Automatica 135 (2022) 109990

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

A unified identification algorithm of FIR systems based on binary observations with time-varying thresholds☆

Ying Wang [a,b], Yanlong Zhao [a,b,*], Ji-Feng Zhang [a,b], Jin Guo [c]

[a] Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, PR China
[b] School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, PR China
[c] The School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, PR China

ARTICLE INFO

Article history:
Received 23 August 2020
Received in revised form 5 July 2021
Accepted 1 September 2021
Available online 15 November 2021

Keywords:
System identification
Binary-valued observations
Stochastic framework
Deterministic framework

ABSTRACT

This paper is concerned with parameter estimation of finite impulse response (FIR) systems with binary observations. Combining a suitable design of the time-varying thresholds, a kind of sign-error type unified algorithm with projection is investigated for either deterministic systems or stochastic systems. The convergence properties of the studied algorithm are established under bounded persistent excitations. Specifically, for the case without noise, the square convergence rate is proved to be close to $O\left(\frac{1}{k^2}\right)$ with respect to the time step k. For the case with bounded noises, the upper bound of the estimation error is obtained, which depends on the bound of the noises and the lower bound of the input persistent excitation condition. For the case with independent and identically distributed (i.i.d.) stochastic noises, the estimate is shown to converge to the true parameter in the sense of mean square and almost surely. Besides, the mean square convergence rate of the estimation error is of the order $O\left(\frac{1}{k}\right)$. Numerical examples are supplied to demonstrate the theoretical results.

1. Introduction

Recently, set-valued systems have increasingly emerged in our daily life and attracted widespread attention, which could be found in plentiful important application fields such as engineering area or bio-medical fields (Akyildiz et al., 2002; Wang et al., 2010). For instance, in communication systems like ATM (asynchronous transmission mode) networks, the traffic information, e.g., bit rate, queue length, is measured through binary sensors with appropriate thresholds (Wang et al., 2003). In automotive systems, binary sensors are also often used in such as switching sensors for exhaust gas oxygen (Brailsford et al., 1993).

Motivated by extensive applications, the past two decades have witnessed considerable amounts of works on system identification using binary-valued data (Carbone et al., 2020; Colinet & Juillard, 2010; Depraetere et al., 2012; Godoy et al., 2011; Tan et al., 2021; Wang & Yin, 2007; You, 2015). In the primary work, Wang et al. (2003) introduced this issue and proposed respectively stochastic and deterministic frameworks to identify the unknown parameter. Moreover, they provided exact or approximate solutions to almost all of the basic issues, like estimation quality evaluation, optimal input design, time complexity or estimate convergence, etc. Since then, stochastic and deterministic frameworks are developed almost separately under binary-valued measurements.

Under the stochastic framework, there are a series of achievements about system identification with binary observations. For example, empirical measure method was one of the earliest off-line methods to identify the system parameters with binary-valued signals (Wang & Yin, 2007; Wang et al., 2003). The supervised learning algorithm such as Support Vector Machines, was used to estimate unknown parameters for binary-valued systems by formulating the identification problem as a classification problem (Goudjil et al., 2015). Expectation maximization method was proposed to estimate system parameters with set-valued data (Godoy et al., 2011). All of the above identification methods are off-line. As for online methods, Guo and Zhao (2013) provided a kind of recursive projection algorithm for FIR systems and gave

☆ The work is supported by National Key R&D Program of China under Grant 2018YFA0703800, National Natural Science Foundation of China (61877057 and 62025306), and Open Fund Program of Beijing National Research Center for Information Science and Technology. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Wei Xing Zheng under the direction of Editor Torsten Söderström.

* Corresponding author at: Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, PR China.

E-mail addresses: wangying96@amss.ac.cn (Y. Wang), ylzhao@amss.ac.cn (Y. Zhao), jif@iss.ac.cn (J.-F. Zhang), guojin@ustb.edu.cn (J. Guo).

https://doi.org/10.1016/j.automatica.2021.109990
0005-1098/

the convergence properties of the algorithm under certain conditions. Pouliquen et al. (2016) proposed a recursive least squares algorithm and gave the bound of the estimation error under binary-valued observations. Song (2018) developed a stochastic approximation algorithm with expanding truncations for the systems with additive ARMA noises and binary measurements. It is worth noticing that almost all of the above methods depend on the probability distribution function (PDF) or the statistical properties of the noises. Once the noises are removed or the properties of noises are not enough good, the above methods are out of operation.

Under the deterministic framework, system identification with binary-valued measurements also has some developments (Casini et al., 2007, 2011; Jafari et al., 2010; Jafari et al., 2012). For instance, Casini et al. (2007, 2011) discussed the optimal input design problem for FIR systems and gave a smaller upper bound of time complexity than Wang et al. (2003). Jafari et al. (2010) proposed a recursive method to solve the parameter estimation problem for the noise-free FIR systems with the known norm of unknown parameters under binary observations, and then, Jafari et al. (2012) improved the algorithm using adaptive regulative coefficient and analyzed the algorithm convergence. However, the former requires designable inputs, the latter requires that the norm of unknown parameters is known. These are exactly what most practical systems cannot assure.

There are some works attempting to combine the two frameworks. For example, Zhao et al. (2009) proposed a joint identification algorithm by use of the features of both stochastic and deterministic frameworks, to solve the estimation problem with bounded stochastic noises and binary-valued observations. They introduced a joint framework, in which two frameworks play complementary roles in improving identification accuracy. This is a combination of the advantages of two frameworks rather than a unified method to solve identification problems under the two frameworks. In essence, it is a deterministic method to solve the identification problem under the deterministic framework, and vice versa.

From the above analysis, almost all of the previous works are for the stochastic framework or deterministic framework, or combine the merits of these two frameworks to improve identification accuracy. Due to different structures and analysis methods of identification algorithms under two frameworks, it is difficult to work them with a unified algorithm. So far, no method can solve simultaneously the system identification problem under two frameworks.

Motivated by the above-mentioned facts, this paper investigates a unified identification algorithm for FIR systems with binary-valued observations under two frameworks. In particular, a kind of binary quantizer with time-varying thresholds is chosen and designed to supply a little more information than binary, uniform, and logarithmic quantizers with fixed thresholds. This kind of binary quantizer makes it possible to estimate the unknown parameter under bounded noises as well as stochastic noises. Under stochastic framework, there are some studies about the sign-error type estimation algorithm and system identification with time-varying binary-valued observations (Chen & Yin, 2003; Csáji & Weyer, 2012; Yin et al., 2003; You, 2015; Zhao et al., 2017). For instance, Chen and Yin (2003) investigated the asymptotic properties of a sign-error type estimation algorithm with expanding truncation bounds under stationary and ergodic signals with bounded variance. With the help of this sign-error type estimation algorithm with expanding truncations, Csáji and Weyer (2012) and Zhao et al. (2017) considered the problem about identifying ARX systems and identifying nonlinear FIR systems respectively with adaptive binary-valued measurements under i.i.d. input signals with bounded variance. In addition, You (2015) developed a stochastic approximation type recursive estimator for FIR systems with adaptive binary observations under i.i.d. input signals.

However, all of the above works require i.i.d. inputs or stationary and ergodic inputs, resulting in their results are limited to be applied to the adaptive control problem. Therefore, this paper studies a control-oriented identification algorithm for FIR systems with binary-valued observations under a persistent excitation condition, which does not need independent and identically distributed or stationary and ergodic assumptions on the system input signals. The main contributions of this paper can be summarized as follows:

- The sign-error type recursive projection algorithm, which is a kind of sign-error type unified algorithm with projection and has the same gain as Guo and Zhao (2013), is investigated for FIR systems with binary-valued observations under both stochastic framework and deterministic framework. Actually, it is the first attempt to unify these two frameworks with one algorithm under persistent excitations. Furthermore, this is also the first result of sign-error type algorithm in a deterministic setting to the best of our knowledge.
- For the noise-free case, the square convergence of the algorithm is established and the convergence rate is proved to be close to $O\left(\frac{1}{k^2}\right)$, which is not only faster than that for the case with stochastic noises, but also the same order as that of gradient algorithm in Chen and Guo (1991) with accurate measurements. In contrast with Jafari et al. (2012), the algorithm studied in this paper does not require knowing the norm of the unknown parameter in advance.
- For the case with bounded noises, the upper bound of estimation error is given under persistent excitations. The condition on system inputs is broadened from the input designable condition compared with Casini et al. (2007) and Casini et al. (2011).
- For the case with stochastic noises, the algorithm convergence is proved in the sense of almost sure and mean square under persistent excitations. In addition, the mean square convergence rate of estimation error is established as $O\left(\frac{1}{k}\right)$, which is faster than Guo and Zhao (2013). Usually, the existing results about sign-error type algorithms need the inputs are i.i.d. (You, 2015) or stationary and ergodic (Chen & Yin, 2003), which is not required in this paper.

The rest of this paper is organized as follows. Section 2 describes the identification problem and introduces the sign-error type recursive projection algorithm. Section 3 starts the investigation on system identification in the deterministic framework, including the noise-free case and the bounded noise case, the convergence and the convergence rate of the algorithm are established. Section 4 is in the stochastic framework and gives the convergence properties of the algorithm studied for the case with stochastic noises. Section 5 gives three numerical examples to demonstrate the main results. Section 6 gives the concluding remarks and related future works.

Notations. In this paper, $\mathbb{Z}$ and $\mathbb{Z}_+$ are the set of integers and positive integers, respectively. $\mathbb{R}$ and $\mathbb{R}^n$ are the sets of real number and n-dimensional real vectors, respectively. $\|x\| = \|x\|_2$ is the Euclidean norm. I_n is an n-dimension identity matrix. Besides, sign$\{x\}$ is the sign of scalar x. $\lfloor x \rfloor = \max\{a \in \mathbb{Z} | a \leq x\}$ and $\lceil x \rceil = \min\{a \in \mathbb{Z} | a \geq x\}$ for $x \in \mathbb{R}$. The function $I_{\{\cdot\}}$ denotes the indicator function, whose value is 1 if its argument (a formula) is true, and 0, otherwise.

2. Problem formulation

Consider the FIR system, described by

$$y_{k+1}=\varphi_k^T\theta+d_{k+1},\ k\geq 0, \tag{1}$$

where $\varphi_k\in\mathbb{R}^n$, $\theta\in\mathbb{R}^n$ and $d_{k+1}\in\mathbb{R}$ are the system input, time-invariant parameter and noise, respectively. And $y_{k+1}\in\mathbb{R}$ is the system output, which cannot be exactly measured and could only be measured by quantized information

$$s_{k+1}=I_{\{y_{k+1}>\varphi_k^T\hat{\theta}_k\}}-I_{\{y_{k+1}<\varphi_k^T\hat{\theta}_k\}}, \tag{2}$$

where $\hat{\theta}_k$ is the estimate of θ at time k.

Though this problem has been separately studied in the deterministic framework (including the noise-free case and the bounded noise case) and the stochastic framework (i.e., the case with stochastic noises), there is no method that can achieve system identification simultaneously in these two frameworks so far. The goal of this paper is to develop a unified algorithm to estimate the unknown parameter θ based on the system input φ_k, the binary-valued output s_{k+1} and the properties of noises in both the deterministic framework and the stochastic framework.

2.1. Assumptions

In order to proceed our analysis, we introduce some assumptions concerning the inputs, the priori information of the unknown parameter and the noises.

Assumption 2.1. The input sequence $\{\varphi_k\}$ satisfies $\sup_{k\geq 1}\|\varphi_k\|\triangleq\bar{\varphi}<\infty$. Besides, there exist a positive integer N and a positive constant δ such that

$$\frac{1}{N}\sum_{i=k}^{k+N-1}\varphi_i\varphi_i^T\geq\delta^2 I_n,\ \forall k\geq 1. \tag{3}$$

Assumption 2.2. There is a known bounded convex compact set $\Theta\subset\mathbb{R}^n$ such that the unknown parameter $\theta\in\Theta$. And denote $\bar{\theta}=2\sup_{\eta\in\Theta}\|\eta\|$.

Assumption 2.3. The noise sequence $\{d_k\}$ is assumed to be bounded by a known quantity $\bar{d}$, i.e., $|d_k|<\bar{d}$ for all $k\geq 1$.

Assumption 2.4. The noise $\{d_k\}$ is a sequence of i.i.d. random variables with zero median. Besides, the density function of d_1 is denoted by $f(x)=\frac{dF(x)}{dx}$, which satisfies

$$\inf_{x\in[-\bar{\varphi}\bar{\theta},\bar{\varphi}\bar{\theta}]}f(x)\triangleq f_{\bar{\theta}}>0, \tag{4}$$

where $F(\cdot)$ is the PDF of d_1.

Remark 2.1. It is worth pointing that the existence of the lower bound $f_{\bar{\theta}}$ is required instead of knowing the probability distribution function $F(\cdot)$ and density function $f(\cdot)$ of the noises, when analyzing the convergence of the algorithm. Thus, $F(\cdot)$ and $f(\cdot)$ can be unknown in this paper. Moreover, if the median of d_k is $\nu\neq 0$, let $\check{\varphi}_k=[\varphi_k,1]^T$, $\check{\theta}=[\theta,\nu]^T$ and $\check{d}_{k+1}=d_{k+1}-\nu$. Then the original FIR system is converted to $y_{k+1}=\check{\varphi}_k^T\check{\theta}+\check{d}_{k+1}$, which satisfies Assumption 2.4. Since the algorithm to be designed is independent of ν, we are able to estimate the median ν of the noise as well. Therefore, without loss of generality, we assume that $\nu=0$ throughout the paper. Furthermore, there are many common noises conforming Assumption 2.4, such as Gaussian white noises.

2.2. Algorithm

For a given initial estimate $\hat{\theta}_0$, based on the input φ_k and the binary-valued observation s_{k+1}, the sign-error type recursive projection algorithm is recursively designed at any iteration $k\geq 0$ as follows:

$$\begin{cases}\hat{\theta}_{k+1}=\Pi_\Theta\left(\hat{\theta}_k+\beta\dfrac{\varphi_k}{r_{k+1}}s_{k+1}\right),\\ r_{k+1}=1+\displaystyle\sum_{i=1}^{k}\varphi_i^T\varphi_i,\end{cases} \tag{5}$$

where $\beta>0$ is a constant scalar and $\Pi_\Theta(\cdot)$ is the projection mapping from $\mathbb{R}^n$ to Θ, which is defined as

$$\Pi_\Theta(\xi)=\arg\min_{\zeta\in\Theta}\|\xi-\zeta\|,\quad\forall\xi\in\mathbb{R}^n. \tag{6}$$

Remark 2.2. The algorithm is motivated by Guo and Zhao (2013), where $s_{k+1}=F(C-\varphi_k^T\hat{\theta}_k)-I_{\{y_{k+1}\leq C\}}$ and C is the fixed threshold of binary quantizer. Noting that $\mathbb{E}[s_{k+1}]=F(C-\varphi_k^T\hat{\theta}_k)-F(C-\varphi_k^T\theta)$ and the probability distribution function $F(\cdot)$ of noises is monotonous, so $\mathbb{E}[s_{k+1}]$ shows the information about the "innovation" of outputs, i.e., the difference between the output and its estimate. The innovation of output, defined as $y_{k+1}-\varphi_k^T\hat{\theta}_k$, is often used to construct identification algorithms under accurate measurements (Chen & Guo, 1991). These enlighten us with the signum of innovation instead of itself or other information about it to design identification algorithms. Inspired by Guo and Zhao (2014), we choose $\varphi_k^T\hat{\theta}_k$ as the time-varying thresholds of binary quantizer and $s_{k+1}=\text{sign}\{y_{k+1}-\varphi_k^T\hat{\theta}_k\}$ in the algorithm. Compared with Guo and Zhao (2013), the algorithm (5) does not require the noise PDF to update the recursions, which makes it is possible to use this algorithm in the noise-free case or the case that the noise PDF is unknown.

Remark 2.3. (Calamai & Moré, 1987) The projection function given by (6) has the following property

$$\|\Pi_\Theta(\xi_1)-\Pi_\Theta(\xi_2)\|\leq\|\xi_1-\xi_2\|,\quad\forall\xi_1,\xi_2\in\mathbb{R}^n.$$

2.3. Remarks on quantizer's choice

In this paper, the quantizer (2), chosen to measure the system output, is the binary quantizer with time-varying thresholds. Why do we choose this kind of quantizer except for constructing the "innovation" of outputs?

Firstly, we consider three typical quantizers, which are binary quantizer with fixed threshold (Godoy et al., 2011; Wang et al., 2003), uniform quantizer (Fu & Wang, 2012) and logarithmic quantizer (Fu & Xie, 2005).

Binary quantizer with fixed threshold: A binary quantizer $q_b:\mathbb{R}\to\mathbb{R}$ is defined as

$$q_b(x)=I_{\{x\leq C\}}, \tag{7}$$

where $C\in\mathbb{R}$ is the fixed threshold of binary quantizer.

Uniform quantizer: A uniform quantizer $q_u:\mathbb{R}\to\mathbb{R}$ is defined as

$$q_u(x)=\varepsilon\left\lfloor\frac{x}{\varepsilon}+\frac{1}{2}\right\rfloor. \tag{8}$$

where $\varepsilon>0$ is the quantization parameter. For a uniform quantizer, the quantization error is always bounded by $\frac{\varepsilon}{2}$, i.e., $|q_u(x)-x|\leq\frac{\varepsilon}{2}$ for all $x\in\mathbb{R}$.

Logarithmic quantizer: A logarithmic quantizer $q_l : \mathbb{R} \to \mathbb{R}$ is defined as

$$q_l(x) = \begin{cases} \omega_i, & \text{if } \frac{1}{1+\rho}\omega_i < x \le \frac{1}{1-\rho}\omega_i, \\ 0, & \text{if } x = 0, \\ -q(-x), & \text{if } x < 0, \end{cases} \quad (9)$$

where $\omega_i = \omega_0 \left(\frac{1-\rho}{1+\rho}\right)^i$ for $i \in \mathbb{Z}$, $\omega_0 > 0$ and $\rho \in (0, 1)$ is the quantization parameter.

As for the noise-free case, Jing and Zhang (2019) proved estimation error is $O(\varepsilon)$ for noise-free ARMAX systems with uniform quantizer (8) under persistent excitations. Actually, the accurate parameter estimation cannot be realized for the noise-free case with these three quantizers to the best of our knowledge, when the inputs satisfy persistent excitation condition. A simple example is given to show that,

$$x = au,$$

where $a \in \mathbb{R}$ is unknown and $a \ne 0$. Let $u \equiv 1$, which satisfies the bounded persistent excitation condition (i.e., Assumption 2.1).

Using binary quantizer (7), we get $a > C$ or $a \le C$. It can be seen that $a \in \left[q_u(a) - \frac{\varepsilon}{2}, q_u(a) + \frac{\varepsilon}{2}\right)$ with uniform quantizer (8). By logarithmic quantizer (9), we learn $a \in \left(\frac{(1-\rho)^i\omega_0}{(1+\rho)^{i+1}}, \frac{(1-\rho)^{i-1}\omega_0}{(1+\rho)^i}\right]$ or $\left(-\frac{(1-\rho)^{i-1}\omega_0}{(1+\rho)^i}, -\frac{(1-\rho)^i\omega_0}{(1+\rho)^{i+1}}\right]$, $i \in \mathbb{Z}$. These are the only information we can get with $u \equiv 1$ and three quantizers. Whatever we design algorithms, we cannot identify accurately the parameter with these quantizers. Moreover, we learn that the parameter can be accurately estimated if quantizer parameters ε and ρ are zero for the uniform quantizer and the logarithmic quantizer. But it is accurate measurements rather than quantized measurements.

From the above analysis, parameter estimation cannot be realized accurately under persistent excitations for the noise-free case, no matter how we choose binary, uniform or logarithmic quantizer with any fixed thresholds. Thus, we choose the binary quantizer with time-varying thresholds, which can provide more information about the parameter than fixed thresholds. Moreover, there are lots of applications about time-varying thresholds in practical fields (Knudson et al., 2016; Zahabi et al., 2017). An example is the coding process in communications (Akyildiz et al., 2002), which is a kind of protocol that can be adjusted on the basis of actual needs. In this paper, the binary quantizer s_{k+1} with time-varying thresholds tells whether the actual output is bigger than its estimate or not, which replaces the role of the noise PDF, and with which we can adjust the iteration direction of the algorithm to estimate unknown parameters. Therefore, time-varying thresholds are the key to achieve system identification under persistent excitations for the noise-free case.

There are some existing studies of system identification under binary measurements with time-varying thresholds in the stochastic framework (Csáji & Weyer, 2012; You, 2015; Zhao et al., 2017), which are based on the sign-error type estimation algorithms (Chen & Yin, 2003; Yin et al., 2003). In these studies of the sign-error type estimation methods, the following condition is required in convergence analysis. The sequence $\{(y_{k+1}, \varphi_k)\}$ is stationary and ergodic such that $\mathbb{E}\begin{pmatrix} y_{k+1}^2 & y_{k+1}\varphi_k^T \\ y_{k+1}\varphi_k & \varphi_k\varphi_k^T \end{pmatrix} = R > 0$. It is easy to see that this condition cannot be satisfied for the noise-free case, since y_{k+1} is linearly dependent on φ_k. In addition, Csáji and Weyer (2012), You (2015) and Zhao et al. (2017) studied system identification with these time-varying thresholds for i.i.d. stochastic noise case, whose theoretic analyses depended on the statistical properties of noises. Thus, their analysis methods are inapplicable to the noise-free case.

3. Deterministic framework

This section will focus on the convergence analysis of the algorithm (5) under the deterministic framework. The convergence properties will be established for the noise-free and bounded noise case, respectively. It is worth noticing that the bounded noise case includes the bounded stochastic noise case. The reason classifying the bounded noise case as the deterministic framework, is that only the bound of noises is used when analyzing the algorithm performance for the bounded noise case.

3.1. The noise-free case

Generally speaking, parameter estimation for the noise-free case is more difficult than the stochastic noise case under binary-valued observations, since the stochastic noise can excite the information of parameters to the limited binary-valued measurements. As a result, most of the existing recursive algorithms are derived by the noise PDF, or the convergence analysis depends on the good statistical properties of the noises (Goudjil et al., 2015; Guo & Zhao, 2013; Wang et al., 2003). Once the noises are removed, the above methods are no longer applicable.

In the current studies, there are two methods to estimate the unknown parameters for the noise-free case with binary-valued measurements. One is input design method (Casini et al., 2007, 2011; Wang et al., 2003), the other is recursive method (Jafari et al., 2010; Jafari et al., 2012). The former requires designable inputs, the latter requires that the norm of the parameter is known. These limits make the above two methods difficult to be applied in practice.

In this part, we consider the noise-free case, i.e., $d_{k+1} = 0$ for System (1)–(2), which is described as

$$\begin{cases} y_{k+1} = \varphi_k^T\theta, \\ s_{k+1} = -\operatorname{sign}\{\varphi_k^T\tilde{\theta}_k\}. \end{cases} \quad (10)$$

where $\tilde{\theta}_k = \hat{\theta}_k - \theta$ is the estimation error.

Firstly, we show the convergence of the algorithm (5) in the following theorem under the bounded persistent excitation condition.

Theorem 1. *For the noise-free FIR system (10), under Assumptions 2.1 and 2.2, the estimate $\hat{\theta}_k$ defined by the algorithm (5) converges to the true parameter θ for any initial value $\hat{\theta}_0$, i.e.,*

$$\lim_{k\to\infty} \tilde{\theta}_k = 0.$$

Proof. By (5) and Remark 2.3, we have

$$\begin{aligned} \tilde{\theta}_{k+1}^T\tilde{\theta}_{k+1} \le& \tilde{\theta}_k^T\tilde{\theta}_k + \frac{2\beta\varphi_k^T\tilde{\theta}_k}{r_{k+1}}s_{k+1} + \frac{\beta^2\varphi_k^T\varphi_k s_{k+1}^2}{r_{k+1}^2} \\ \le& \tilde{\theta}_k^T\tilde{\theta}_k - \frac{2\beta|\varphi_k^T\tilde{\theta}_k|}{r_{k+1}} + \frac{\beta^2\varphi_k^T\varphi_k s_{k+1}^2}{r_{k+1}^2}. \end{aligned} \quad (11)$$

Then, from $|s_{k+1}| \le 1$ and $r_{k+1} \le 1 + k\bar{\varphi}^2$ it follows that

$$\begin{aligned} \|\tilde{\theta}_{k+1}\|^2 \le& \|\tilde{\theta}_k\|^2 - \frac{2\beta|\varphi_k^T\tilde{\theta}_k|}{r_{k+1}} + \frac{\beta^2\|\varphi_k\|^2}{r_{k+1}^2} \\ \le& \|\tilde{\theta}_0\|^2 - \sum_{i=0}^{k}\frac{2\beta|\varphi_i^T\tilde{\theta}_i|}{r_{i+1}} + \sum_{i=0}^{k}\frac{\beta^2\|\varphi_i\|^2}{r_{i+1}^2} \\ \le& \|\tilde{\theta}_0\|^2 - \sum_{i=0}^{k}\frac{2\beta|\varphi_i^T\tilde{\theta}_i|}{1+i\bar{\varphi}^2} + \sum_{i=0}^{k}\frac{\beta^2\|\varphi_i\|^2}{r_{i+1}^2}. \end{aligned}$$

Noticing $\|\tilde{\theta}_{k+1}\|^2 \geq 0$ and

$$\sum_{i=0}^{k} \frac{\|\varphi_i\|^2}{r_{i+1}^2} \leq \sum_{i=1}^{k} \frac{\|\varphi_i\|^2}{r_{i+1} r_i} \leq \sum_{i=1}^{k} \left(\frac{1}{r_i} - \frac{1}{r_{i+1}} \right) \leq 1 - \frac{1}{r_{k+1}} < \infty,$$

we can get $\sum_{i=0}^{k} \frac{2\beta|\varphi_i^T \tilde{\theta}_i|}{1+i\bar{\varphi}^2} < \infty$, which implies

$$\lim_{k\to\infty} |\varphi_k^T \tilde{\theta}_k| = 0. \tag{12}$$

In addition, by Lemma A.1 we have

$$\begin{aligned} |\varphi_{k+j}^T \tilde{\theta}_k| &= |\varphi_{k+j}^T (\tilde{\theta}_k - \tilde{\theta}_{k+j}) + \varphi_{k+j}^T \tilde{\theta}_{k+j}| \\ &\leq \|\varphi_{k+j}\| \cdot \|\tilde{\theta}_{k+j} - \tilde{\theta}_k\| + |\varphi_{k+j}^T \tilde{\theta}_{k+j}| \\ &\leq \frac{j\beta\bar{\varphi}}{r_{k+1}} + |\varphi_{k+j}^T \tilde{\theta}_{k+j}|, \end{aligned} \tag{13}$$

for $j = 1, 2, \ldots, N-1$. From (3) in Assumption 2.1,

$$\begin{aligned} r_{k+1} &\geq 1 + \sum_{i=1}^{\lfloor \frac{k}{N} \rfloor N} \varphi_i^T \varphi_i = 1 + \sum_{j=1}^{\lfloor \frac{k}{N} \rfloor} \sum_{i=1+(j-1)N}^{jN} \varphi_i^T \varphi_i \\ &\geq 1 + \sum_{j=1}^{\lfloor \frac{k}{N} \rfloor} n\lambda_{\min} \left(\sum_{i=1+(j-1)N}^{jN} \varphi_i \varphi_i^T \right) \\ &\geq 1 + nN\delta^2 \left\lfloor \frac{k}{N} \right\rfloor \geq 1 + n\delta^2 (k - N). \end{aligned} \tag{14}$$

From (12)–(14) it follows that $\lim_{k\to\infty} |\varphi_{k+j}^T \tilde{\theta}_k| = 0$ for $j = 1, 2, \ldots, N-1$, which together with (3) gives

$$\begin{aligned} \delta^2 N \|\tilde{\theta}_k\|^2 &\leq \tilde{\theta}_k^T \left(\sum_{j=0}^{N-1} \varphi_{k+j} \varphi_{k+j}^T \right) \tilde{\theta}_k \\ &= \sum_{j=0}^{N-1} |\varphi_{k+j}^T \tilde{\theta}_k|^2 \to 0, k \to \infty. \end{aligned}$$

Thus, we get $\lim_{k\to\infty} \tilde{\theta}_k = 0$. □

Nextly, we will prove that the square convergence rate of the estimation error can reach $O\left(\frac{1}{k^2}\right)$ under the bounded persistent excitation condition.

Theorem 2. *Under the conditions of Theorem 1, if $\beta > \frac{\bar{\varphi}^2+1}{2\delta^2}$, the estimation error given by the algorithm (5) has the convergence rate*

$$\|\tilde{\theta}_k\|^2 = O\left(\frac{1}{k^r}\right),$$

for arbitrary positive number $r \in (1, 2)$.

Proof. The main idea of proof is as follows.

$$O\left(\frac{1}{k}\right) \to O\left(\frac{1}{k^{\frac{r+1}{2}}}\right) \to \cdots \to O\left(\frac{1}{k^{r+\frac{1-r}{2^m}}}\right) \to O\left(\frac{1}{k^r}\right).$$

Step 1: to prove $\|\tilde{\theta}_k\|^2 = O(\frac{1}{k})$ by Lemma A.3 and Assumption 2.1.

Step 2: to prove $\|\tilde{\theta}_k\|^2 = O\left(\frac{1}{k^{1+t_1}}\right)$ using $\|\tilde{\theta}_k\|^2 = O\left(\frac{1}{k}\right)$, where $t_1 = \frac{r-1}{2} > 0$.

⋮

Step m + 1: to prove $\|\tilde{\theta}_k\|^2 = O\left(\frac{1}{k^{1+t_m}}\right)$ by use of $\|\tilde{\theta}_k\|^2 = O\left(\frac{1}{k^{1+t_{m-1}}}\right)$ similarly to Step 2, where $t_m = t_{m-1} + \frac{1}{2^{m-1}} t_1 = t_1 + \frac{1}{2} t_1 + \cdots + \frac{1}{2^{m-1}} t_1 = \frac{2^m - 1}{2^m}(r-1)$.

Finally, repeating the above process, we can get $\|\tilde{\theta}_k\|^2 = O\left(\frac{1}{k^r}\right)$ for any $r \in (1, 2)$ from $\lim_{m\to\infty} t_m = r - 1$.

The detailed proof is given in Appendix B. □

Remark 3.1. This theorem illustrates that the square convergence rate of the algorithm (5) can get close to $O\left(\frac{1}{k^2}\right)$ for the noise-free system (10). As a matter of fact, it is also the fastest convergence rate that the following gradient algorithm can reach for the noise-free case with even accurate measurements under the bounded persistent excitation condition.

Gradient algorithm (Chen & Guo, 1991):

$$\begin{cases} \hat{\theta}_{k+1} = \hat{\theta}_k + \dfrac{\varphi_k}{r_{k+1}} (y_{k+1} - \varphi_k^T \hat{\theta}_k), \\ r_{k+1} = 1 + \displaystyle\sum_{i=1}^{k} \varphi_i^T \varphi_i. \end{cases} \tag{15}$$

Substituting $y_{k+1} = \varphi_k^T \theta$ into (15) gives

$$\tilde{\theta}_{k+1} = \left(I - \frac{\varphi_k \varphi_k^T}{r_{k+1}} \right) \tilde{\theta}_k. \tag{16}$$

Similarly to Theorem 2, we can get the fastest square convergence rate of (15) is $O\left(\frac{1}{k^2}\right)$. What is more, we claim that (15) is not exponential convergent. As Theorem 2.3.2 in Guo (2020) shows, (16) is exponential stable if and only if there exists $h \in \mathbb{Z}_+$ such that $\inf_{k\geq 0} \lambda_{\min} \left(\sum_{i=k+1}^{k+h} \frac{\varphi_i \varphi_i^T}{r_{i+1}} \right) > 0$. From (14) and the boundness of φ_k, $\inf_{k\geq 0} \lambda_{\min} \left(\sum_{i=k+1}^{k+h} \frac{\varphi_i \varphi_i^T}{r_{i+1}} \right) = 0$ for all $h \in \mathbb{Z}_+$, which indicates our claim.

3.2. The impact of bounded noises

Parameter estimation under the bounded noise case is more complex than the noise-free case. In the existing researches about parameter estimation with binary-valued measurements, the minimum achievable error is given by the input design method for the bounded noise case (Casini et al., 2007, 2011; Wang et al., 2003). But when we broaden the input designable condition to the bounded persistent excitation condition, what conclusions can we draw?

For convenience, redescribe the FIR system (1) with binary-valued observation (2) as follows,

$$\begin{cases} y_{k+1} = \varphi_k^T \theta + d_{k+1}, \\ s_{k+1} = -\operatorname{sign}\{\varphi_k^T \tilde{\theta}_k - d_{k+1}\}. \end{cases} \tag{17}$$

The difficult point is how to deal with the effect of the bounded noise, and the following proposition plays a key role in our analysis.

Proposition 1. *If $\{\varphi_l \in \mathbb{R}, l = 1, 2, \ldots, N\}$ satisfy $\frac{1}{N} \sum_{l=1}^{N} \varphi_l^2 \geq \delta^2 > 0$, then there exists $a \in (0, 1)$ such that*

$$\sum_{|\varphi_l| > a\delta} |\varphi_l| > \sum_{|\varphi_l| \leq a\delta} |\varphi_l|.$$

Proof. Noticing $\frac{1}{N} \sum_{l=1}^{N} \varphi_l^2 \geq \delta^2 > 0 (\delta > 0)$, there are at most $N - 1$ inputs φ_l that satisfy $|\varphi_l| \leq a\delta$. Hence, $\max \sum_{|\varphi_l| \leq a\delta} |\varphi_l|$ is not greater than $(N-1)a\delta$.

Noticing $\frac{1}{N}\sum_{l=1}^{N}\varphi_l^2 \geq \delta^2$, by Pigeonhole principle, there is at least an input $\varphi_{l_0}(l_0 \in \{1, 2, \ldots, N\})$ making $|\varphi_{l_0}| \geq \delta$. Hence, $\min \sum_{|\varphi_l|\geq a\delta} |\varphi_l|$ is not less than δ.

Based on the above analysis, when $a \in (0, \frac{1}{N-1}) \cap (0, 1)$,

$$\sum_{|\varphi_l|>a\delta} |\varphi_l| \geq \delta > a(N-1)\delta \geq \sum_{|\varphi_l|\leq a\delta} |\varphi_l|.$$

Remark 3.2. This proposition shows that if N inputs satisfy the persistent excitation condition, then there exists a constant $a\delta$ such that the sum of the inputs that are bigger than $a\delta$ is greater than the sum of other inputs. With the help of this idea, the estimation error could be proved to be reduced by N iterations for the bounded noise case.

In this part, let $\beta = 1$ in the algorithm (5) for the analysis of simplicity. It is rational since $\beta(\beta > 0)$ only affects the algorithm convergence rate rather than convergence, according to the analysis of the noise-free case.

In the following theorem, the upper bound of estimation error is given, which is related to the noise bound and the lower bound of the persistent excitation condition.

Theorem 3. *For the FIR system (17) with bounded noises, if Assumption 2.1, 2.2 and 2.3 hold, then there exists $k_0 > 0$ for the estimation error given by the algorithm (5) such that,*

$$\|\tilde{\theta}_{k+N}\| \leq \frac{\bar{d}}{\alpha\delta} + \frac{N\bar{\varphi}}{r_{k+1}},$$

for $k > k_0$ and $\forall \alpha \in \left(0, \frac{1}{\sqrt{N^3}}\right)$.

The proof is given in Appendix C. □

Remark 3.3. It is worth mentioning that Theorem 3 tries to theoretically give the smallest upper bound of the estimation error. It contains two items: the first item is the main part, which is a bounded constant $\frac{\bar{d}}{\alpha\delta}$ depending on the inputs and the noises; the second item $\frac{N\bar{\varphi}}{r_{k+1}}$ converges to zero as $k \to \infty$ since $r_{k+1} \geq 1+n\delta^2(k-N)$. Even the second item is small, it is necessary since the norm of the estimation error may be larger than $\frac{\bar{d}}{\alpha\delta}$ in some special case.

Remark 3.4. From the proof of Theorem 3 in Appendix C, we learn k_0 depends on the noise bound $\bar{d}$, so this analysis method cannot be used in the noise-free case. But the estimation error $\|\tilde{\theta}_k\|$ is approximate to $O(\frac{1}{k})$ when $\bar{d}$ is sufficiently small, which confirms the conclusion of Theorem 2 to some extent.

Corollary 3.1. *For the 1-order FIR system (17) with bounded noises, if Assumption 2.1, 2.2 and 2.3 hold, then there exists $k_0 > 0$ such that,*

$$\|\tilde{\theta}_{k+N}\| \leq \frac{\bar{d}}{\alpha\delta} + \frac{N\bar{\varphi}}{r_{k+1}},$$

for $k > k_0$ and $\forall \alpha \in \left(0, \frac{1}{N-1}\right) \cap (0, 1)$.

The proof is similar to that of Theorem 3 except for analyzing φ_k instead of $u_k = \varphi_k^T\tilde{\theta}_k$, so the detail proof is omitted.

Remark 3.5. In this section, we show that the sign-error type recursive projection algorithm can solve identification problems in the deterministic framework, including the noise-free case and the bounded noise case. Actually, it is no influence on the algorithm convergence to get rid of projection in the deterministic framework, since these convergence analyses do not depend on the priori information of the unknown parameter. In other words, we can take $\Theta = \mathbb{R}^n$ in Assumption 2.2 in the deterministic framework.

4. Stochastic framework

In order to show the sign-error type recursive projection algorithm (5) could also solve parameter estimation problems in the stochastic framework, we consider the FIR system (1)–(2) with stochastic noises,

$$\begin{cases} y_{k+1} = \varphi_k^T\theta + d_{k+1}, \\ s_{k+1} = I_{\{d_{k+1}>\varphi_k^T\tilde{\theta}_k\}} - I_{\{d_{k+1}<\varphi_k^T\tilde{\theta}_k\}}. \end{cases} \tag{18}$$

In this part, we will show the convergence of the algorithm (5) and give the convergence rate of estimation error for FIR systems with stochastic noises.

First of all, we establish the almost sure and mean square convergence of the algorithm under the bounded persistent excitation condition in the following theorem.

Theorem 4. *For the FIR system (18) with stochastic noises, if Assumption 2.1, 2.2 and 2.4 hold, then the estimate given by the algorithm (5) is both mean-square and almost surely convergent,*

$$\lim_{k\to\infty} \mathbb{E}\|\tilde{\theta}_k\|^2 = 0,$$
$$\lim_{k\to\infty} \tilde{\theta}_k = 0, \ \ a.s.$$

Proof. Considering the algorithm (5), from (18) and Remark 2.3, we have

$$\tilde{\theta}_{k+1}^T\tilde{\theta}_{k+1} \leq \tilde{\theta}_k^T\tilde{\theta}_k + \frac{2\beta\varphi_k^T\tilde{\theta}_k}{r_{k+1}}s_{k+1} + \frac{\beta^2\varphi_k^T\varphi_k s_{k+1}^2}{r_{k+1}^2}.$$

Then, by $|s_{k+1}| \leq 1$ we have

$$\|\tilde{\theta}_{k+1}\|^2 \leq \|\tilde{\theta}_k\|^2 + \frac{2\beta\varphi_k^T\tilde{\theta}_k s_{k+1}}{r_{k+1}} + \frac{\beta^2\|\varphi_k\|^2}{r_{k+1}^2}. \tag{19}$$

We define $\mathcal{F}_k$ as a σ-algebra generated by $d_1, d_2, \ldots, d_k$. Since s_{k+1} only depends on d_{k+1}, s_{k+1} is independent of $\mathcal{F}_k$. Then from Assumption 2.4, it can be seen

$$\begin{aligned}\mathbb{E}[s_{k+1}|\mathcal{F}_k] =& \mathbb{E}[s_{k+1}] = \mathbb{E}\left[I_{\{d_{k+1}>\varphi_k^T\tilde{\theta}_k\}} - I_{\{d_{k+1}<\varphi_k^T\tilde{\theta}_k\}}\right] \\ =& 1 - 2F(\varphi_k^T\tilde{\theta}_k) = 2\left(F(0) - F(\varphi_k^T\tilde{\theta}_k)\right),\end{aligned}$$

where $F(\cdot)$ is the PDF of noises. By differential mean value theorem (Zemouche et al., 2005), there exists ξ_k between 0 and $\varphi_k^T\tilde{\theta}_k$ such that

$$\mathbb{E}[s_{k+1}|\mathcal{F}_k] = -2f(\xi_k)\varphi_k^T\tilde{\theta}_k. \tag{20}$$

By Assumption 2.1, 2.2 and 2.4, we have $r_{k+1} \leq 1 + k\bar{\varphi}^2$, $-\bar{\varphi}\bar{\theta} \leq \xi_k \leq \bar{\varphi}\bar{\theta}$, and hence,

$$f(\xi_k) \geq f_{\bar{\theta}}. \tag{21}$$

Thus, from (19)–(21), we get

$$\begin{aligned}&\mathbb{E}[\|\tilde{\theta}_{k+1}\|^2|\mathcal{F}_k] \leq \|\tilde{\theta}_k\|^2 - \frac{4\beta f(\xi_k)}{r_{k+1}}(\varphi_k^T\tilde{\theta}_k)^2 + \frac{\beta^2\|\varphi_k\|^2}{r_{k+1}^2} \\ &\leq \|\tilde{\theta}_k\|^2 - \frac{4\beta f_{\bar{\theta}}}{r_{k+1}}(\varphi_k^T\tilde{\theta}_k)^2 + \frac{\beta^2\|\varphi_k\|^2}{r_{k+1}^2},\end{aligned} \tag{22}$$

and

$$\begin{aligned}&\mathbb{E}[\|\tilde{\theta}_{k+1}\|^2] \\ \leq& \mathbb{E}[\|\tilde{\theta}_k\|^2] - \frac{4\beta f_{\bar{\theta}}}{r_{k+1}}\mathbb{E}(\varphi_k^T\tilde{\theta}_k)^2 + \frac{\beta^2\|\varphi_k\|^2}{r_{k+1}^2} \\ \leq& \mathbb{E}[\|\tilde{\theta}_k\|^2] - \frac{4\beta f_{\bar{\theta}}}{1+k\bar{\varphi}^2}\mathbb{E}(\varphi_k^T\tilde{\theta}_k)^2 + \frac{\beta^2\|\varphi_k\|^2}{r_{k+1}^2}\end{aligned}$$

$$\le \mathbb{E}[\|\tilde{\theta}_0\|^2] - \sum_{i=0}^{k} \frac{4\beta f_{\bar{\theta}}}{1+i\bar{\varphi}^2}\mathbb{E}(\varphi_i^T\tilde{\theta}_i)^2 + \sum_{i=0}^{k}\frac{\beta^2\|\varphi_i\|^2}{r_{i+1}^2}.$$

Noticing $\mathbb{E}\|\tilde{\theta}_{k+1}\|^2 \ge 0$ and

$$\sum_{i=0}^{k}\frac{\|\varphi_i\|^2}{r_{i+1}^2} \le \sum_{i=1}^{k}\frac{\|\varphi_i\|^2}{r_{i+1}r_i} \le \sum_{i=1}^{k}\left(\frac{1}{r_i}-\frac{1}{r_{i+1}}\right) \le 1-\frac{1}{r_{k+1}} < \infty,$$

we can get $\sum_{i=0}^{k}\frac{\mathbb{E}(\varphi_k^T\tilde{\theta}_k)^2}{1+i\bar{\varphi}^2} < \infty$, which implies

$$\lim_{k\to\infty}\mathbb{E}(\varphi_k^T\tilde{\theta}_k)^2 = 0. \tag{23}$$

In addition, for $j = 1, 2, \ldots, N-1$, we have

$$\begin{aligned}&\mathbb{E}(\varphi_{k+j}^T\tilde{\theta}_k)^2\\ \le&\mathbb{E}\left(\varphi_{k+j}^T\tilde{\theta}_{k+j}\right)^2 + \mathbb{E}\left(\varphi_{k+j}^T(\tilde{\theta}_{k+j}-\tilde{\theta}_k)\right)^2\\ &+2\mathbb{E}\tilde{\theta}_{k+j}^T\varphi_{k+j}\varphi_{k+j}^T(\tilde{\theta}_{k+j}-\tilde{\theta}_k).\\ \le&\mathbb{E}\left(\varphi_{k+j}^T\tilde{\theta}_{k+j}\right)^2 + \mathbb{E}\left(\varphi_{k+j}^T(\tilde{\theta}_{k+j}-\tilde{\theta}_k)\right)^2\\ &+2\left[\mathbb{E}\left(\varphi_{k+j}^T\tilde{\theta}_{k+j}\right)^2\right]^{\frac{1}{2}}\left[\mathbb{E}\left(\varphi_{k+j}^T(\tilde{\theta}_{k+j}-\tilde{\theta}_k)\right)^2\right]^{\frac{1}{2}}.\end{aligned}$$

By Lemma A.1, we can get $\mathbb{E}\left(\varphi_{k+j}^T(\tilde{\theta}_{k+j}-\tilde{\theta}_k)\right)^2 \le \mathbb{E}\beta^2\bar{\varphi}^4j^2/r_{k+1}^2$, which indicates

$$\begin{aligned}\mathbb{E}(\varphi_{k+j}^T\tilde{\theta}_k)^2 \le& \mathbb{E}\left(\varphi_{k+j}^T\tilde{\theta}_{k+j}\right)^2 + \mathbb{E}\frac{\beta^2\bar{\varphi}^4j^2}{r_{k+1}^2}\\ &+2\left[\mathbb{E}\left(\varphi_{k+j}^T\tilde{\theta}_{k+j}\right)^2\right]^{\frac{1}{2}}\left[\mathbb{E}\frac{\beta^2\bar{\varphi}^4j^2}{r_{k+1}^2}\right]^{\frac{1}{2}}.\end{aligned} \tag{24}$$

From (23), (24) and $r_{k+1} \ge 1 + n\delta^2(k-N)$ we have

$$\lim_{k\to\infty}\mathbb{E}(\varphi_{k+j}^T\tilde{\theta}_k)^2 = 0, j = 1, 2, \ldots, N-1. \tag{25}$$

Hence, by Assumption 2.1 and (25), we get

$$\begin{aligned}\delta^2\mathbb{E}\|\tilde{\theta}_k\|^2 \le& \mathbb{E}\left[\tilde{\theta}_k^T\left(\sum_{j=0}^{N-1}\varphi_{k+j}\varphi_{k+j}^T\right)\tilde{\theta}_k\right]\\ =&\sum_{j=0}^{N-1}\mathbb{E}(\varphi_{k+j}^T\tilde{\theta}_k)^2 \to 0, k\to\infty.\end{aligned}$$

which implies

$$\lim_{k\to\infty}\mathbb{E}\|\tilde{\theta}_k\|^2 = 0.$$

On the other hand, $\mathbb{E}[\|\tilde{\theta}_{k+1}\|^2|\mathcal{F}_k] \le \|\tilde{\theta}_k\|^2 + \frac{\beta^2\|\varphi_k\|^2}{r_{k+1}^2}$, which together with Lemma A.2 and $\sum_{i=0}^{k}\frac{\|\varphi_i\|^2}{r_{i+1}^2} < \infty$ implies that $\tilde{\theta}_k$ converges a.s. to a bounded limit. Note that $\lim_{k\to\infty}\mathbb{E}\|\tilde{\theta}_k\|^2 = 0$. Then, there is a subsequence of $\tilde{\theta}_k$ that converges almost surely to 0. Consequently, $\tilde{\theta}_k$ almost surely converges to 0. □

Remark 4.1. In fact, the deterministic framework and the stochastic framework complement each other precisely in some situations. We demonstrate it by an example, consider the first-order system $y_{k+1} = \varphi_k\theta + d_{k+1}$, where $\Theta = (-5, 5)$, $\{d_k\}$ is i.i.d. and follows $U(-3, 3)$, $\{\varphi_k, \varphi_k \equiv 3\}$ satisfies Assumption 2.1. It can be calculated that $f_{\bar{\theta}} = 0$, making the results for the stochastic noise case are invalid. But by use of the results for the bounded noise case (Theorem 3), the estimation error of the algorithm (5) can be smaller than $\acute{\theta} = \frac{2}{3}$ after k_0 steps. What is more, from (20)–(21), the condition (4) in Assumption 2.4 could be broadened to $\inf_{x\in[-\bar{\varphi}\acute{\theta},\bar{\varphi}\acute{\theta}]} f(x) \triangleq f_{\acute{\theta}} > 0$ if $\|\tilde{\theta}_k\| \le \acute{\theta} \le \bar{\theta}$. Besides, the density function of the noises has a non-zero lower bound $f_{\acute{\theta}} = \frac{1}{6}$ on $[-\bar{\varphi}\acute{\theta}, \bar{\varphi}\acute{\theta}] = [-2, 2]$. Then the estimate given by the algorithm (5) could converge to the true parameter according to Theorem 4. In this example, the estimation task can be finished by the conjoint results of both the stochastic and bounded noise cases, while it cannot be completed by each deterministic framework or stochastic framework.

Then, the mean square convergence rate of the algorithm (5) is given in the following theorem with the bounded persistent excitation condition and stochastic noises.

Theorem 5. *For the FIR system (18) with stochastic noises, under the conditions of Theorem 4, the estimation error given by the algorithm (5) has the following property:*

$$\mathbb{E}\|\tilde{\theta}_k\|^2 = O\left(\frac{1}{k}\right),$$

provided that $\beta > \frac{\bar{\varphi}^2+1}{4\delta^2 f_{\bar{\theta}}}$ with $f_{\bar{\theta}}$ being defined in Assumption 2.4.

Proof. From (22) we can get

$$\begin{aligned}\mathbb{E}\|\tilde{\theta}_{k+1}\|^2 \le& \mathbb{E}\|\tilde{\theta}_k\|^2 - \frac{4\beta f_{\bar{\theta}}}{r_{k+1}}\mathbb{E}(\varphi_k^T\tilde{\theta}_k)^2 + \mathbb{E}\left[\frac{\beta^2\bar{\varphi}^2}{r_{k+1}^2}\right]\\ \le&\mathbb{E}\|\tilde{\theta}_{k-N+1}\|^2 - 4\beta f_{\bar{\theta}}\sum_{i=k-N+1}^{k}\mathbb{E}\left[\frac{\tilde{\theta}_i^T\varphi_i\varphi_i^T\tilde{\theta}_i}{r_{k+1}}\right]\\ &+\sum_{i=k-N+1}^{k}\mathbb{E}\left[\frac{\beta^2\bar{\varphi}^2}{r_{k+1}^2}\right].\end{aligned} \tag{26}$$

From (B.4), for $k \ge N$ we have

$$\begin{aligned}&-\sum_{i=k-N+1}^{k}\frac{\tilde{\theta}_i^T\varphi_i\varphi_i^T\tilde{\theta}_i}{r_{i+1}}\\ \le& -\frac{\delta^2N\|\tilde{\theta}_{k-N+1}\|^2}{r_{k+1}} + O\left(\frac{1}{(k-N+1)^2}\right).\end{aligned}$$

Thus, we can get

$$\begin{aligned}&-4\beta f_{\bar{\theta}}\sum_{i=k-N+1}^{k}\mathbb{E}\left[\frac{\tilde{\theta}_i^T\varphi_i\varphi_i^T\tilde{\theta}_i}{r_{i+1}}\right]\\ \le& -\frac{4\delta^2\beta f_{\bar{\theta}}N}{r_{k+1}}\mathbb{E}\|\tilde{\theta}_{k-N+1}\|^2 + O\left(\frac{1}{(k-N+1)^2}\right).\end{aligned} \tag{27}$$

Substituting (27) into (26) gives

$$\begin{aligned}\mathbb{E}\|\tilde{\theta}_{k+1}\|^2 \le& \left(1-\frac{4\delta^2\beta f_{\bar{\theta}}N}{r_{k+1}}\right)\mathbb{E}\|\tilde{\theta}_{k-N+1}\|^2 + O\left(\frac{1}{(k-N+1)^2}\right)\\ \le&\prod_{j=0}^{\left\lfloor\frac{k}{N}\right\rfloor-1}\left(1-\frac{4\delta^2\beta f_{\bar{\theta}}N}{r_{k+1-jN}}\right)\mathbb{E}\|\tilde{\theta}_{k+1-\left\lfloor\frac{k}{N}\right\rfloor N}\|^2\\ &+\sum_{j=1}^{\left\lfloor\frac{k}{N}\right\rfloor}\prod_{i=0}^{j-1}\left(1-\frac{4\delta^2\beta f_{\bar{\theta}}N}{r_{k+1-iN}}\right)O\left(\frac{1}{(k-jN+1)^2}\right).\end{aligned} \tag{28}$$

On one hand, the first item on the right side of (28) is

$$\prod_{j=0}^{\left\lfloor\frac{k}{N}\right\rfloor-1}\left(1-\frac{4\delta^2\beta f_{\bar{\theta}}N}{r_{k+1-jN}}\right) \le \prod_{m=\kappa+1}^{\left\lceil\frac{k}{N}\right\rceil}\left(1-\frac{4\delta^2\beta f_{\bar{\theta}}N}{r_{mN+1}}\right)$$

$$
\begin{aligned}
&\le \prod_{m=\kappa+1}^{\lceil \frac{k}{N} \rceil} \left(1 - \frac{4\delta^2 \beta f_{\bar{\theta}}}{1 + m\bar{\varphi}^2}\right) \\
&= O\left(\prod_{m=\kappa+1}^{\lceil \frac{k}{N} \rceil} \left(1 - \frac{\frac{4\delta^2\beta f_{\bar{\theta}}}{\bar{\varphi}^2+1}}{m}\right)\right), \qquad (29)
\end{aligned}
$$

where $\kappa = \lceil \frac{k}{N} \rceil - \lfloor \frac{k}{N} \rfloor$.

On the other hand, the second item on the right side of (28) is

$$
\begin{aligned}
&\sum_{j=1}^{\lfloor \frac{k}{N} \rfloor} \prod_{i=0}^{j-1} \left(1 - \frac{4\delta^2\beta f_{\bar{\theta}} N}{r_{k+1-iN}}\right) O\left(\frac{1}{(k - jN + 1)^2}\right) \\
&\le \sum_{j=1}^{\lfloor \frac{k}{N} \rfloor} \prod_{i=0}^{j-1} \left(1 - \frac{4\delta^2\beta f_{\bar{\theta}} N}{r_{k+1-iN}}\right) O\left(\frac{1}{(\lfloor \frac{k}{N} \rfloor - j + 1)^2 N^2}\right) \\
&\le O\left(\prod_{j=0}^{\lfloor \frac{k}{N} \rfloor - 1} \left(1 - \frac{4\delta^2\beta f_{\bar{\theta}} N}{r_{k+1-jN}}\right)\right) \\
&\quad + \sum_{m=1}^{\lfloor \frac{k}{N} \rfloor - 1} \prod_{p=\kappa+m+1}^{\lceil \frac{k}{N} \rceil} \left(1 - \frac{\frac{4\delta^2\beta f_{\bar{\theta}}}{\bar{\varphi}^2+1}}{p}\right) O\left(\frac{1}{m^2 N^2}\right) \\
&= O\left(\prod_{m=\kappa+1}^{\lceil \frac{k}{N} \rceil} \left(1 - \frac{\frac{4\delta^2\beta f_{\bar{\theta}}}{\bar{\varphi}^2+1}}{m}\right)\right) \\
&\quad + \sum_{m=1}^{\lfloor \frac{k}{N} \rfloor - 1} \prod_{p=\kappa+m+1}^{\lceil \frac{k}{N} \rceil} \left(1 - \frac{\frac{4\delta^2\beta f_{\bar{\theta}}}{\bar{\varphi}^2+1}}{p}\right) O\left(\frac{1}{m^2}\right). \qquad (30)
\end{aligned}
$$

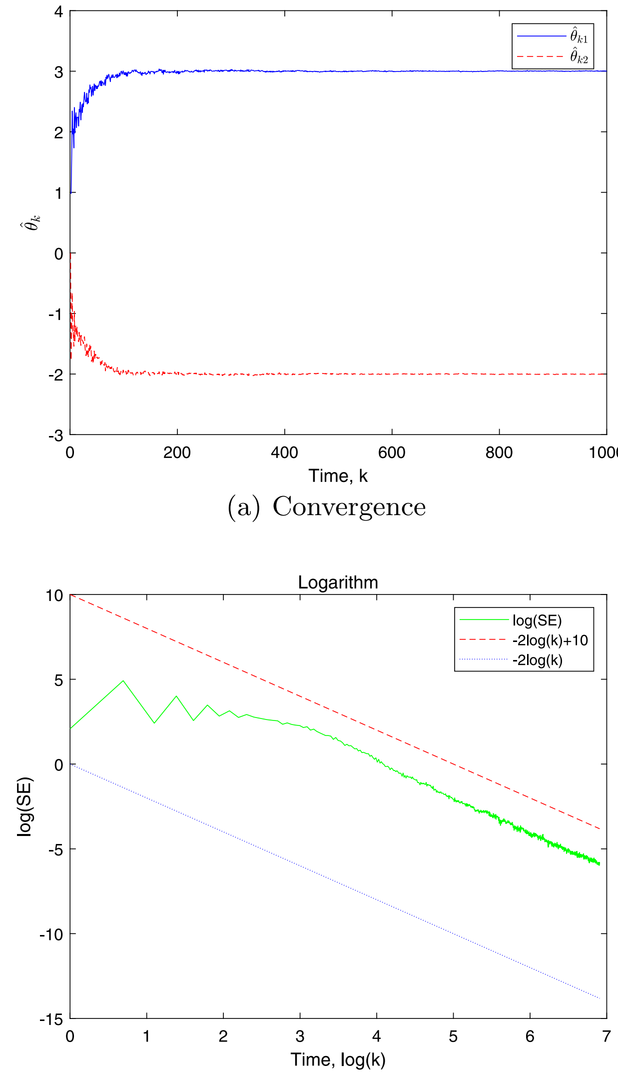
Thus, from Lemma A.3, substituting (29) and (30) into (28) yields $\mathbb{E}\|\tilde{\theta}_k\|^2 = O(\frac{1}{k})$ when $\frac{4\beta\delta^2 f_{\bar{\theta}}}{\bar{\varphi}^2+1} > 1$. □

Remark 4.2. Comparing with Guo and Zhao (2013), the mean square convergence rate is improved from $O\left(\frac{\ln k}{k}\right)$ to $O\left(\frac{1}{k}\right)$. Moreover, the algorithm (5) does not need stationary ergodicity of signals and can deal with more types of stochastic noises, contrasted with the sign-error type estimation algorithm (Chen & Yin, 2003). Moreover, this paper does not require the independent and identically distributed inputs compared to other identification works with time-varying binary observations (Csáji & Weyer, 2012; You, 2015; Zhao et al., 2017), which makes the results be applied to adaptive control problems. Besides, the results of this paper adapt to more types of noises. For example, Zhao et al. (2017) need the noises are i.i.d. random variables with symmetric, continuous, and bounded probability density function. Accordingly, the noise condition in Assumption 2.4 is more general.

5. Simulations

In this section, we will illustrate the theoretical results with three simulation examples.

Example 1: Consider a second-order noise-free FIR system: $y_{k+1} = \varphi_k^T\theta$, where y_{k+1} only can be measured by binary observation (2). The true parameter $\theta = [3, -2]^T$ is unknown, only can we learn $\theta \in \Theta = [-10, 10] \times [-10, 10]$. We apply the sign-error type recursive projection algorithm (5) to give the estimate $\hat{\theta}_k$. Then we choose the inputs $\varphi_k = [\varphi_k^1, \varphi_k^2]^T$ satisfying Assumption 2.1 and $\beta = 50$ satisfying the condition of Theorem 2, where $\varphi_k^i (i = 1, 2)$ are randomly generated in the



(a) Convergence

(b) Convergence rate

Fig. 1. The noise-free case.

interval [0.2, 0.7]. Here we repeat the simulation for 500 times to eliminate the effects of random inputs with same initial estimate $\hat{\theta}_0 = [1, 0]^T$.

The simulation results are given in Fig. 1. From it, we can see that the estimate converges to the true value, and the logarithm of the square error (SE) is linear with $2\log k$, which implies the convergence rate of this algorithm is close to $O(\frac{1}{k^2})$.

Example 2: Consider a second-order FIR system with bounded noise: $y_{k+1} = \varphi_k^T\theta + d_{k+1}$, whose output only can be measured by binary observation (2). The true parameter $\theta = [3, -2]^T$ is unknown, but we learn that $\theta \in \Theta = [-10, 10] \times [-10, 10]$. The system noise d_k is generated as

$$
d_{k+1} = \begin{cases} -\frac{1}{2} \cdot (-1)^k + \frac{1}{2^k}, & k = 3l, \\ -\frac{2}{3} \cdot (-1)^k - \frac{1}{3^k}, & k = 3l + 1, \\ \frac{1}{3} \cdot (-1)^k - \frac{1}{3^k}, & k = 3l + 2, \end{cases}
$$

where $l = 1, 2, 3, \ldots$. Parameter estimate $\hat{\theta}_k$ is given by the algorithm (5). The inputs are periodically generated by $\varphi_{2k+1} = [2, 0]^T$, $\varphi_{2k} = [0, -2]^T$, which satisfy Assumption 2.1. Then, we calculate the upper bound of estimation error by Theorem 3, which is $B_k = \frac{d}{\alpha\delta} + \frac{N\bar{\varphi}}{r_{k-N+1}} = \frac{55}{81} + \frac{4}{r_{k-1}}$. From Fig. 2, we can learn

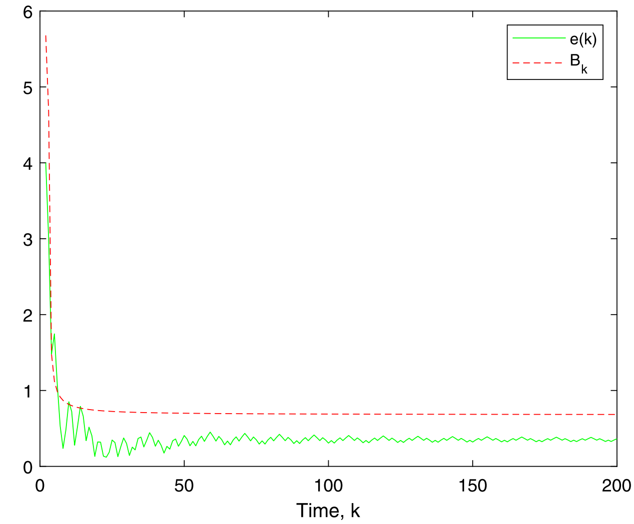


Fig. 2. The bounded noise case.

that the estimation error $e(k)(\triangleq \|\tilde{\theta}_k\|)$ finally undulates in $[0, B_k]$, which confirms the results of Theorem 3.

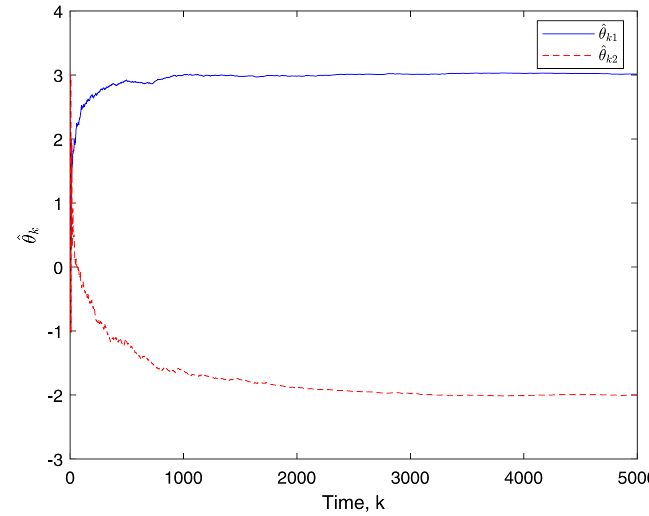
Example 3: Consider a second-order FIR system with stochastic noise: $y_{k+1} = \varphi_k^T\theta + d_{k+1}$, where the output only can be measured by binary observation (2). The true parameter $\theta = [3, -2]^T$ is unknown, only can we learn $\theta \in \Theta = [-4, 4] \times [-3, 3]$. The system noise d_{k+1} follows $N(0, 6^2)$. The inputs are periodically generated by $\varphi_{2k+1} = [-1, 0]^T$, $\varphi_{2k} = [0, 1]^T$, satisfying Assumption 2.1. The estimate $\hat{\theta}_k$ is given by the algorithm (5), where $\beta = 50$ satisfies Theorem 5. Here we repeat the simulation for 500 times to eliminate effects of stochastic noises with the same initial estimate $\hat{\theta}_0 = [1, -1]^T$.

The simulation results are given in Fig. 3. From it, we can learn the estimate converges to the true value, and the logarithm of the mean square error (MSE) is linear with the logarithm of the index k, which indicates the mean square convergence rate of estimation error is $O(\frac{1}{k})$.

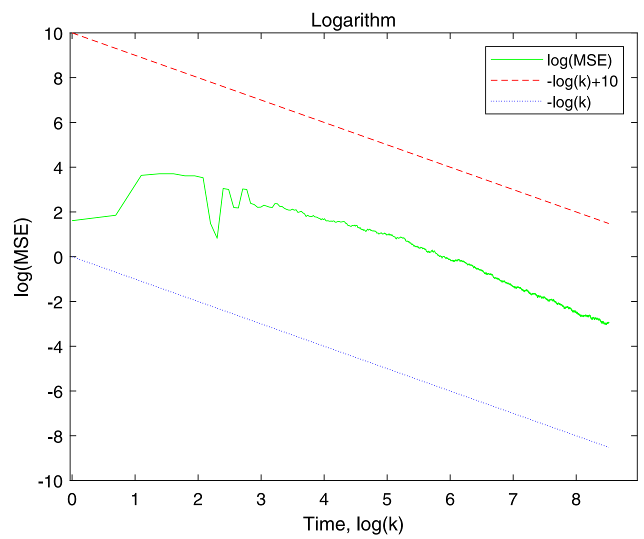
6. Conclusion

In this paper, we have investigated the sign-error type recursive projection algorithm, which could identify FIR systems with binary-valued observations under the bounded persistent excitation condition. It is worth mentioning that this is a unified algorithm, which not only could solve system identification problems under the deterministic framework but also can estimate the unknown parameter under the stochastic framework. Moreover, this is the first result of the sign-error type algorithm in a deterministic setting to the best of our knowledge. Under the deterministic framework, the convergence of the unified algorithm is established and the convergence rate is proved to be close to $O\left(\frac{1}{k^2}\right)$ for the noise-free case, the upper bound of the estimation error is obtained for the bounded noise case. Meanwhile, under the stochastic framework, the algorithm convergence is given in the sense of mean square and almost sure with i.i.d. stochastic noises. Besides, the mean square convergence rate of the estimation error is proved as $O(\frac{1}{k})$.

There are many meaningful topics for future works, for example, whether the unified algorithm is suitable for more complex systems with binary-valued observations, such as ARMAX systems? How do we remove the assumption about the boundness of inputs?



(a) Convergence



(b) Convergence rate

Fig. 3. The stochastic noise case.

Appendix A. Useful tools

In this part, some lemmas are collected and established, which are frequently used in the analysis of convergence and convergence rate.

Lemma A.1. *If Assumption 2.1 holds, then for any positive integer p we have*

$$\|\tilde{\theta}_{k+p} - \tilde{\theta}_k\| \le \sum_{i=k}^{k+p-1} \frac{\beta\bar{\varphi}}{r_{i+1}} \le \frac{p\beta\bar{\varphi}}{r_{k+1}}.$$

The proof is similar to Lemma 8 in Guo and Zhao (2013), so the detail proof is omitted.

Lemma A.2. *(Chen, 2002) Let $(v_k, \mathcal{F}_k)$, $(w_k, \mathcal{F}_k)$ be two nonnegative adapted sequences. If $\mathbb{E}(v_{k+1}|\mathcal{F}_k) \le v_k + w_k$ and $\mathbb{E}\sum_{i=1}^{\infty} w_k < \infty$, then v_k converges a.s. to a finite limit.*

Lemma A.3. *(Zhao et al., 2019) For any given $a \in \mathbb{R}$, we have the assertions:*

(i)

$$\prod_{i=1}^{k}\left(1-\frac{a}{i}\right)=O\left(\frac{1}{k^a}\right),$$

(ii)

$$\sum_{l=1}^{k}\prod_{i=l+1}^{k}\left(1-\frac{a}{i}\right)\frac{1}{l^2}=\begin{cases}O\left(\frac{1}{k^a}\right), & a<1,\\ O\left(\frac{\ln k}{k}\right), & a=1,\\ O\left(\frac{1}{k}\right), & a>1.\end{cases}$$

Lemma A.4. *For any given $a\in(0,1)$, we have*

$$\sum_{l=1}^{k}\prod_{i=l}^{k}\left(1-\frac{1}{i^a}\right)\frac{1}{l^2}=O\left(\frac{1}{k^{2-a}}\right).$$

Proof. Note that

$$\prod_{i=l}^{k}\left(1-\frac{1}{i^a}\right)=e^{\sum_{i=l}^{k}\ln(1-\frac{1}{i^a})}=O\left(e^{-\sum_{i=l}^{k}\frac{1}{i^a}}\right)$$
$$=O\left(e^{\frac{1}{1-a}(l^{1-a}-k^{1-\alpha})}\right).$$

Then,

$$\sum_{l=1}^{k}\prod_{i=l}^{k}\left(1-\frac{1}{i^a}\right)\frac{1}{l^2}=O\left(\frac{\sum_{l=1}^{k}\frac{e^{\frac{1}{1-a}l^{1-a}}}{l^2}}{e^{\frac{1}{1-a}k^{1-a}}}\right)$$
$$=O\left(\frac{\frac{e^{\frac{1}{1-a}k^{1-\alpha}}}{k^2}}{k^{-a}e^{\frac{1}{1-a}k^{1-a}}}\right)=O\left(\frac{1}{k^{2-a}}\right).$$

Hence, the conclusion is established. □

Appendix B. Proof of Theorem 2

The proof can be divided into the following steps.

Step 1: To prove $\|\tilde{\theta}_k\|^2=O\left(\frac{1}{k}\right)$.

From Theorem 1 we know $\lim_{k\to\infty}|\varphi_k^T\tilde{\theta}_k|=0$. So, there exists $\kappa_1=k_1N>0$ such that when $k>\kappa_1$, $|\varphi_k^T\tilde{\theta}_k|\geq|\varphi_k^T\tilde{\theta}_k|^2$. Thus, by (11), when $k>\kappa_1$, we have

$$\begin{aligned}\|\tilde{\theta}_{k+1}\|^2\leq&\|\tilde{\theta}_k\|^2-\frac{2\beta|\varphi_k^T\tilde{\theta}_k|^2}{r_{k+1}}+\frac{\beta^2\|\varphi_k\|^2}{r_{k+1}^2}\\ \leq&\|\tilde{\theta}_{k-N+1}\|^2-\sum_{i=k-N+1}^{k}\frac{2\beta\tilde{\theta}_i^T\varphi_i\varphi_i^T\tilde{\theta}_i}{r_{i+1}}\\ &+\sum_{i=k-N+1}^{k}\frac{\beta^2\|\varphi_i\|^2}{r_{i+1}^2}.\end{aligned}\tag{B.1}$$

From Assumption 2.2, we have

$$\|\tilde{\theta}_k\|=\|\hat{\theta}_k-\theta\|\leq\|\hat{\theta}_k\|+\|\theta\|\leq\bar{\theta},\quad\forall k\geq 0.\tag{B.2}$$

The above result can be also obtained without Assumption 2.2, since $\tilde{\theta}_k$ converges to 0.

Noticing

$$\begin{aligned}\tilde{\theta}_i^T\varphi_i\varphi_i^T\tilde{\theta}_i=&\tilde{\theta}_{k-N+1}^T\varphi_i\varphi_i^T\tilde{\theta}_{k-N+1}+((\tilde{\theta}_i-\tilde{\theta}_{k-N+1})^T\varphi_i)^2\\ &+2(\tilde{\theta}_i-\tilde{\theta}_{k-N+1})^T\varphi_i\varphi_i^T\tilde{\theta}_{k-N+1}\\ \geq&\tilde{\theta}_{k-N+1}^T\varphi_i\varphi_i^T\tilde{\theta}_{k-N+1}\\ &+2(\tilde{\theta}_i-\tilde{\theta}_{k-N+1})^T\varphi_i\varphi_i^T\tilde{\theta}_{k-N+1},\end{aligned}\tag{B.3}$$

and from (14), (B.2) and Lemma A.1, it follows that,

$$\begin{aligned}&-\sum_{i=k-N+1}^{k}\frac{\tilde{\theta}_i^T\varphi_i\varphi_i^T\tilde{\theta}_i}{r_{i+1}}\\ \leq&-\sum_{i=k-N+1}^{k}\frac{\tilde{\theta}_{k-N+1}^T\varphi_i\varphi_i^T\tilde{\theta}_{k-N+1}}{r_{i+1}}\\ &+2\bar{\varphi}^2\bar{\theta}\sum_{i=k-N+1}^{k}\frac{\|\tilde{\theta}_i-\tilde{\theta}_{k-N+1}\|}{r_{i+1}}\\ \leq&-\frac{\delta^2N\|\tilde{\theta}_{k-N+1}\|^2}{r_{k+1}}+2\bar{\varphi}^2\bar{\theta}\sum_{i=k-N+1}^{k}\frac{\beta N\bar{\varphi}}{r_{k-N+2}^2}\\ \leq&-\frac{\delta^2N\|\tilde{\theta}_{k-N+1}\|^2}{r_{k+1}}+O\left(\frac{1}{(k-N+1)^2}\right),\end{aligned}\tag{B.4}$$

for $k\geq N$. Substituting (B.4) into (B.1) gives

$$\begin{aligned}&\|\tilde{\theta}_{k+1}\|^2\\ \leq&\left(1-\frac{2\delta^2\beta N}{r_{k+1}}\right)\|\tilde{\theta}_{k-N+1}\|^2+O\left(\frac{1}{(k-N+1)^2}\right)\\ \leq&\prod_{j=0}^{\lfloor\frac{k}{N}\rfloor-k_1-1}\left(1-\frac{2\delta^2\beta N}{r_{k+1-jN}}\right)\left\|\tilde{\theta}_{k+1-\lfloor\frac{k-\kappa_1}{N}\rfloor N}\right\|^2\\ &+\sum_{j=1}^{\lfloor\frac{k}{N}\rfloor-k_1}\prod_{i=0}^{j-1}\left(1-\frac{2\delta^2\beta N}{r_{k+1-iN}}\right)O\left(\frac{1}{(k-jN+1)^2}\right).\end{aligned}\tag{B.5}$$

For brevity, let $\kappa=\lceil\frac{k}{N}\rceil-\lfloor\frac{k}{N}\rfloor$. On one hand, the first item on the right side of inequality (B.5) is

$$\begin{aligned}&\prod_{j=0}^{\lfloor\frac{k}{N}\rfloor-k_1-1}\left(1-\frac{2\delta^2\beta N}{r_{k+1-jN}}\right)\\ \leq&\prod_{m=\kappa+1}^{\lceil\frac{k}{N}\rceil-k_1}\left(1-\frac{2\delta^2\beta N}{r_{(m+k_1)N+1}}\right)\\ \leq&\prod_{m=\kappa+1}^{\lceil\frac{k}{N}\rceil-k_1}\left(1-\frac{2\delta^2\beta N}{1+(m+k_1)N\bar{\varphi}^2}\right)\\ \leq&\prod_{m=\kappa+1}^{\lceil\frac{k}{N}\rceil-k_1}\left(1-\frac{2\delta^2\beta}{(m+k_1)(\bar{\varphi}^2+1)}\right)\\ =&O\left(\prod_{m=\kappa+1}^{\lceil\frac{k}{N}\rceil-k_1}\left(1-\frac{\frac{2\delta^2\beta}{\bar{\varphi}^2+1}}{m+k_1}\right)\right).\end{aligned}\tag{B.6}$$

On the other hand, the second item on the right side of inequality (B.6) is

$$\begin{aligned}&\sum_{j=1}^{\lfloor\frac{k}{N}\rfloor-k_1}\prod_{i=0}^{j-1}\left(1-\frac{2\delta^2\beta N}{r_{k+1-iN}}\right)O\left(\frac{1}{(k-jN+1)^2}\right)\\ \leq&\sum_{j=1}^{\lfloor\frac{k}{N}\rfloor-k_1}\prod_{i=0}^{j-1}\left(1-\frac{2\delta^2\beta N}{r_{k+1-iN}}\right)O\left(\frac{1}{(\lfloor\frac{k}{N}\rfloor-j+1)^2N^2}\right)\end{aligned}$$

$$\le O\left(\prod_{j=0}^{\lfloor \frac{k}{N}\rfloor - k_1 - 1}\left(1-\frac{2\delta^2\beta N}{r_{k+1-jN}}\right)\right)$$

$$+\sum_{j=1}^{\lfloor \frac{k}{N}\rfloor - k_1 - 1}\prod_{i=0}^{j-1}\left(1-\frac{2\delta^2\beta N}{r_{k+1-iN}}\right)O\left(\frac{1}{(\lfloor \frac{k}{N}\rfloor - j+1)^2 N^2}\right)$$

$$\le O\left(\prod_{m=\kappa+1}^{\lceil \frac{k}{N}\rceil - k_1}\left(1-\frac{\frac{2\delta^2\beta}{\bar\varphi^2+1}}{m+k_1}\right)\right)$$

$$+\sum_{m=1}^{\lfloor \frac{k}{N}\rfloor - k_1 - 1}\prod_{p=\kappa+m+1}^{\lceil \frac{k}{N}\rceil - k_1}\left(1-\frac{\frac{2\delta^2\beta}{\bar\varphi^2+1}}{p+k_1}\right)O\left(\frac{1}{(m+k_1)^2N^2}\right)$$

$$\le O\left(\prod_{m=k_1+\kappa+1}^{\lceil \frac{k}{N}\rceil}\left(1-\frac{\frac{2\delta^2\beta}{\bar\varphi^2+1}}{m}\right)\right)$$

$$+\sum_{m=k_1}^{\lfloor \frac{k}{N}\rfloor - 1}\prod_{p=\kappa+m+1}^{\lceil \frac{k}{N}\rceil}\left(1-\frac{\frac{2\delta^2\beta}{\bar\varphi^2+1}}{p}\right)O\left(\frac{1}{m^2}\right). \quad \text{(B.7)}$$

As above, from (B.5)–(B.7) and Lemma A.3, when $\frac{2\delta^2\beta}{\bar\varphi^2+1} > 1$, we obtain $\|\tilde\theta_k\|^2 = O(\frac{1}{k})$.

Step 2: Prove $\|\tilde\theta_k\|^2 = O\left(\frac{1}{k^{1+t_1}}\right)$ using $\|\tilde\theta_k\|^2 = O\left(\frac{1}{k}\right)$, where $t_1 = \frac{r-1}{2} > 0$.

From $\|\tilde\theta_k\|^2 = O(\frac{1}{k})$, $\|\varphi_k\| \le \bar\varphi$ and $1+n\delta^2(k-N) \le r_{k+1} \le k\bar\varphi^2+1$, there exist $\kappa_2 = k_2N$ and $t_1 = \frac{r-1}{2} \in (0, \frac{1}{2})$ such that $|\varphi_k^T\tilde\theta_k| \ge r_{k+1}^{t_1}|\varphi_k^T\tilde\theta_k|^2$ and $\|\tilde\theta_k\| < \frac{1}{r_{k+1}^{t_1}}$ for $k > \kappa_2$. Thus, when $k > \kappa_2$, we have

$$\begin{aligned}\|\tilde\theta_{k+1}\|^2 \le & \|\tilde\theta_k\|^2 - \frac{2\beta|\varphi_k^T\tilde\theta_k|^2}{r_{k+1}^{1-t_1}} + \frac{\beta^2\|\varphi_k\|^2}{r_{k+1}^2}\\ \le & \|\tilde\theta_{k-N+1}\|^2 - \sum_{i=k-N+1}^{k}\frac{2\beta\tilde\theta_i^T\varphi_i\varphi_i^T\tilde\theta_i}{r_{i+1}^{1-t_1}}\\ & + \sum_{i=k-N+1}^{k}\frac{\beta^2\|\varphi_i\|^2}{r_{i+1}^2}. \end{aligned} \quad \text{(B.8)}$$

By (B.3) and $\|\tilde\theta_k\| < \frac{1}{r_{k+1}^{t_1}}$, similarly to (B.4) we get

$$\begin{aligned} & -\sum_{i=k-N+1}^{k}\frac{\tilde\theta_i^T\varphi_i\varphi_i^T\tilde\theta_i}{r_{i+1}^{1-t_1}}\\ \le & -\sum_{i=k-N+1}^{k}\frac{\tilde\theta_{k-N+1}^T\varphi_i\varphi_i^T\tilde\theta_{k-N+1}}{r_{i+1}^{1-t_1}}\\ & -\sum_{i=k-N+1}^{k}\frac{2(\tilde\theta_i-\tilde\theta_{k-N+1})^T\varphi_i\varphi_i^T\tilde\theta_{k-N+1}}{r_{i+1}^{1-t_1}}\\ \le & -\frac{\delta^2N\|\tilde\theta_{k-N+1}\|^2}{r_{k+1}^{1-t_1}} + O\left(\frac{1}{(k-N+1)^2}\right). \end{aligned} \quad \text{(B.9)}$$

Substituting (B.9) into (B.8) gives

$$\|\tilde\theta_{k+1}\|^2 \le \prod_{j=0}^{\lfloor \frac{k}{N}\rfloor - k_2 - 1}\left(1-\frac{2\delta^2\beta N}{r_{k+1-jN}^{1-t_1}}\right)\left\|\tilde\theta_{k+1-\lfloor \frac{k-\kappa_2}{N}\rfloor N}\right\|^2$$

$$+\sum_{j=1}^{\lfloor \frac{k}{N}\rfloor - k_2}\prod_{i=0}^{j-1}\left(1-\frac{2\delta^2\beta N}{r_{k+1-iN}^{1-t_1}}\right)O\left(\frac{1}{(k-jN+1)^2}\right). \quad \text{(B.10)}$$

Similarly to (B.6) and (B.7), we have

$$\prod_{j=0}^{\lfloor \frac{k}{N}\rfloor - k_2 - 1}\left(1-\frac{2\delta^2\beta N}{r_{k+1-jN}^{1-t_1}}\right)$$

$$= O\left(\prod_{m=\kappa+1}^{\lceil \frac{k}{N}\rceil - k_2}\left(1-\frac{\frac{2\delta^2\beta N^{t_1}}{(\bar\varphi^2+1)^{1-t_1}}}{(m+k_2)^{1-t_1}}\right)\right), \quad \text{(B.11)}$$

where $\kappa = \lceil \frac{k}{N}\rceil - \lfloor \frac{k}{N}\rfloor$, and

$$\sum_{j=1}^{\lfloor \frac{k}{N}\rfloor - k_2}\prod_{i=0}^{j-1}\left(1-\frac{2\delta^2\beta N}{r_{k+1-iN}^{1-t_1}}\right)O\left(\frac{1}{(k-jN+1)^2}\right)$$

$$= O\left(\prod_{m=k_2+\kappa+1}^{\lceil \frac{k}{N}\rceil}\left(1-\frac{\frac{2\delta^2\beta N^{t_1}}{(\bar\varphi^2+1)^{1-t_1}}}{m^{1-t_1}}\right)\right)$$

$$+\sum_{m=k_2}^{\lfloor \frac{k}{N}\rfloor - 1}\prod_{p=\kappa+m+1}^{\lceil \frac{k}{N}\rceil}\left(1-\frac{\frac{2\delta^2\beta N^{t_1}}{(\bar\varphi^2+1)^{1-t_1}}}{p^{1-t_1}}\right)O\left(\frac{1}{m^2}\right). \quad \text{(B.12)}$$

As above, by $\frac{2\delta^2\beta N^{t_1}}{(\bar\varphi^2+1)^{1-t_1}} > \frac{2\delta^2\beta}{\bar\varphi^2+1} > 1$, Lemma A.4 and (B.10)–(B.12), we obtain

$$\|\tilde\theta_k\|^2 = O\left(\frac{1}{k^{1+t_1}}\right).$$

Step 3: Prove $\|\tilde\theta_k\|^2 = O\left(\frac{1}{k^{1+t_2}}\right)$, where $t_2 = t_1 + \frac{1}{2}t_1 = \frac{3(r-1)}{4}$. Similarly to Step 2, we can prove that by $\|\tilde\theta_k\|^2 = O\left(\frac{1}{k^{1+t_1}}\right)$.

⋮

Step m + 1: Prove $\|\tilde\theta_k\|^2 = O\left(\frac{1}{k^{1+t_m}}\right)$, where $t_m = t_1 + \frac{1}{2}t_1 + \cdots + \frac{1}{2^{m-1}}t_1 = \frac{2^m-1}{2^m}(r-1)$. Similarly to Step 2, it also can be proved by using $\|\tilde\theta_k\|^2 = O\left(\frac{1}{k^{1+t_{m-1}}}\right)$.

By the definition of t_m, $\lim_{n\to\infty} t_m = r-1$. Repeating the above process, we can get $\|\tilde\theta_k\|^2 = O\left(\frac{1}{k^r}\right)$. □

Appendix C. Proof of Theorem 3

Let $u_k = \varphi_k^T\tilde\theta_k$. For $\alpha \in \left(0, \frac{1}{\sqrt{N^3}}\right)$, let $c > \frac{1}{1-N^3\alpha^2} > 1$. Then, from Assumption 2.1, Lemma A.1 and $\beta = 1$, we have

$$\begin{aligned} u_k^2 + u_{k+1}^2 + \cdots + u_{k+N-1}^2 = & \sum_{l=k}^{k+N-1}\tilde\theta_l^T\varphi_l\varphi_l^T\tilde\theta_l\\ \ge & \tilde\theta_k^T\left(\sum_{l=k}^{k+N-1}\varphi_l\varphi_l^T\right)\tilde\theta_k + \sum_{l=k}^{k+N-1}2(\tilde\theta_l-\tilde\theta_k)^T\varphi_l\varphi_l^T\tilde\theta_k\\ & + \sum_{l=k}^{k+N-1}(\tilde\theta_l-\tilde\theta_k)^T\varphi_l\varphi_l^T(\tilde\theta_l-\tilde\theta_k) \end{aligned}$$

$$\geq \left(1-\frac{1}{c}\right)\tilde{\theta}_k^T\left(\sum_{l=k}^{k+N-1}\varphi_l\varphi_l^T\right)\tilde{\theta}_k$$

$$-\ (c-1)\sum_{l=k}^{k+N-1}(\tilde{\theta}_l-\tilde{\theta}_k)^T\varphi_l\varphi_l^T(\tilde{\theta}_l-\tilde{\theta}_k)$$

$$\geq \left(1-\frac{1}{c}\right)\delta^2\|\tilde{\theta}_k\|^2-(c-1)\frac{\gamma\bar{\varphi}^4}{r_{k+1}^2}.$$

where $\gamma = \sum_{l=1}^{N-1} l^2$. Hence, from $r_{k+1} \geq 1 + n\delta^2(k-N)$, when $k > N + \sqrt{\frac{(c-1)\gamma}{1-\frac{1}{c}-N^3\alpha^2}\frac{\alpha\bar{\varphi}^2}{n\delta^2}}$ and $\|\tilde{\theta}_k\| > \frac{\bar{d}}{\alpha\delta}$, we get

$$\frac{1}{N}\sum_{l=k}^{k+N-1}u_l^2 \geq \left(1-\frac{1}{c}\right)\frac{\delta^2}{N}\|\tilde{\theta}_k\|^2-(c-1)\frac{\gamma\bar{\varphi}^4}{Nr_{k+1}^2}$$

$$\geq (N\bar{d})^2. \tag{C.1}$$

And by (5), $\beta = 1$ and $u_k = \varphi_k^T\tilde{\theta}_k$, we have

$$\|\tilde{\theta}_{k+1}\|^2 \leq \|\tilde{\theta}_k\|^2 - \frac{2\varphi_k^T\tilde{\theta}_k\,\text{sign}\{\varphi_k^T\tilde{\theta}_k-d_{k+1}\}}{r_{k+1}}+\frac{\|\varphi_k\|^2}{r_{k+1}^2}$$

$$\leq \|\tilde{\theta}_k\|^2 - \frac{2u_k\,\text{sign}\{u_k-d_{k+1}\}}{r_{k+1}}+\frac{\bar{\varphi}^2}{r_{k+1}^2}. \tag{C.2}$$

Define $t = \max\{N + \sqrt{\frac{(c-1)\gamma}{1-\frac{1}{c}-N^3\alpha^2}\frac{\alpha\bar{\varphi}^2}{n\delta^2}}, N + \frac{2N^2\bar{\varphi}^2}{n\delta^2}+\frac{2N\bar{\varphi}^2}{n\bar{d}\delta^2}+\frac{N^2\bar{\varphi}^3}{n\alpha\delta^3}\}$, where $c > \frac{1}{1-N^3\alpha^2}$.

Then we divide the proof into two steps.

Step 1: Prove there exists $k_0 > t$, such that $\|\tilde{\theta}_{k_0}\| \leq \frac{\bar{d}}{\alpha\delta}$.

Without loss of generality, we suppose $\|\tilde{\theta}_l\| > \frac{\bar{d}}{\alpha\delta}$ for all $l \in L_t \triangleq \{t, t+1, \ldots, t+N-1\}$, otherwise the conclusion is obvious. By $\|\tilde{\theta}_t\| > \frac{\bar{d}}{\alpha\delta}$, from (C.1) and Proposition 1 we learn

$$\sum_{|u_l|>\bar{d},l\in L_t}|u_l| > N\bar{d} \geq (N-1)\bar{d} \geq \sum_{|u_l|\leq\bar{d},l\in L_t}|u_l|.$$

From (C.2), we can get

(1) when $|u_l| \leq \bar{d}$, we have $\|\tilde{\theta}_{l+1}\|^2 \leq \|\tilde{\theta}_l\|^2 + \frac{2|u_l|}{r_{l+1}}+\frac{\|\varphi_l\|^2}{r_{l+1}^2}$;

(2) when $|u_l| > \bar{d}$, we have $\|\tilde{\theta}_{l+1}\|^2 \leq \|\tilde{\theta}_l\|^2 - \frac{2|u_l|}{r_{l+1}}+\frac{\|\varphi_l\|^2}{r_{l+1}^2}$.

Therefore,

$$\|\tilde{\theta}_{t+N}\|^2$$

$$\leq \|\tilde{\theta}_t\|^2 + \sum_{l=t}^{t+N-1}\frac{\|\varphi_l\|^2}{r_{l+1}^2} - \sum_{|u_l|>\bar{d},l\in L_t}\frac{2|u_l|}{r_{l+1}}+\sum_{|u_l|\leq\bar{d},l\in L_t}\frac{2|u_l|}{r_{l+1}}$$

$$\leq \|\tilde{\theta}_t\|^2 + \sum_{l=t}^{t+N-1}\frac{\|\varphi_l\|^2}{r_{l+1}r_l} - \frac{2N\bar{d}}{r_{t+N}}+\frac{2(N-1)\bar{d}}{r_{t+1}}$$

$$\leq \|\tilde{\theta}_t\|^2 + \sum_{l=t}^{t+N-1}\left(\frac{1}{r_l}-\frac{1}{r_{l+1}}\right) - \frac{2N\bar{d}}{r_{t+N}}+\frac{2(N-1)\bar{d}}{r_{t+1}}$$

$$\leq \|\tilde{\theta}_t\|^2 - \frac{2N\bar{d}-\frac{r_{t+N}}{r_{t+1}}2(N-1)\bar{d}-\frac{r_{t+N}-r_t}{r_t}}{r_{t+N}}$$

$$\leq \|\tilde{\theta}_t\|^2 - \frac{2\bar{d}-\frac{2(N-1)^2\bar{\varphi}^2\bar{d}}{r_{t+1}}-\frac{N\bar{\varphi}^2}{r_t}}{r_{t+N}} \leq \|\tilde{\theta}_t\|^2 - \frac{\bar{d}}{r_{t+N}},$$

where the last inequality is got by $r_{t+1} \geq 1 + n(t-N)\delta^2$ and $t > N + \frac{2N^2\bar{\varphi}^2}{n\delta^2}+\frac{2N\bar{\varphi}^2}{n\bar{d}\delta^2}+\frac{N^2\bar{\varphi}^3}{n\alpha\delta^3} > 2N + \frac{2N^2\bar{\varphi}^2}{n\delta^2}+\frac{N\bar{\varphi}^2}{n\bar{d}\delta^2}$.

For $r_{t+1} \geq 1+n\delta^2(t-N)$ and harmonic progression $\sum_{l=1}^{\infty}\frac{1}{l} = \infty$, there exists $k_0 \geq t$ such that $\|\tilde{\theta}_{k_0}\| \leq \frac{\bar{d}}{\alpha\delta}$.

Step 2: Show $\|\tilde{\theta}_{k+N}\| \leq \frac{\bar{d}}{\alpha\delta}+\frac{N\bar{\varphi}}{r_{k+1}}$ for all $k \geq k_0$.

From (5), (18), $\beta = 1$ and Remark 2.3, we can get

$$\|\tilde{\theta}_{k+1}\| = \|\hat{\theta}_{k+1}-\theta\| \leq \left\|\tilde{\theta}_k+\frac{\varphi_k}{r_{k+1}}s_{k+1}\right\|$$

$$\leq \|\tilde{\theta}_k\| + \frac{\|\varphi_k\|}{r_{k+1}}.$$

Hence, for all $l \in \{k_0+1,\ldots,k_0+N\}$, using $\|\tilde{\theta}_{k_0}\| \leq \frac{\bar{d}}{\alpha\delta}$ we get

$$\|\tilde{\theta}_l\| \leq \|\tilde{\theta}_{k_0}\| + \sum_{i=k_0}^{l-1}\frac{\|\varphi_i\|}{r_{i+1}}$$

$$\leq \|\tilde{\theta}_{k_0}\| + \frac{(l-k_0)\bar{\varphi}}{r_{k_0+1}} \leq \frac{\bar{d}}{\alpha\delta}+\frac{N\bar{\varphi}}{r_{l-N+1}}. \tag{C.3}$$

Assume $\|\tilde{\theta}_{k_1}\| \leq \frac{\bar{d}}{\alpha\delta}+\frac{N\bar{\varphi}}{r_{k_1-N+1}}$ when $k = k_1 \geq k_0+N$. Then, we just need to prove $\|\tilde{\theta}_{k_1+N}\| \leq \frac{\bar{d}}{\alpha\delta}+\frac{N\bar{\varphi}}{r_{k_1+1}}$. We need to consider two cases.

One is $\|\tilde{\theta}_{k_1}\| \leq \frac{\bar{d}}{\alpha\delta}$. In this case, we can prove $\|\tilde{\theta}_{k_1+N}\| \leq \frac{\bar{d}}{\alpha\delta}+\frac{N\bar{\varphi}}{r_{k_1+1}}$ similarly to (C.3).

The other case is $\|\tilde{\theta}_{k_1}\| > \frac{\bar{d}}{\alpha\delta}$. In this case, by (C.1) and $k_1 > k_0 \geq t > N + \sqrt{\frac{(c-1)\gamma}{1-\frac{1}{c}-N^3\alpha^2}\frac{\alpha\bar{\varphi}^2}{n\delta^2}}$, we have

$$\frac{1}{N}\sum_{l=k_1}^{k_1+N-1}u_l^2 \geq (N\bar{d})^2.$$

Similarly to Step 1, we can learn

$$\|\tilde{\theta}_{k_1+N}\|^2 \leq \|\tilde{\theta}_{k_1}\|^2 + \sum_{l=k_1}^{k_1+N-1}\frac{\|\varphi_l\|^2}{r_{l+1}^2}$$

$$-\sum_{|u_l|>\bar{d},l\in L_{k_1}}\frac{2|u_l|}{r_{l+1}}+\sum_{|u_l|\leq\bar{d},l\in L_{k_1}}\frac{2|u_l|}{r_{l+1}}$$

$$\leq \|\tilde{\theta}_{k_1}\|^2 + \sum_{l=k_1}^{k_1+N-1}\left(\frac{1}{r_l}-\frac{1}{r_{l+1}}\right)-\frac{2N\bar{d}}{r_{k_1+N}}+\frac{2(N-1)\bar{d}}{r_{k_1+1}}$$

$$\leq \left(\frac{\bar{d}}{\alpha\delta}+\frac{N\bar{\varphi}}{r_{k_1-N+1}}\right)^2+\frac{N\bar{\varphi}^2}{r_{k_1}r_{k_1+N}}-\frac{2N\bar{d}}{r_{k_1+N}}+\frac{2(N-1)\bar{d}}{r_{k_1+1}},$$

where $L_{k_1} \triangleq \{k_1, k_1+1, \ldots, k_1+N-1\}$.

To prove $\|\tilde{\theta}_{k_1+N}\| \leq \frac{\bar{d}}{\alpha\delta}+\frac{N\bar{\varphi}}{r_{k_1+1}}$, we only need to prove

$$\left(\frac{\bar{d}}{\alpha\delta}+\frac{N\bar{\varphi}}{r_{k_1-N+1}}\right)^2+\frac{N\bar{\varphi}^2}{r_{k_1}r_{k_1+N}}-\frac{2N\bar{d}}{r_{k_1+N}}+\frac{2(N-1)\bar{d}}{r_{k_1+1}}$$

$$\leq \left(\frac{\bar{d}}{\alpha\delta}+\frac{N\bar{\varphi}}{r_{k_1+1}}\right)^2. \tag{C.4}$$

By $k_1 > k_0 \geq t \geq N + \frac{2N^2\bar{\varphi}^2}{n\delta^2}+\frac{N\bar{\varphi}^2}{n\bar{d}\delta^2}+\frac{2N^2\bar{\varphi}^3}{n\alpha\delta^3}$, we have $r_{k_1+1} \geq 2N^2\bar{\varphi}^2+\frac{2N\bar{\varphi}^2}{\bar{d}}+\frac{N^2\bar{\varphi}^3}{\alpha\delta}$, $r_{k_1-N+1} \geq N^2\bar{\varphi}^2 \geq N\bar{\varphi}^2$, and then

$$\frac{\bar{d}r_{k_1+1}-(N-1)^2\bar{\varphi}^2\bar{d}}{2N\bar{\varphi}^2+\frac{N^2\bar{\varphi}^3\bar{d}}{\alpha\delta}} \geq \frac{r_{k_1+1}+(N-1)\bar{\varphi}^2}{r_{k_1+1}-N\bar{\varphi}^2} \geq \frac{r_{k_1+N}}{r_{k_1-N+1}}.$$

By $r_{k_1-N+1} \geq N^2\bar{\varphi}^2$, we have $\frac{2N^3\bar{\varphi}^4}{r_{k_1-N+1}} \leq 3N\bar{\varphi}^2$. And then, by $\frac{r_{k_1-N+1}+r_{k_1+1}}{r_{k_1+1}} \leq 2$ and $r_{k_1+1} - r_{k_1-N+1} \leq N\bar{\varphi}^2$, we have

$$N^2\bar{\varphi}^2\left(\frac{1}{r^2_{k_1-N+1}} - \frac{1}{r^2_{k_1+1}}\right) + \frac{N\bar{\varphi}^2 + \frac{2N^2\bar{\varphi}^3\bar{d}}{\alpha\delta}}{r_{k_1+1}r_{k_1-N+1}}$$

$$\leq \frac{2\bar{d}r_{k_1+1} - 2(N-1)^2\bar{\varphi}^2\bar{d}}{r_{k_1+1}r_{k_1+N}}.$$

Thus, by $r_{k_1+1}r_{k_1-N+1} \leq r_{k_1}r_{k_1+N}$, we get (C.4), and then,

$$\|\tilde{\theta}_{k_1+N}\| \leq \frac{\bar{d}}{\alpha\delta} + \frac{N\bar{\varphi}}{r_{k_1+1}}.$$

Therefore, by mathematical induction, we have

$$\|\tilde{\theta}_{k+N}\| \leq \frac{\bar{d}}{\alpha\delta} + \frac{N\bar{\varphi}}{r_{k+1}}, \quad \forall k \geq k_0,$$

for all $\alpha \in \left(0, \frac{1}{\sqrt{N^3}}\right)$.

References

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: a survey. *Computer Networks*, *38*(4), 393–422.

Brailsford, A. D., Yussouff, M., & Logothetis, E. M. (1993). Theory of gas sensors. *Sensors and Actuators B (Chemical)*, *13*(1), 135–138.

Calamai, P. H., & Moré, J. J. (1987). Projected gradient methods for linearly constrained problems. *Mathematical Programming*, *39*, 93–116.

Carbone, P., Schoukens, J., & Moschitta, A. (2020). Quick estimation of periodic signal parameters from 1-bit measurements. *IEEE Transactions on Instrumentation and Measurement*, *69*(2), 339–353.

Casini, M., Garulli, A., & Vicino, A. (2007). Time complexity and input design in worst-case identification using binary sensors. In *Proceedings of the 46th IEEE conference on decision and control* (pp. 5528–5533).

Casini, M., Garulli, A., & Vicino, A. (2011). Input design in worst-case system identification using binary sensors. *IEEE Transactions on Automatic Control*, *56*(5), 1186–1191.

Chen, H. F. (2002). *Stochastic approximation and its applications*. Dordrecht: Kluwer Academic Publishers.

Chen, H. F., & Guo, L. (1991). *Identification and stochastic adaptive control*. Boston: Birkhauser.

Chen, H. F., & Yin, G. (2003). Asymptotic properties of sign algorithms for adaptive filtering. *IEEE Transactions on Automatic Control*, *48*(9), 1545–1556.

Colinet, E., & Juillard, J. (2010). A weighted least-squares approach to parameter estimation problems based on binary measurements. *IEEE Transactions on Automatic Control*, *55*(1), 148–152.

Csáji, B. C., & Weyer, E. (2012). Recursive estimation of ARX systems using binary sensors with adjustable thresholds. *IFAC Proceedings Volumes*, *45*(16), 1185–1190.

Depraetere, B., Stoev, J., Pinte, G., & Swevers, J. (2012). Identification of linear systems with binary outputs using short independent experiments. *IFAC Proceedings Volumes*, *45*(16), 7–12.

Fu, J., & Wang, J. (2012). Adaptive coordinated tracking of multi-agent systems with quantized information. *Systems & Control Letters*, *74*, 115–125.

Fu, M., & Xie, L. (2005). The sector bound approach to quantized feedback control. *IEEE Transactions on Automatic Control*, *50*(11), 1698–1711.

Godoy, B., Goodwin, G., Agüero, J., Marelli, D., & Wigren, T. (2011). On identification of FIR systems having quantized output data. *Automatica*, *47*(9), 1905–1915.

Goudjil, A., Pouliquen, M., Pigeon, E., Gehan, O., & M'Saad, M. (2015). Identification of systems using binary sensors via Support Vector Machines. In *2015 54th IEEE conference on decision and control* (pp. 3385–3390).

Guo, L. (2020). *Time-varying stochastic systems, stability and adaptive theory* (2nd ed.). Beijing: Science Press.

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

Guo, J., & Zhao, Y. (2014). Identification of the gain system with quantized observations and bounded persistent excitations. *Science China. Information Sciences*, *57*, Article 012205.

Jafari, K., Juillard, J., & Colinet, E. (2010). A recursive system identification method based on binary measurements. In *49th IEEE conference on decision and control* (pp. 1154–1158).

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

Jing, L., & Zhang, J. F. (2019). Tracking control and parameter identification with quantized ARMAX systems. *Science China. Information Sciences*, *62*, Article 199203.

Knudson, K., Saab, R., & Ward, R. (2016). One-bit compressive sensing with norm estimation. *IEEE Transactions on Information Theory*, *62*(5), 2748–2758.

Pouliquen, M., Menard, T., Pigeon, E., Gehan, O., & Goudjil, A. (2016). Recursive system identification algorithm using binary measurements. In *2016 European control conference* (pp. 1353–1358).

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

Tan, S., Guo, J., Zhao, Y., & Zhang, J. F. (2021). Adaptive control with saturation-constrainted observations for drag-free satellites – a set-valued identification approach. *Science China. Information Sciences*, *64*, Article 202202.

Wang, L. Y., & Yin, G. (2007). Asymptotically efficient parameter estimation using quantized output observations. *Automatica*, *43*(7), 1178–1191.

Wang, L. Y., Yin, G., Zhang, J. F., & Zhao, Y. (2010). *System identification with quantized observations*. Boston: Birkhauser.

Wang, L. Y., Zhang, J. F., & Yin, G. (2003). System identification using binary sensors. *IEEE Transactions on Automatic Control*, *48*(11), 1892–1907.

Yin, G., Krishnamurthy, V., & Ion, C. (2003). Iterate-averaging sign algorithms for adaptive filtering with applications to blind multiuser detection. *IEEE Transactions on Information Theory*, *49*(3), 657–671.

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

Zahabi, S. J., Mahdi Naghsh, M., Modarres-Hashemi, M., & Li, J. (2017). Compressive pulse-doppler radar sensing via 1-bit sampling with time-varying threshold. In *2017 IEEE international conference on acoustics, speech and signal processing* (pp. 3419–3423).

Zemouche, A., Boutayeb, M., & Bara, G. I. (2005). Observer design for nonlinear systems: An approach based on the differential mean value theorem. In *Proceedings of the 44th IEEE conference on decision and control* (pp. 6353–6358).

Zhao, W., Chen, H. F., Tempo, R., & Dabbene, F. (2017). Recursive nonparametric identification of nonlinear systems with adaptive binary sensors. *IEEE Transactions on Automatic Control*, *62*(8), 3959–3971.

Zhao, Y., Wang, T., & Bi, W. (2019). Consensus protocol of multi-agent systems with undirected topologies and binary-valued communications. *IEEE Transactions on Automatic Control*, *64*(1), 206–221.

Zhao, Y., Wang, L. Y., Zhang, J. F., & Yin, G. (2009). Jointly deterministic and stochastic identification of linear systems using binary-valued observations. In *Proceedings of the 15th IFAC symposium on system identification* (pp. 60–65).

Ying Wang received the B.S. degree in Mathematics from Wuhan University, Wuhan, China, in 2017. She is currently working toward the Ph.D. degree majoring in system theory at Academy of Mathematics and Systems Science, Chinese Academy of Science, Beijing, China. Her research interests include the identification and control of quantized systems.

Yanlong Zhao received the B.S. degree in mathematics from Shandong University, Jinan, China, in 2002, and the Ph.D. degree in systems theory from the Academy of Mathematics and Systems Science (AMSS), Chinese Academy of Sciences (CAS), Beijing, China, in 2007. Since 2007, he has been with the AMSS, CAS, where he is currently a Professor. His research interests include identification and control of quantized systems, and modeling of financial systems. Dr. Zhao has been a Deputy Editor-in-Chief of the *Journal of Systems and Science and Complexity*, an Associate Editor of *Automatica* and *IEEE Transactions on Systems, Man and Cybernetics*; and was the Editor of the *IFAC Symposium on System Identification* in 2015.

Ji-Feng Zhang received the B.S. degree in mathematics from Shandong University, China, in 1985, and the Ph.D. degree from the Institute of Systems Science (ISS), Chinese Academy of Sciences (CAS), China, in 1991. Since 1985, he has been with the ISS, CAS, and now is the Director of ISS. His current research interests include system modeling, adaptive control, stochastic systems, and multi-agent systems. He is an IEEE Fellow, IFAC Fellow, CAA Fellow, member of the European Academy of Sciences and Arts, and Academician of the International Academy for Systems and Cybernetic Sciences. He received the Second Prize of the State Natural Science Award of China in 2010 and 2015, respectively. He is a Vice-President of the Chinese Mathematical Society and the Chinese Association of Automation. He was a Vice-Chair of the IFAC Technical Board, member of the Board of Governors, IEEE Control Systems Society; Convenor of Systems Science Discipline, Academic Degree Committee of the State Council of China; Vice-President of the Systems Engineering Society of China. He served as Editor-in-Chief, Deputy Editor-in-Chief or Associate Editor for more than 10 journals, including *Science China Information Sciences*, *IEEE Transactions on Automatic Control* and *SIAM Journal on Control and Optimization* etc.

Jin Guo received his B.S. degree in Mathematics from Shandong University, China, in 2008, and Ph.D. in System Modeling and Control Theory from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences in 2013. He is currently a professor in the School of Automation and Electrical Engineering, University of Science and Technology Beijing. His research interests are system modeling and identification, adaptive control, and systems biology. He is an Associate Editor of *Asian Journal of Control and Journal of Systems Science and Complexity*.