt{M^2+\max_{i=1,\ldots,q}C_i^2}\,\|\vartheta\|\right),\ \text{a.s.}
$$

Therefore, given any $\varepsilon>0$, there exists k' such that for all $k\geq k'$, $\check{f}_{i,k}>f_0\left(\sqrt{M^2+\max_{i=1,\ldots,q}C_i^2}\,\|\vartheta\|\right)-\varepsilon$ almost surely. Then, the theorem can be proved by Lemma 5.2 and (23). ☐

Remark 5.8 *Theorem 5.2 demonstrates that the convergence rate of the algorithm* (4)-(6) *is related to the excitation condition on the inputs. When $\eta=\frac{1}{3}$, the condition* (11) *is Assumption 4. The condition is weakest among existing SA-based identification algorithms [11, 14, 18]. Under the condition, the algorithm* (4)-(6) *can achieve convergence, and a logarithmic convergence rate can be obtained. This is consistent with the unquantized case [26]. When $\eta=0$, the condition* (11) *is PE. Under the PE condition, the algorithm* (4)-(6) *converges at a polynomial rate almost surely. Besides, under uniform PE condition* (17)*, similar to [18], the algorithm* (4)-(6) *can achieve an almost sure convergence rate of $O(\sqrt{\ln\ln k/k})$, which is the best among existing identification algorithms for quantized systems [11, 13, 14].*

Now, we analyzes the L^p convergence for any positive integer p of the decomposition-recombination algorithm (4)-(6). The proof consists of three steps.

Firstly, we proves the L^p boundedness of the algorithm.

Lemma 5.3 *Under Assumptions 1-3, for the algorithm* (4)-(6)*, $\mathbb{E}\|\tilde{\vartheta}_k\|_p^p$ is bounded for any positive integer p.*

Proof. Firstly, we prove the boundedness of $\mathbb{E}\|\tilde{\vartheta}_k\|_2^{2r}$ for any non-negative integer r. When $r=0$, we have $\|\tilde{\vartheta}_k\|_2^0=1$. When $\mathbb{E}\|\tilde{\vartheta}_k\|_2^{2r-2}$ is bounded, by Lyapunov inequality [39], $\mathbb{E}\|\tilde{\vartheta}_k\|_2^s$ is bounded for all $s\leq 2r-2$. Then, it holds that

$$
\begin{aligned}
&\mathbb{E}\|\tilde{\vartheta}_k\|_2^{2r}\\
=&\mathbb{E}\left(\|\tilde{\vartheta}_{k-1}\|_2^2+\frac{1}{k}\sum_{i=1}^{q}\varphi_{i,k}^{\top}\tilde{\vartheta}_{k-1}\hat{s}_{i,k}+O\left(\frac{1}{k^2}\right)\right)^r\\
=&\mathbb{E}\|\tilde{\vartheta}_{k-1}\|_2^{2r}+O\left(\sum_{s=0}^{2r-2}\mathbb{E}\frac{\|\tilde{\vartheta}_{k-1}\|_2^s}{k^{2r-s}}\right)\\
&+\frac{r}{k}\sum_{i=1}^{q}\mathbb{E}\left[\|\tilde{\vartheta}_{k-1}\|_2^{2r-2}\varphi_{i,k}^{\top}\tilde{\vartheta}_{k-1}\left(F_0(\varphi_{i,k}^{\top}\vartheta)-F_0(\varphi_{i,k}^{\top}\hat{\vartheta})\right)\right]\\
\leq&\mathbb{E}\|\tilde{\vartheta}_{k-1}\|_2^{2r}+O\left(\frac{1}{k^2}\right),
\end{aligned}
$$

which implies the boundedness of $\mathbb{E}\|\tilde{\vartheta}_k\|_2^{2r}$.

Denote $\tilde{\vartheta}_{k,j}$ as the j-th component of $\tilde{\vartheta}_k$. Then, we have

$$
\mathbb{E}\|\tilde{\vartheta}_k\|_{2r}^{2r}=\mathbb{E}\left(\sum_{j=1}^{n+1}\tilde{\vartheta}_{k,j}^{2r}\right)\leq\mathbb{E}\left(\sum_{j=1}^{n+1}\tilde{\vartheta}_{k,j}^{2}\right)^r=\mathbb{E}\|\tilde{\vartheta}_k\|_2^{2r}.
$$

Therefore, $\mathbb{E}\|\tilde{\vartheta}_k\|_p^p$ is bounded for all $p = 2r$, which together with Lyapunov inequality [39] implies the boundedness of $\mathbb{E}\|\tilde{\vartheta}_k\|_p^p$ for all integer p. □

Secondly, we proves the uniform integrability for $\|\tilde{\vartheta}_k\|_p^p$.

Lemma 5.4 *If a random sequence $\{x_k\}$ satisfies $\sup_k \mathbb{E}\|x_k\|_{2p}^{2p} < \infty$, then $\|x_k\|_p^p$ is uniformly integrable.*

Proof. By C_r-inequality [35], we have

$$\mathbb{E}\|x_k\|_p^{2p} = \mathbb{E}\left(\sum_{i=1}^n x_{k,i}^p\right)^2 \le n\mathbb{E}\sum_{i=1}^n x_{k,i}^{2p} = n\mathbb{E}\|x_k\|_{2p}^{2p},$$

where $x_{k,i}$ is the i-th component of x_k. Then, the lemma can be proved by de La Vallée Poussin criterion [40]. □

Corollary 5.1 *Under the condition of Lemma 5.3, $\|\tilde{\vartheta}_k\|_p^p$ is uniformly integrable for any positive integer p.*

Proof. The corollary can be obtained by Lemmas 5.3 and 5.4. □

Lastly, we proves the L^p convergence of the algorithm using Corollary 5.1.

Theorem 5.3 *Under the condition of Theorem 5.1, the estimate $\hat{\vartheta}_k$ given by the decomposition-recombination algorithm* (4)-(6) *converges to the true parameters ϑ in the L^p sense for any positive integer p.*

Proof. By Theorem 5.1, it holds that $\|\tilde{\vartheta}_k\|_p^p$ converges to 0 almost surely. Then, the theorem can be obtained by Corollary 5.1 and Theorem 2.6.4 in [39]. □

Remark 5.9 *Theorems 5.1 and 5.3 analyze the strong consistency and L^p convergence, respectively. Strong consistency describes the convergence property of the estimates in one random experiment with probability 1. In some practical engineering scenarios, such as the adaptive control problem of drag-free satellites [41], repeated experiments are difficult to be achieved. In this case, the strong consistency of the estimates is important to be investigated. On the other hand, L^p convergence focuses on the average property of L^p estimation error. When $p = 2$, L^p convergence is the mean square convergence, which further implies the convergence of the estimate's variance. In some practical identification problem for set-valued systems, such as the genome-wide association studies [1], the estimate's variance is an important metric for the effectiveness of identification algorithms.*

5.3 Known mean value case

When the mean value μ of the noises is known, the joint identification algorithm is similar to Algorithm (4)-(6) for the unknown mean value case. The main difference is that in (4),

$$\vartheta = \begin{bmatrix} \frac{\theta}{\sigma} & \frac{1}{\sigma} \end{bmatrix}^\top, \ \varphi_{i,k} = \begin{bmatrix} \phi_k^\top & \mu - C_i \end{bmatrix}^\top.$$

It is worth mentioning that the modified algorithm can converges to the true value under weaker conditions. For the quantizer, Assumption 2 can be relaxed as follows.

Assumption 2' *For the quantized observation* (2), *there exists at least one threshold C_i that is not equal to μ.*

For the excitation condition, Assumption 4 can be also relaxed.

Assumption 4' *When $q = 1$, the regressor sequence $\{\phi_k\}$ satisfies the non-PE condition* (7). *When $q \ge 2$, $\{\phi_k\}$ satisfies that there exist $k_0 \in \mathbb{N}$ and $\delta > 0$ such that for all $k \ge k_0$,*

$$\frac{1}{k}\sum_{t=1}^k \phi_t\phi_t^\top \ge \delta (\ln k)^{-1/3} I_n, \ a.s. \tag{24}$$

Similar to Theorems 5.1-5.3 and by Corollary A.1 in Appendix A, the modified algorithm for the known mean value case can also achieve almost sure and L^p convergence under Assumptions 1, 2', 3 and 4'. A similar almost sure convergence rate can also be obtained.

6 Simulation

This section will give a numerical example for the main results. Consider the quantized system (1)-(2), where $q = 2$, $C_1 = 0$, and $C_2 = \frac{3}{2}$. The unknown system parameter $\theta = [2 \ -2]^\top$. For the system noise d_k, both the variance $\sigma^2 = 1$ and mean value $\mu = 1$ are unknown. The regressor $\phi_k = [u_k \ u_{k-1}]^\top$, and the input u_k is set to be

$$\begin{cases} u_{3l-2} = \frac{1}{2}\left(1 + \frac{1}{\ln^{1/6}(3l-1)}\right); \\ u_{3l-1} = \frac{1}{2}\left(1 + \frac{1}{2\ln^{1/6}(3l)}\right); \\ u_{3l} = \frac{1}{2}\left(1 - \frac{1}{\ln^{1/6}(3l+1)}\right), \end{cases}$$

for all positive integer l. We can prove that the regressor ϕ_k satisfies Assumption 4'. Then, we apply the decomposition-recombination algorithm (4)-(6) to identify the unknown system parameter θ, the variance σ^2 and the mean value μ. The initial values $\hat{\vartheta}_0 = [0 \ 0 \ 0 \ 0]^\top$, and $\beta_1 = \beta_2 = 5$. For each k, if $\hat{\vartheta}_k^{(3)} > 0$, the estimates of θ, σ and μ are $\left[\hat{\vartheta}_k^{(1)}/\hat{\vartheta}_k^{(3)} \ \hat{\vartheta}_k^{(2)}/\hat{\vartheta}_k^{(3)}\right]^\top$, $1/\hat{\vartheta}_k^{(3)}$, and

$\hat{\vartheta}_k^{(4)}/\hat{\vartheta}_k^{(3)}$ respectively, where $\hat{\theta}_k^{(i)}$ is the i-th component of $\hat{\theta}_k$. Otherwise, the estimates are set as $[0\ \ 0]^\top$, 0 and 0.

The trajectories of the estimates of θ, σ and μ are shown in Figures 1-3. All of the estimates converge to the true values.

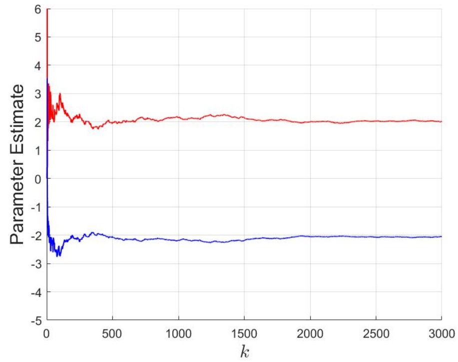

Fig. 1. Estimate of θ

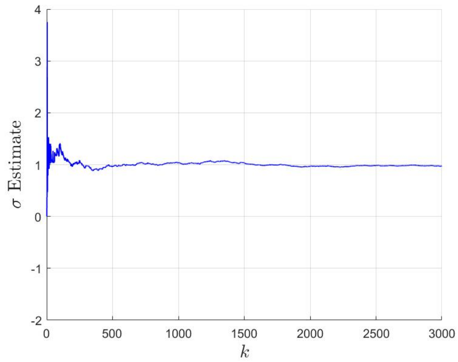

Fig. 2. Estimate of σ

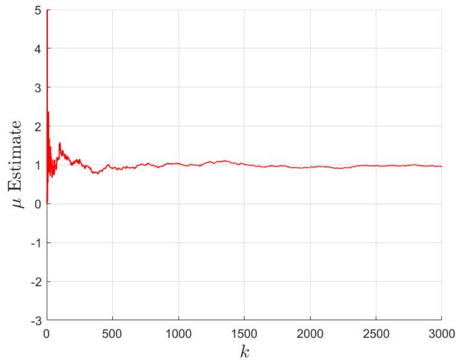

Fig. 3. Estimate of μ

7 Conclusion

This paper investigates the identification problem of quantized systems with a non-PE condition and Gaussian noises with unknown noise variance and mean value. The equivalent condition of the joint identifiability for unknown system parameter and noise parameters is given. Then, a decomposition-recombination identification algorithm is designed for the joint identification problem. It is shown that the estimates of the algorithm converge to the true values in the almost sure and L^p sense. The almost sure convergence rates of the algorithm are also obtained.

For the future research, the results of the paper lay a foundation for the consensus protocol design under quantized communications [16, 23]. Besides, the main results can be extended to the ARMA system case by applying the technique of [38].

Appendix A Lemmas and corollary

This appendix will give lemmas to prove the convergence and convergence rates of the algorithm.

Lemma A.1 *Under Assumption 2, for any $\eta \geq 0$, the following two propositions are equivalent.*

a) There exist $k_0 \in \mathbb{N}$ and $\delta > 0$ such that for all $k \geq k_0$,

$$\frac{1}{k}\sum_{t=1}^{k}\sum_{i=1}^{q}\begin{bmatrix}\phi_t\phi_t^\top & \phi_t\\ \phi_t^\top & 1\end{bmatrix} \geq \delta(\ln k)^{-\eta} I_{n+1},\ a.s. \quad \text{(A.1)}$$

b) There exist $k_0 \in \mathbb{N}$ and $\delta' > 0$ such that for all $k \geq k_0$,

$$\frac{1}{k}\sum_{t=1}^{k}\sum_{i=1}^{q}\begin{bmatrix}\phi_t\phi_t^\top & -C_i\phi_t & \phi_t\\ -C_i\phi_t^\top & C_i^2 & -C_i\\ \phi_t^\top & -C_i & 1\end{bmatrix} \geq \delta'(\ln k)^{-\eta} I_{n+2},\ a.s. \quad \text{(A.2)}$$

Proof. If (A.2) holds, then we have

$$\begin{aligned}
&q\sum_{t=1}^{k}\begin{bmatrix}\phi_t\phi_t^\top & \phi_t & 0\\ \phi_t^\top & 1 & 0\\ 0 & 0 & 0\end{bmatrix}\\
=&\begin{bmatrix}I_n & 0 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}\left(\sum_{t=1}^{k}\sum_{i=1}^{q}\begin{bmatrix}\phi_t\phi_t^\top & -C_i\phi_t & \phi_t\\ -C_i\phi_t^\top & C_i^2 & -C_i\\ \phi_t^\top & -C_i & 1\end{bmatrix}\right)\begin{bmatrix}I_n & 0 & 0\\ 0 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\\
\geq&\delta' k(\ln k)^{-\eta}\begin{bmatrix}I_{n+1} & \\ & 0\end{bmatrix},
\end{aligned}$$

which implies

$$\frac{1}{k}\sum_{t=1}^{k}\begin{bmatrix}\phi_t\phi_t^\top & \phi_t\\ \phi_t^\top & 1\end{bmatrix} \geq \frac{\delta}{q}(\ln k)^{-\eta} I_n,\ \text{a.s.}$$

Therefore, (A.1) holds for $\delta \le \delta'/q$.

If (A.1) holds, then by denoting $\rho = \frac{(C_2-C_1)^2}{C_1^2+C_2^2+2}$, we have $(2-\rho)(C_1^2+C_2^2-\rho) = \rho^2 + (C_1+C_2)^2 > (C_1+C_2)^2$, which implies

$$\begin{aligned}
&\begin{bmatrix} 2\phi_t\phi_t^\top & -(C_1+C_2)\phi_t & 2\phi_t \\ -(C_1+C_2)\phi_t^\top & C_1^2+C_2^2 & -C_1-C_2 \\ \phi_t^\top & -C_1-C_2 & 2 \end{bmatrix} \\
\ge&\rho \begin{bmatrix} \phi_t\phi_t^\top & 0 & \phi_t \\ 0 & 1 & 0 \\ \phi_t^\top & 0 & 1 \end{bmatrix} + H_tH_t^\top,
\end{aligned}$$

where $H_t = \left[\sqrt{2-\rho}\phi_t^\top \;\; -\frac{C_1+C_2}{\sqrt{2-\rho}} \;\; \sqrt{2-\rho}\right]^\top$. Hence, by Assumption 2, one can get

$$\begin{aligned}
&\sum_{i=1}^{q} \begin{bmatrix} \phi_t\phi_t^\top & -C_i\phi_t & \phi_t \\ -C_i\phi_t^\top & C_i^2 & -C_i \\ \phi_t^\top & -C_i & 1 \end{bmatrix} \\
\ge& \begin{bmatrix} 2\phi_t\phi_t^\top & -(C_1+C_2)\phi_t & 2\phi_t \\ -(C_1+C_2)\phi_t^\top & C_1^2+C_2^2 & -C_1-C_2 \\ \phi_t^\top & -C_1-C_2 & 2 \end{bmatrix} \\
\ge&\rho \begin{bmatrix} \phi_t\phi_t^\top & 0 & \phi_t \\ 0 & 1 & 0 \\ \phi_t^\top & 0 & 1 \end{bmatrix},
\end{aligned}$$

which implies for sufficiently large k,

$$\begin{aligned}
&\frac{1}{k}\sum_{t=1}^{k}\sum_{i=1}^{q} \begin{bmatrix} \phi_t\phi_t^\top & -C_i\phi_t & \phi_t \\ -C_i\phi_t^\top & C_i^2 & -C_i \\ \phi_t^\top & -C_i & 1 \end{bmatrix} \\
\ge&\frac{\rho}{k}\sum_{t=1}^{k} \begin{bmatrix} I_n & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \phi_t\phi_t^\top & \phi_t & 0 \\ \phi_t^\top & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I_n & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \\
\ge& \begin{bmatrix} I_n & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \sum_{t=1}^{k}\rho\delta\,(\ln k)^{-\eta} I_{n+1} & \\ & \rho \end{bmatrix} \begin{bmatrix} I_n & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \\
\ge&\rho\delta\,(\ln k)^{-\eta} I_{n+2}.
\end{aligned}$$

Therefore, (A.2) holds for $\delta' \le \rho\delta$. □

Corollary A.1 *Under Assumption 2', for any $\eta \ge 0$, the following two propositions are equivalent.*

a) When $q = 1$, the condition (A.1) holds. When $q \ge 2$, there exist $k_0 \in \mathbb{N}$ and $\delta > 0$ such that for all $k \ge k_0$,

$$\frac{1}{k}\sum_{t=1}^{k}\sum_{i=1}^{q} \phi_t\phi_t^\top \ge \delta(\ln k)^{-\eta} I_{n+1},\ a.s.$$

b) There exist $k_0 \in \mathbb{N}$ and $\delta' > 0$ such that for all $k \ge k_0$,

$$\frac{1}{k}\sum_{t=1}^{k}\sum_{i=1}^{q} \begin{bmatrix} \phi_t\phi_t^\top & \breve{C}_i\phi_t \\ \breve{C}_i\phi_t^\top & \breve{C}_i^2 \end{bmatrix} \ge \delta'(\ln k)^{-1/3} I_{n+2},\ \ a.s.,$$

where $\breve{C}_i = \mu - C_i$.

The proof of Corollary A.1 is similar to that of Lemma A.1, and thereby omitted here.

Lemma A.2 *For $0 < h \le 1$, $J, \alpha > 0$, and $H, N \ge 0$, we have*

$$\begin{aligned}
&\prod_{l=t}^{k-1}\left(1 - \frac{1}{H + J(\alpha l + N)^h}\right) \\
\le&\begin{cases} \left(\frac{H+J(\alpha t+N)}{H+J(\alpha k+N)}\right)^{\frac{1}{J\alpha}}, & h = 1; \\ \dfrac{\exp\left(\frac{\alpha t+N}{\alpha H+\alpha(1-h)J(\alpha t+N)^h}\right)}{\exp\left(\frac{H+J(\alpha t+N)^h}{\alpha H+\alpha(1-h)J(\alpha t+N)^h}\cdot\frac{\alpha k+N}{H+J(\alpha k+N)^h}\right)}, & h \in (0,1). \end{cases}
\end{aligned}$$

Proof. Note that

$$\begin{aligned}
&\prod_{l=t}^{k-1}\left(1 - \frac{1}{H + J(\alpha l + N)^h}\right) \\
=&\exp\left(\sum_{l=t}^{k-1}\ln\left(1 - \frac{1}{H + J(\alpha l + N)^h}\right)\right) \\
\le&\exp\left(-\sum_{l=t}^{k-1}\frac{1}{H + J(\alpha l + N)^h}\right). \quad \text{(A.3)}
\end{aligned}$$

Then, when $h = 1$, it holds that

$$\begin{aligned}
&\sum_{l=t}^{k-1}\frac{1}{H + J(\alpha l + N)} \ge \int_t^k \frac{1}{H + J(\alpha x + N)}\mathrm{d}x \\
=&\frac{1}{J\alpha}\left(\ln\left(H + J(\alpha k + N)\right) - \ln\left(H + J(\alpha t + N)\right)\right),
\end{aligned}$$

which together with (A.3) implies the lemma.

When $h \in (0,1)$, it holds that

$$\sum_{l=t}^{k-1}\frac{1}{H + J(\alpha l + N)^h} \ge \int_t^k \frac{1}{H + J(\alpha x + N)^h}\mathrm{d}x. \quad \text{(A.4)}$$

Note that

$$\begin{aligned}&\frac{\alpha H+\alpha(1-h)J(\alpha t+N)^h}{H+J(\alpha t+N)^h}\int_t^k\frac{1}{H+J(\alpha x+N)^h}\mathrm{d}x\\ \geq&\int_t^k\frac{\alpha H+\alpha(1-h)J(\alpha x+N)^h}{(H+J(\alpha x+N)^h)^2}\mathrm{d}x\\ =&\frac{\alpha k+N}{H+J(\alpha k+N)^h}-\frac{\alpha t+N}{H+J(\alpha t+N)^h}.\end{aligned}$$

Then, the lemma can be proved by (A.3) and (A.4). □

Remark A.1 *The proof of Lemma A.2 is inspired by that of Lemma 3.2 in [32].*

Lemma A.3 *Assume that $\{X_k\}$ is a martingale difference sequence and bounded almost surely. Then, $\sum_{t=1}^{k}\frac{X_t}{t}$ converges almost surely, and $\sum_{t=k}^{\infty}\frac{X_t}{t}=O\left(\sqrt{\ln\ln k/k}\right)$ almost surely.*

Proof. The lemma can be proved by

$$\sum_{t=k_1}^{k_2}\frac{X_t}{t}=\frac{\sum_{t=1}^{k_2}X_t}{k_2}-\frac{\sum_{t=1}^{k_1-1}X_t}{k_1-1}+\sum_{t=k_1}^{k_2}\frac{\sum_{l=1}^{t-1}X_l}{t(t-1)}$$

for any integers $k_2\geq k_1\geq 2$, and the law of iterated logarithm [37]. □

References

[1] G. Kang, W. Bi, H. Zhang, S. Pounds, C. Cheng, S. Shete, F. Zou, Y.L. Zhao, J.F. Zhang, W. Yue, A robust and powerful set-valued approach to rare variant association analyses of secondary traits in case-control sequencing studies, Genetics 205 (3) (2017) 1049–1062.

[2] J. Li, L. Wu, W. Lü, T. Wang, Y. Kang, D. Feng, H. Zhou, Lithology classification based on set-valued identification method, J. Syst. Sci. Complex. 35 (2022) 1637–1652.

[3] G. Gagliardi, D. Mari, F. Tedesco, A. Casavola, An air-to-fuel ratio estimation strategy for turbocharged spark-ignition engines based on sparse binary HEGO sensor measures and hybrid linear observers, Control Eng. Pract. 107 (2021) 104694.

[4] L.Y. Wang, J.F. Zhang, G. Yin, System identification using binary sensors, IEEE Trans. Automat. Control, 48 (11) (2003) 1892–1907.

[5] K. Fu, H.F. Chen, W.X. Zhao, Distributed system identification for linear stochastic systems with binary sensors, Automatica, 141 (2022) 110298.

[6] Y. Wang, Y.L. Zhao, J.F. Zhang, J. Guo, A unified identification algorithm of FIR systems based on binary observations with time-varying thresholds, Automatica 135 (2022) 109990.

[7] K. You Recursive algorithms for parameter estimation with adaptive quantizer, Automatica, 52 (2015) 192–201.

[8] J. Guo, L.Y. Wang, G. Yin, Y.L. Zhao, J.F. Zhang Asymptotically efficient identification of FIR systems with quantized observations and general quantized inputs, Automatica 57 (2015) 113–122.

[9] D. Marelli, K. You, M. Fu, Identification of ARMA models using intermittent and quantized output observations, Automatica, 49 (2) (2013) 360–369.

[10] Y.L. Zhao, H. Zhang, T. Wang, G. Kang, System identification under saturated precise or set-valued measurements, Sci. China Inf. Sci. 66 (2023) 112204.

[11] J. Guo, Y.L. Zhao (2013). Recursive projection algorithm on FIR system identification with binary-valued observations, Automatica, 49, 3396–3401.

[12] H. Zhang, T. Wang, Y.L. Zhao, Asymptotically efficient recursive identification of FIR systems with binary-valued observations, IEEE Trans. Syst. Man Cybern. Syst. 51 (5) (2021) 2687–2700.

[13] L.T. Zhang, Y.L. Zhao, L. Guo. Identification and adaptation with binary-valued observations under non-persistent excitation condition, Automatica 138 (2022) 110158.

[14] Q.J. Song, Recursive identification of systems with binary-valued outputs and with ARMA noises, Automatica 93 (2018) 106–113.

[15] Z. Huang, Q.J. Song (2024). Identification of linear systems using binary sensors with random thresholds, J. Syst. Sci. Complex., 37 (3) (2024), 907-923.

[16] C.L. Kong, Y. Wang, Y.L. Zhao. Asymptotic consensus of multi-agent systems under binary-valued observations and observation uncertainty, Systems Control Lett. 182 (2023) 105656.

[17] J. Guo, J. Chen, J.D. Diao, System identification with binary-valued output observations under either-or communication and data packet dropout, Systems Control Lett. 156 (2021) 105010.

[18] J.M. Ke, Y. Wang, Y.L. Zhao, J.F. Zhang, Recursive identification of binary-valued systems under uniform persistent excitations, IEEE Trans. Automat. Control (2024), https://doi.org/10.1109/TAC.2024.3399968.

[19] L.Y. Wang, G. Yin, J.F. Zhang, Joint identification of plant rational models and noise distribution functions using binary-valued observations, Automatica 42 (4) (2006) 535–547.

[20] L.Y. Wang, G. Yin, Y.L. Zhao, J.F. Zhang, Identification input design for consistent parameter estimation of linear systems with binary-valued output observations, IEEE Trans. Automat. Control 53 (4) (2008) 867–880.

[21] K. Jafari, J. Juillard, M. Roger, Convergence analysis of an online approach to parameter estimation problems based on binary observations, Automatica 48 (11) (2012) 2837–2842.

[22] B.I. Godoy, G.C. Goodwin, J.C. Agüero, D. Marelli, T. Wigren, On identification of FIR systems having

quantized output data, Automatica 47 (9) (2011) 1905–1915.
[23] Y.L. Zhao, T. Wang, W. Bi, Consensus protocol for multi-agent systems with undirected topologies and binary-valued communications, IEEE Trans. Automat. Control 64 (1) (2019) 206–221.
[24] H.F. Chen, L. Guo, Strong consistency of recursive identification by no use of persistent excitation condition, Acta Mathematicae Applicatae Sinica, 2 (2) (1985) 133–145.
[25] H.F. Chen, L. Guo, Asymptotically optimal adaptive control with consistent parameter estimates, SIAM J. Control Optim. 25 (3) (1987) 558–575.
[26] D. Gan, Z.X. Liu, Convergence of the distributed SG algorithm under cooperative excitation condition, IEEE Trans. Neural Netw. Learn. Syst. 35 (5) (2024) 7087–7101.
[27] R.S. Risuleo, G. Bottegal, H. Hjalmarsson, Identification of linear models from quantized data: a midpoint-projection approach, IEEE Trans. Automat. Control 65 (7) (2020) 2801–2813.
[28] Y. Wang, X. Li, Y.L. Zhao, J.F. Zhang, Threshold selection and resource allocation for quantized identification, J. Syst. Sci. Complex. 37 (1) (2024) 204–229.
[29] V. Cerone, D. Piga, D. Regruto, Fixed-order FIR approximation of linear systems from quantized input and output data, Systems Control Lett. 62 (12) (2013) 1136-1142.
[30] M. Fu, L. Xie, Finite-level quantized feedback control for linear systems, IEEE Trans. Automat. Control 54 (5) (2009) 1165–1170.
[31] V.A. Zorich, Mathematical Analysis I, second ed., Springer-Verlag, Berlin, 2015.
[32] J.M. Wang, J.M. Ke, J.F. Zhang, Differentially private bipartite consensus over signed networks with time-varying noises, IEEE Trans. Automat. Control (2024) http://dx.doi.org/10.1109/TAC.2024.3351869.
[33] H.F. Chen, Stochastic Approximation and Its Applications, Kluwer Academic Publishers, Dordrecht, 2003.
[34] L. Ljung, System Identification: Theory for the User, second ed., Prentice-Hall, Englewood Cliffs, NJ, 1999.
[35] L. Guo, Time-Varying Stochastic Systems: Stability and Adaptive Theory, second ed., Science Press, Beijing, 2020.
[36] H. Chen, L. Guo, Convergence of adaptive MPC for linear stochastic systems, Sci. China Inf. Sci. 66 (2023) 152201.
[37] E. Fisher, On the law of the iterated logarithm for martingales, Ann. Probab. 20 (2) (1992) 675-680.
[38] T. Wang, X. Li, J. Guo, Y.L. Zhao, Identification of ARMA models with binary-valued observations, Automatica 149 (2023) 110832.
[39] A.N. Shiryaev, Probability, third ed., Springer Science & Business Media, New York, 2016.
[40] T.C. Hu, A. Rosalsky A note on the de La Vallée Poussin criterion for uniform integrability, Stat. Probabil. Lett. 81 (1) (2011) 169–174.
[41] S.P. Tan, J. Guo, Y.L. Zhao, J.F. Zhang, Adaptive control with saturation-constrainted observations for drag-free satellites: a set-valued identification approach, Sci. China Inf. Sci. 64 (2021) 202202.