Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 59-19 (2025) 454–459

Recursive Identification of FIR Systems with Binary Outputs under Data Tampering Attacks

Jian Guo [*,**] **Wenchao Xue** [**] **Ji-Feng Zhang** [***,**]

[*] *CAS AMSS-PolyU Joint Laboratory of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China*
[**] *State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China and School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*
[***] *School of Automation and Electrical Engineering, Zhongyuan University of Technology, Zhengzhou 450007, China*
(e-mail: j.guo@amss.ac.cn, wenchaoxue@amss.ac.cn, jif@iss.ac.cn)

Abstract: This paper addresses the issue of parameter estimation in Cyber-Physical Systems where binary outputs are subject to data tampering attacks. A novel recursive identification algorithm is designed to estimate unknown parameters under tampered binary outputs, ensuring asymptotically convergent parameter estimation. Theoretical guarantees are provided for the algorithm's performance, including mean-square convergence, almost sure convergence, and convergence rate. Numerical simulations show the effectiveness of the proposed method.

Keywords: Binary-valued observations, tampering attack, system identification, recursive estimate, convergence rate.

1. INTRODUCTION

Cyber–physical systems (CPS) enable seamless interaction and coordination among humans, machines, environments, and other elements across both physical and cyber domains (Rawat et al., 2015). With their capability for real-time response, rapid iteration, and dynamic optimization, CPS have become integral in various fields, including industrial automation, healthcare, energy systems, and smart infrastructure (Lee et al., 2015; Hahn et al., 2013).

Despite these advantages, CPS face significant security risks due to the distributed nature of their networks and the use of low-cost sensors, which are often vulnerable to cyber threats (Peng et al., 2017; Pasqualetti et al., 2013; Teixeira et al., 2015; Kwon and Hwang, 2018). Two primary types of attacks are commonly observed (Peng et al., 2017): denial-of-service (DoS) attacks and deception attacks. Unlike DoS attacks, which result in data loss, deception attacks manipulate transmitted data to evade detection, leading to errors in decision-making, production management, and quality assurance. These errors can have far-reaching consequences, ranging from financial losses to major accidents, highlighting the urgent need for robust defense mechanisms against such attacks(Oliva et al., 2018; Cherpantier, 2011; Yamada et al., 2019).

In CPS, sensors often provide quantized data instead of precise measurements due to limitations in accuracy and cost. Binary outputs from devices such as switching sensors or optical detectors are common examples (Wang et al., 2010). Although quantization helps reduce data transmission and bandwidth usage, it introduces challenges in system analysis and control due to its inherent nonlinearity (Wu et al., 2013). This has led to substantial research efforts focused on the theory and methodologies for system identification using quantized observations (Wang et al., 2003; Guo and Zhao, 2013). In recent years, various studies have explored system identification with quantized data (Guo et al., 2015; Bottegal et al., 2017), covering applications of binary sensors (Pouliquen et al., 2020) within both stochastic and deterministic frameworks (Wang et al., 2022) and analyzing worst-case set-membership methods (Casini et al., 2011) and asymptotically efficiency (Wang et al., 2024b). These contributions have established a solid foundation for addressing identification problems in quantized systems.

In the context of binary systems subject to attacks, a significant body of literature has emerged. For example, Guo et al. (2020) addresses the defense against data tampering attacks in CPS by proposing a compensation-oriented defense approach and designing a corresponding identification algorithm. The algorithm is proven to exhibit strong consistency and asymptotic normality, achieving optimal defense. Guo et al. (2023) extends this analysis and algorithm to multi-dimensional scenarios. Additionally, Guo et al. (2021) tackles the issue of packet loss in system identification with binary-valued observations, proposing

⋆ This work was supported by the National Key Research and Development Program of China (Grant No. 2022YFA1004703), the National Natural Science Foundation of China under Grants 92471204 and 62433020, and the Youth Innovation Promotion Association, CAS. (Corresponding author: Wenchao Xue.)

10.1016/j.ifacol.2025.11.076

algorithms for both known and unknown packet loss probabilities and offering a compensation-oriented approach to minimize communication while ensuring accurate estimation. Moreover, Guo et al. (2025) investigates identifying FIR systems against random replay attacks with binary-valued observations, designing a defense algorithm that remains consistent during attacks, and deriving the optimal defense strategy based on asymptotic normality.

However, existing defense solutions are predominantly based on empirical measures. This article explores the design of recursive algorithms for parameter estimation in CPS when binary outputs are compromised by data tampering attacks. The proposed method ensures asymptotically convergent parameter estimation and establishes theoretical guarantees for the algorithm's performance.

The main contributions of this work are summarized as follows:

- This paper proposes a recursive identification algorithm for estimating unknown parameters under tampered binary outputs. The algorithm processes data sequentially and effectively addresses the challenges posed by binary observations and adversarial attacks.
- The algorithm's performance is rigorously analyzed, with guarantees for mean-square convergence, almost sure convergence, and convergence rate. Typical simulations show its robustness, ensuring reliable parameter estimation even under high tampering probabilities.

The structure of this paper is as follows: Section 2 introduces the system identification problem under tampered binary observations. Section 3 develops the recursive identification algorithm, proving its convergence properties and convergence rate. Section 4 supports the theoretical findings with numerical simulations. Finally, Section 5 concludes the paper and discusses future research directions.

2. PROBLEM STATEMENT

In this section, we will first introduce the identification problem with tampering attack. Consider the parameter identification of the following stochastic FIR systems:

$$y_{k+1} = \varphi_k^T \theta + w_{k+1}, \quad k = 0, 1, \ldots, \tag{1}$$

where θ is unknown but time invariant p-dimensional parameter vector with p being known, $\varphi_k \in \mathbb{R}^p$ is the regressor vector consisting current and past inputs, y_{k+1} and w_{k+1} are the system output and noise, respectively.

Let $\mathcal{F}_k$ be some σ fields sequence defined by

$$\mathcal{F}_k \triangleq \sigma\left\{\varphi_k, \ldots, \varphi_0, w_k, \ldots, w_0\right\}, \quad k \geq 0.$$

In our settings, however, on the one hand, the system output y_k is unknown and it can only be measured by a binary-valued sensor with some threshold $C \in (-\infty, \infty)$. Mathematically, the binary-valued output can be represented by

$$s_{k+1}^0 = I_{[y_{k+1} \leq C]} = \begin{cases} 1, & y_{k+1} \leq C \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

On the other hand, s_{k+1}^0 is delivered to the processing center through the communication network, but is subject to data tampering attacks. The data received at the processing center is denoted as s_{k+1}. The tampering method and the corresponding probability are given by the following equation:

$$\begin{cases} \Pr\left\{s_k = 0 \mid s_k^0 = 1\right\} = p \\ \Pr\left\{s_k = 1 \mid s_k^0 = 0\right\} = q. \end{cases} \tag{3}$$

In the framework of such a tampering attack, the following question arises: how can the processing center use the tampered observations to identify the parameter of the real system.

Identification objective. The goal of this paper is to construct a recursive identification algorithm for estimating the unknown parameter vector θ based on the input φ_k and the possibly tampered binary observations s_k. For this problem, we consider the scenarios where the processing center knows the attacker's strategy (p, q) in advance.

Remark 1. In practice, if the probability is unknown, it can be estimated by conducting multiple tests to measure the channel error rate and determine the probability of correct transmission.

3. IDENTIFICATION ALGORITHM

3.1 Motivation and Algorithm

Classical identification algorithms, such as the least squares (LS) algorithm and stochastic approximation algorithm, can produce consistent estimates of unknown parameters as the number of data tends to infinity. However, when only a possibly tampered binary observation can be obtained, these methods will not work. Due to the limitation of the output information, to achieve the recognition of the target, we need to know more information, such as more noise properties. Before giving the algorithm, we first review the classical least squares method. It obtain an estimate of θ by minimizing the following "accumulated forecast error"

$$I_n(\beta) \triangleq \sum_{k=0}^{n} \left(y_{i+1} - \beta^T \varphi_k\right)^2. \tag{4}$$

Assume that the noise $\{w_{k+1}, \mathcal{F}\}$ is a martingale difference sequence, i.e. $\mathbb{E}\left[w_{k+1} \mid \mathcal{F}_k\right] = 0$. Since

$$\mathbb{E}\left[y_{k+1} \mid \mathcal{F}_k\right] = \theta^T \varphi_k,$$

the least square method is to minimize the "variance" of the measured output data.

Denote the conditional probability distribution function of $w_{k+1}|\mathcal{F}_k$ as $F_k(\cdot)$. For the possibly tampered binary output case, by Law of total probability, one can calculate that

$$\begin{aligned} &P(s_{k+1} = 0|\mathcal{F}_k) \\ &= P(s_{k+1}^0 = 1|\mathcal{F}_k) P(s_{k+1} = 0|s_{k+1}^0 = 1, \mathcal{F}_k) \\ &\quad + P(s_{k+1}^0 = 0|\mathcal{F}_k) P(s_{k+1} = 0|s_{k+1}^0 = 0, \mathcal{F}_k) \\ &= pF_k(C - \theta^T \varphi_k) + (1 - q)(1 - F_k(C - \theta^T \varphi_k)) \\ &= (p + q - 1) F_k(C - \theta^T \varphi_k) + 1 - q \end{aligned} \tag{5}$$

and

$$\begin{aligned} P(s_{k+1} = 1|\mathcal{F}_k) &= 1 - P(s_{k+1} = 0|\mathcal{F}_k) \\ &= (1 - (p + q)) F_k(C - \theta^T \varphi_k) + q. \end{aligned} \tag{6}$$

Identifiability. First, we discuss the identifiability of the binary system (1)-(3) under attack. From the distributions of s_k in (5)-(6), the identifiability of this system requires that $p+q \neq 1$ and the excitation condition regarding φ_k, which will be provided in subsequent sections.

Motivation. The following presents the idea for constructing the algorithm. By (5)-(6), we have

$$\mathbb{E}(s_{k+1} \mid \mathcal{F}_k) = (1-(p+q))F_k(C-\theta^T\varphi_k)+q. \quad (7)$$

This, together with the structure of the least square, motivates us to construct the following "accumulated forecast error" for the possibly *Tampered Binary* case

$$J_{TB,n}(\beta)=\sum_{k=1}^{n}\left(s_{k+1}-\left((1-(p+q))F_k(C-\beta^T\varphi_k)+q\right)\right)^2.$$

Based on this objective function and the principles of stochastic approximation, a recursive algorithm for the estimation of the unknown parameter θ can be formulated. To initiate the process, we first define the following projection operator.

Definition 1. For a given convex compact set $\Omega \subseteq \mathbb{R}^n$, the projection operator $\Pi_\Omega(\cdot)$ is defined as

$$\Pi_\Omega(x) = \operatorname*{argmin}_{\omega\in\Omega}\|x-\omega\|, \quad \forall x \in \mathbb{R}^n.$$

Proposition 1. The projection operator given by Definition 1 follows

$$\|\Pi_\Omega(x)-\Pi_\Omega(y)\| \leq \|x-y\|, \quad \forall x,y \in \mathbb{R}^n.$$

Given some positive step-size $\{b_k\}_{k\geq 1}$ and $\beta > 0$, the recursive projection algorithm is designed as follows.

Algorithm.

$$\hat{\theta}_{k+1} = \Pi_\Theta\left\{\hat{\theta}_k + b_k\varphi_k\tilde{s}_{k+1}\right\}, \quad (8)$$

$$\tilde{s}_{k+1} = \beta(1-(p+q))\times \left((1-(p+q))F_k(C-\hat{\theta}_k^T\varphi_k)+q-s_{k+1}\right). \quad (9)$$

The following subsections will establish the almost sure and mean square convergence of the identification algorithm, along with deriving the convergence rate of the estimation error.

3.2 Convergence properties.

Assumption 1. A priori information of the unknown parameter is that $\theta \in \Theta \subseteq \mathbb{R}^n$, where Θ is a bounded convex compact set with $B = \sup_{\nu\in\Theta}\|\nu\|$ and $\|\cdot\|$ is the Euclidean norm.

Assumption 2. For the observation noise sequence $\{w_k\}$, its conditional probability distribution and density functions given $\mathcal{F}_k$ are known and denoted by $F_k(\cdot)$ and $f_k(\cdot)$.

Assumption 3. The input φ_k satisfies

$$\sup_{k\geq 1}\|\varphi_k\| \leq M < \infty \quad (10)$$

and there exist a positive integer $h \geq n$ and a constant $\delta > 0$ such that

$$\frac{1}{h}\mathbb{E}\left[\sum_{l=k}^{k+h-1}\varphi_l\varphi_l^T\middle|\mathcal{F}_{k-1}\right] \geq \delta I, \quad k=1,2,\ldots \quad (11)$$

where I is an $n\times n$ identity matrix.

Assumption 4 The step size b_k satisfies:

$$\sum_{k=0}^{\infty} b_k = \infty, \quad \lim_{k\to\infty} b_k = 0, \quad \text{and} \quad b_k = O(b_{k+1}).$$

Assumption 5. The corresponding conditional density functions $f_k(\cdot)$ given $\mathcal{F}_{k-1}$ of the observation noise w_k satisfies: $\underline{f} = \inf_{k\geq 1}\inf_{|x|\leq C+MB} f_k(x) > 0$.

Remark 2. Assumption 2 is widely applied in binary classification fields, such as Wang et al. (2003); Guo and Zhao (2013). Condition (11) is commonly referred to as the "sufficiently rich condition" , upon which an identification algorithm can be developed to guarantee the identifiability of θ. This assumption is less restrictive than the traditional persistent excitation assumption.

Remark 3. The conditions $\sum_{k=0}^{\infty} b_k = \infty$ and $\lim_{k\to\infty} b_k = 0$ are typical for stochastic approximation algorithms. Additionally, the condition $b_k = O(b_{k+1})$ ensures that the step sizes change gradually, which helps handle the conditional expectation-type excitation condition.

Denote the estimate error $\tilde{\theta}_k = \hat{\theta}_k - \theta,\ k = 0,1,\ldots$. First, we give the following useful lemmas.

Lemma 3.1. (Pedregal, 2004) Let $\{p_k\},\{q_k\}$ and $\{\alpha_k\}$ be real sequences satisfying $p_{k+1} \leq (1-q_k)p_k + \alpha_k$, where $0 < q_k \leq 1$, $\sum_{k=0}^{\infty} q_k = \infty$, $\alpha_k \geq 0$, and $\lim_{k\to\infty}\frac{\alpha_k}{q_k} = 0$. Then, $\limsup_{k\to\infty} p_k \leq 0$.

Lemma 3.2. (Wang et al., 2024a) For $0 < b \leq 1, a > 0, k_0 \geq 0$ and sufficiently large l, we have

$$\begin{cases} \prod_{i=l}^{k}\left(1-\dfrac{a}{(i+k_0)^b}\right) \\ \leq \begin{cases} \left(\dfrac{l+k_0}{k+k_0}\right)^a, & b=1, \\ e^{\frac{a}{1-b}\left((l+k_0)^{1-b}-(k+k_0+1)^{1-b}\right)}, & b\in(0,1); \end{cases} \\ \sum_{l=1}^{k}\prod_{i=l}^{k}\left(1-\dfrac{a}{(i+k_0)^b}\right)O\left(\dfrac{1}{(l+k_0)^{2b}}\right) = O\left(\dfrac{1}{k^b}\right), \\ \qquad\qquad b\in(0,1). \end{cases}$$

Lemma 3.3. (Zhang et al., 2019) For any given positive integer l and $a,b \in \mathbb{R}$, the following results hold

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

Lemma 3.4. If Assumptions 1–3 hold, then

$$\|\tilde{\theta}_{k+l} - \tilde{\theta}_k\| = O(b_{k+l}), \text{ for } k,\ l \in \mathbb{N}.$$

Proof. First, we have

$$\|\tilde{\theta}_{k+l} - \tilde{\theta}_k\| = \|\hat{\theta}_{k+l} - \hat{\theta}_k\| = \left\|\sum_{j=1}^{l}(\hat{\theta}_{k+j} - \hat{\theta}_{k+j-1})\right\|$$

$$\leq \sum_{j=1}^{l}\left\|(\hat{\theta}_{k+j} - \hat{\theta}_{k+j-1})\right\|. \quad (12)$$

Since $0 \leq (1-(p+q))F_k(C-\theta^T\varphi_k)+q \leq 1$, it follows that $|\tilde{s}_k| \leq \beta$. Combining this with Proposition

1 and the condition $\|\phi_k\| \le M$, we obtain $\|\hat{\theta}_{l+1} - \hat{\theta}_l\| \le b_{l+1}\|\phi_{l+1}\tilde{s}_{l+1}\| \le b_k\beta M$ for $l \ge 1$. This result, together with (12) and Assumption 4, implies the lemma. □

The following theorem establishes the almost sure and mean square convergence of the identification algorithm.

Theorem 1. For system (1) with the binary-valued observation (2) and the tampering attack (3), under Assumptions 1–5, the parameter estimate given by the algorithm (8)–(9) is mean-square convergent, i.e. $\lim_{k\to\infty} E[\tilde{\theta}_k^T\tilde{\theta}_k] = 0$. Moreover, if $\sum_{k=1}^{\infty} b_k^2 < \infty$, then the estimate $\hat{\theta}_k$ is also almost surely convergent, i.e. $\lim_{k\to\infty}\tilde{\theta}_k = 0$, a.s.

Proof. By $\tilde{s}_k^2 \le \beta^2$, Proposition 1 and (8), we have

$$\begin{aligned}\|\tilde{\theta}_{k+1}\|^2 &\le \|\tilde{\theta}_k\|^2 + 2b_k\tilde{s}_{k+1}\phi_k^T\tilde{\theta}_k + b_k^2\|\phi_k\|^2\beta^2 \\ &= \|\tilde{\theta}_k\|^2 + 2b_k\tilde{s}_{k+1}\phi_k^T\tilde{\theta}_k + O(b_k^2). \end{aligned} \tag{13}$$

From (7) and (9), it follows that

$$\begin{aligned}&\mathbb{E}[\tilde{s}_{k+1}|\mathcal{F}_k] \\ &= \beta(1-p-q)^2\left(F(C-\phi_k^T\hat{\theta}_k) - F(C-\phi_k^T\theta)\right). \end{aligned} \tag{14}$$

This together with Assumption 2 and the differential mean value theorem, it leads to

$$\begin{aligned}&\mathbb{E}\left[2b_k\tilde{s}_{k+1}\phi_k^T\tilde{\theta}_k|\mathcal{F}_k\right] = 2b_k\phi_k^T\tilde{\theta}_k\mathbb{E}[\tilde{s}_{k+1}|\mathcal{F}_k] \\ &= 2b_k\phi_k^T\tilde{\theta}_k\beta(1-p-q)^2\left(F_k(C-\phi_k^T\hat{\theta}_k) - F_k(C-\phi_k^T\theta)\right) \\ &= -2b_k\beta(1-p-q)^2 f_k(\xi_k)\tilde{\theta}_k^T\phi_k\phi_k^T\tilde{\theta}_k \\ &\le -2b_k\beta(1-p-q)^2\underline{f}\tilde{\theta}_k^T\phi_k\phi_k^T\tilde{\theta}_k, \end{aligned} \tag{15}$$

where ξ_k is in the interval between $C-\phi_k^T\hat{\theta}_k$ and $C-\phi_k^T\theta$ such that $F_k(C-\phi_k^T\hat{\theta}_k) - F_k(C-\phi_k^T\theta) = f_k(\xi_k)\tilde{\theta}_k^T\phi_k\phi_k^T\tilde{\theta}_k$. Taking the expectation on both sides of (13) and substituting (15) into it, we can obtain

$$\begin{aligned}&\mathbb{E}\|\tilde{\theta}_{k+1}\|^2 \le \mathbb{E}\|\tilde{\theta}_k\|^2 + 2b_k\mathbb{E}\tilde{s}_{k+1}\phi_k^T\tilde{\theta}_k + O(b_k^2) \\ &= \mathbb{E}\|\tilde{\theta}_k\|^2 + \mathbb{E}\left[\mathbb{E}\left[2b_k\tilde{s}_{k+1}\phi_k^T\tilde{\theta}_k|\mathcal{F}_k\right]\right] + O(b_k^2). \\ &\le \mathbb{E}\|\tilde{\theta}_k\|^2 - 2b_k\beta(1-p-q)^2\underline{f}\mathbb{E}[\tilde{\theta}_k^T\phi_k\phi_k^T\tilde{\theta}_k] \\ &\quad + O(b_k^2). \end{aligned} \tag{16}$$

By iterating (16) h times and noting $b_k = O(b_{k+1})$, we obtain

$$\begin{aligned}&\mathbb{E}\|\tilde{\theta}_{k+h}\|^2 \\ &\le \mathbb{E}\|\tilde{\theta}_k\|^2 - 2\beta(1-p-q)^2\underline{f}\mathbb{E}\left[\sum_{l=k}^{k+h-1}[b_l\tilde{\theta}_l^T\phi_l\phi_l^T\tilde{\theta}_l]\right] \\ &\quad + O(b_{k+h}^2) \\ &\le \mathbb{E}\|\tilde{\theta}_k\|^2 - 2\beta(1-p-q)^2\underline{f}\mathbb{E}\left[\sum_{l=k}^{k+h-1}[b_l\tilde{\theta}_k^T\phi_l\phi_l^T\tilde{\theta}_k]\right] \\ &\quad - 2\beta(1-p-q)^2\underline{f}\mathbb{E}\left[\sum_{l=k}^{k+h-1}[b_l(\tilde{\theta}_l-\tilde{\theta}_k)^T\phi_l\phi_l^T(\tilde{\theta}_l-\tilde{\theta}_k)]\right] \\ &\quad + O(b_{k+h}^2). \end{aligned} \tag{17}$$

By Lemma 3.4 and (10), the last two terms of (16) are of order $O(b_{k+h}^2)$. In addition,

$$\begin{aligned}&\mathbb{E}\left[\sum_{l=k}^{k+h-1}[b_l\tilde{\theta}_k^T\phi_l\phi_l^T\tilde{\theta}_k]\right] \\ &= \mathbb{E}\left[\tilde{\theta}_k^T\mathbb{E}\left[\sum_{l=k}^{k+h-1}b_l\phi_l\phi_l^T\middle|\mathcal{F}_k\right]\tilde{\theta}_k\right]. \end{aligned} \tag{18}$$

By Assumption 3, we have

$$\begin{aligned}&\mathbb{E}\left[\sum_{l=k}^{k+h-1}b_l\phi_l\phi_l^T\middle|\mathcal{F}_k\right] \\ &= \sum_{l=k}^{k+h-1}b_l\frac{1}{h}\mathbb{E}\left[\sum_{l=k}^{k+h-1}\phi_l\phi_l^T\middle|\mathcal{F}_k\right] + O(b_{k+h}^2) \\ &\ge \delta\sum_{l=k}^{k+h-1}b_l I + O(b_{k+h}^2). \end{aligned} \tag{19}$$

Substituting (18) and (19) into (17) yields

$$\begin{aligned}&\mathbb{E}\|\tilde{\theta}_{k+h}\|^2 \\ &\le \mathbb{E}\|\tilde{\theta}_k\|^2 - 2\beta(1-p-q)^2\underline{f}\delta\sum_{l=k}^{k+h-1}b_l\mathbb{E}\|\tilde{\theta}_k\|^2 + O(b_{k+h}^2). \end{aligned} \tag{20}$$

Then, based on Lemma 3.1 and Assumption 4, and noting $\sum_{k=1}^{\infty}b_k = \infty$ and $\lim_{k\to\infty}\frac{b_k^2}{\sum_{l=k-h}^{k-1}b_{l+1}} = 0$, it follows that $\lim_{k\to\infty}E[\|\tilde{\theta}_k\|^2] = 0$.

On the other hand, by (17) we have

$$E[\|\tilde{\theta}_{k+1}\|^2|\mathcal{F}_k] \le \|\tilde{\theta}_k\|^2 + O(b_k^2),$$

which together with Lemma 1.2.2 in Chen (2021) and $\sum_{k=1}^{\infty}b_k^2 < \infty$ implies that $\|\tilde{\theta}_k\|$ converges to a bounded limit a.s. Notice that $\lim_{k\to\infty}E[\tilde{\theta}_k^T\tilde{\theta}_k] = 0$. Then, there is a subsequence of $\tilde{\theta}_k$ that converges almost surely to 0. Consequently, $\tilde{\theta}_k$ almost surely converges to 0. □

3.3 Convergence rate

This subsection gives the convergence rate of the algorithm (8)–(9).

Theorem 2. Under Assumption 1-5, the algorithm (8)–(9) has the following mean square convergence rate:

- If the step-size $b_k = \frac{1}{k^\gamma}$ with $\frac{1}{2} < \gamma < 1$, then
$$\mathbb{E}\|\tilde{\theta}_k\|^2 = O\left(\frac{1}{k^\gamma}\right).$$
- If the step-size $b_k = \frac{1}{k}$ and $\beta > \frac{1}{2(1-p-q)^2\underline{f}\delta}$, then
$$\mathbb{E}\|\tilde{\theta}_k\|^2 = O\left(\frac{1}{k}\right).$$

Proof. When $b_k = \frac{1}{k^\gamma}$, $\gamma \in (1/2, 1)$, letting $\alpha = 2\beta(1-p-q)^2\underline{f}\delta$ and based on (20), we have

$$\mathbb{E}\|\tilde{\theta}_k\|^2 \le \left(1-\alpha\sum_{l=k-h}^{k-1}\frac{1}{(l+1)^\gamma}\right)\mathbb{E}\|\tilde{\theta}_k\|^2+O\left(\frac{1}{k^{2\gamma}}\right),$$
$$\le\left(1-\frac{\alpha h}{k^\gamma}\right)\mathbb{E}\|\tilde{\theta}_{k-h}\|^2+O\left(\frac{1}{k^{2\gamma}}\right)$$
$$\le\prod_{l=0}^{\lfloor\frac{k-K}{h}\rfloor-1}\left(1-\frac{\alpha h}{(k-lh)^\gamma}\right)\mathbb{E}\left\|\tilde{\theta}_{k-\lfloor\frac{k-K}{h}\rfloor h}\right\|^2$$
$$+\sum_{l=1}^{\lfloor\frac{k-K}{h}\rfloor}\prod_{j=0}^{l-1}\left(1-\frac{\alpha h}{(k-jh)^\gamma}\right)O\left(\frac{1}{(k-lh)^{2\gamma}}\right)$$
$$\le\prod_{l=\lceil\frac{K}{h}\rceil+\kappa+1}^{\lceil\frac{k}{h}\rceil}\left(1-\frac{\alpha h^{1-\gamma}}{l^\gamma}\right)\mathbb{E}\left\|\tilde{\theta}_{k-\lfloor\frac{k-K}{h}\rfloor h}\right\|^2$$
$$+\sum_{l=\lceil\frac{K}{h}\rceil+1}^{\lfloor\frac{k}{h}\rfloor-1}\prod_{q=\lceil\frac{K}{h}\rceil+\kappa+l+1}^{\lceil\frac{k}{h}\rceil}\left(1-\frac{\alpha h^{1-\gamma}}{j^\gamma}\right)O\left(\frac{1}{l^{2\gamma}}\right),$$

where $\kappa=\lceil\frac{k-K}{h}\rceil-\lfloor\frac{k-K}{h}\rfloor$. This together with Lemma 3.2 yields $\mathbb{E}\|\tilde{\theta}_k\|^2=O\left(\frac{1}{k^\gamma}\right)$.

When $b_k=\frac{1}{k}$, letting $\alpha=2\beta(1-p-q)^2\underline{f}\delta$ and by (20), we have

$$\mathbb{E}\|\tilde{\theta}_k\|^2 \le \left(1-\alpha\sum_{l=k-h}^{k-1}\frac{1}{l+1}\right)\mathbb{E}\|\tilde{\theta}_{k-h}\|^2+O\left(\frac{1}{k^2}\right),$$
$$\le\left(1-\frac{\alpha h}{k}\right)\mathbb{E}\|\tilde{\theta}_{k-h}\|^2+O\left(\frac{1}{k^2}\right)$$
$$\le\prod_{l=0}^{\lfloor\frac{k-K}{h}\rfloor-1}\left(1-\frac{\alpha h}{k-lh}\right)\mathbb{E}\|\tilde{\theta}_{k-\lfloor\frac{k-K}{h}\rfloor h\|}\|^2$$
$$+\sum_{l=1}^{\lfloor\frac{k-K}{h}\rfloor}\prod_{q=0}^{l-1}\left(1-\frac{\alpha h}{k-qh}\right)O\left(\frac{1}{(k-lh)^2}\right)$$
$$\le\prod_{l=}^{\lceil\frac{K}{h}\rceil+\kappa+1}\left(1-\frac{\alpha}{l}\right)\mathbb{E}\|\tilde{\theta}_{k-\lfloor\frac{k-K}{h}\rfloor h\|}\|^2$$
$$+\sum_{l=\lceil\frac{K}{h}\rceil+1}^{\lfloor\frac{k}{h}\rfloor-1}\sum_{q=\lceil\frac{K}{h}\rceil+\kappa+l+1}^{\lceil\frac{k}{h}\rceil}\left(1-\frac{\alpha}{q}\right)O\left(\frac{1}{l^2}\right)$$

where $\kappa=\lceil\frac{k-K}{h}\rceil-\lfloor\frac{k-K}{h}\rfloor$. Since $\beta>\frac{1}{2(1-p-q)^2\underline{f}\delta}$, i.e. $\alpha>1$. Thus, by Lemma 3.3, we have $\mathbb{E}\left\|\tilde{\theta}_k\right\|^2=O\left(\frac{1}{k}\right)$. This completes this part's proof.

4. SIMULATION STUDIES

Consider the system $y_{k+1}=\varphi_k^T\theta+w_{k+1}$ with the binary observation

$$s_k=I_{[y_k\le C]}=\begin{cases}1, & y_k\le C\\ 0, & \text{otherwise},\end{cases}$$

where $\theta=[3,-1]^T$ is unknown but known as in $\Theta=\{(x,y):|x|<6,|y|<6\}$. The threshold $C=1$, and the system noise w_{k+1} obeys the standard normal distribution. The inputs $\varphi_k=\{u_k,u_{k-1}\}$ with u_k obeying the uniform distribution of $N(0,2)$. Algorithm (8)-(9) has a step size of $\beta=80$, $b_k=1/k^\gamma$ with $\gamma=1,\ 0.8$, and an initial value of $\theta_0=[1,1]^T$. All the simulations are looped 50 times. Figure 1 presents the estimation results of the algorithm for attack strategy shown in (3) as $p=0.2,q=0.3$ and $p=0.8,q=0.9$. From Figure 1, it can be seen that even when the tampering probability is close to 1, the proposed recursive defense algorithm still converges to the true value. Figure 2 shows the convergence rate for $\gamma=1$ and $\gamma=0.8$, validating the results derived in Theorem 2.

5. CONCLUDING REMARKS

This paper presents a novel recursive identification algorithm for parameter estimation in CPS with binary outputs, compromised by data tampering attacks. The proposed method guarantees asymptotically convergent parameter estimation, even in the presence of high-probability tampering, and provides theoretical guarantees, including mean-square convergence, almost convergence and convergence rate. Future research could consider the following issues: (i) designing recursive identification algorithms when the tampering attack probability is unknown, (ii) studying recursive identification algorithms for multi-agent systems under distributed communication and tampering attacks, and (iii) investigating second-order recursive defense identification algorithms.

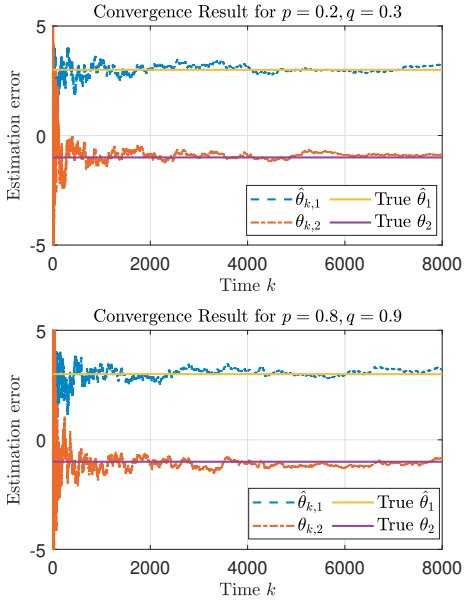

Fig. 1. Convergence of the estimation shown by a trajectory of $\hat{\theta}_{n+1}$ for the case $p=0.2,q=0.3$ and $p=0.8,q=0.9$.

REFERENCES

Bottegal, G., Hjalmarsson, H., and Pillonetto, G. (2017). A new kernel-based approach to system identification with quantized output data. *Automatica*, 85, 145–152.

Casini, M., Garulli, A., and Vicino, A. (2011). Input design in worst-case system identification using binary sensors.

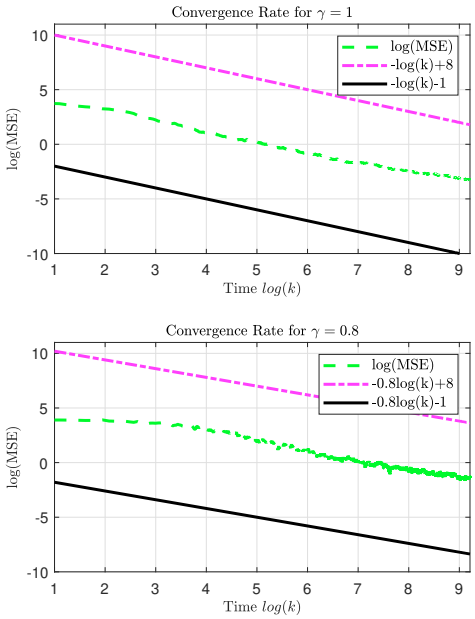

Fig. 2. Convergence rate of the estimation shown by a trajectory of $k\tilde{\theta}_k^T\tilde{\theta}_k/\ln k$ for the case $p = 0.1, q = 0.2$.

IEEE Transactions on Automatic Control, 56(5), 1186–1191.

Chen, H.F. (2021). Stochastic approximation with applications. In *Encyclopedia of Systems and Control*, 2154–2160. Springer.

Cherpantier, F. (2011). Tamper reactive memory device to secure data from tamper attacks.

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

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

Guo, J., Cheng, J., and Diao, J.D. (2021). System identification with binary-valued output observations under either-or communication and data packet dropout. *Systems & Control Letters*, 156, 105010.

Guo, J., Jia, R., Su, R., and Zhao, Y. (2023). Identification of fir systems with binary-valued observations against data tampering attacks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(9), 5861–5873.

Guo, J., Wang, X., Xue, W., and Zhao, Y. (2020). System identification with binary-valued observations under data tampering attacks. *IEEE Transactions on Automatic Control*, 66(8), 3825–3832.

Guo, J., Zhang, Q., and Zhao, Y. (2025). Identification of fir systems with binary-valued observations under replay attacks. *Automatica*, 172, 112001.

Hahn, A., Ashok, A., Sridhar, S., and Govindarasu, M. (2013). Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid. *IEEE Transactions on Smart Grid*, 4(2), 847–855.

Kwon, C. and Hwang, I. (2018). Reachability analysis for safety assurance of cyber-physical systems against cyber attacks. *IEEE Transactions on Automatic Control*, 63(7), 2272–2279.

Lee, J., Bagheri, B., and Kao, H.A. (2015). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing letters*, 3, 18–23.

Oliva, G., Cioabă, S., and Hadjicostis, C.N. (2018). Distributed calculation of edge-disjoint spanning trees for robustifying distributed algorithms against man-in-the-middle attacks. *IEEE Transactions on Control of Networked Systems*, 5(4), 1646–1656.

Pasqualetti, F., Dörfler, F., and Bullo, F. (2013). Attack detection and identification in cyber-physical systems. *IEEE Transactions on Automatic Control*, 58(11), 2715–2729.

Pedregal, P. (2004). *Introduction to optimization*, volume 46. Springer.

Peng, L., Shi, L., Cao, X., and Sun, C. (2017). Optimal attack energy allocation against remote state estimation. *IEEE Transactions on Automatic Control*, 63(7), 2199–2205.

Pouliquen, M., Pigeon, E., Gehan, O., and Goudjil, A. (2020). Identification using binary measurements for iir systems. *IEEE Transactions on Automatic Control*, 65(2), 786–793.

Rawat, D.B., Rodrigues, J.J., and Stojmenovic, I. (2015). *Cyber-physical systems: from theory to practice*. CRC Press.

Teixeira, A., Shames, I., Sandberg, H., and Johansson, K.H. (2015). A secure control framework for resource-limited adversaries. *Automatica*, 51, 135–148.

Wang, J., Ke, J., and Zhang, J.F. (2024a). Differentially private bipartite consensus over signed networks with time-varying noises. *IEEE Transactions on Automatic Control*.

Wang, L.Y., Yin, G.G., Zhang, J.F., and Zhao, Y.L. (2010). *System Identification with Quantized Observations*. Birkhäuser, Boston, MA, USA.

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

Wang, Y., Zhao, Y., and Zhang, J.F. (2024b). Asymptotically efficient quasi-newton type identification with quantized observations under bounded persistent excitations. *Automatica*, 166, 111722.

Wang, Y., Zhao, Y., Zhang, J.F., and Guo, J. (2022). A unified identification algorithm of fir systems based on binary observations with time-varying thresholds. *Automatica*, 135, 109990.

Wu, J., Jia, Q., Johansson, K., and Shi, L. (2013). Event-based sensor data scheduling: Trade-off between sensor communication rate and estimation quality. *IEEE Transactions on Automatic Control*, 58(4), 1041–1046.

Yamada, K., Hoshino, J., and Kubo, R. (2019). Detection of data tampering attacks using redundant network paths with different delays for networked control systems. *Nonlinear Theory and Its Applications*, 10(2), 140–156.

Zhang, H., Wang, T., and Zhao, Y. (2019). Asymptotically efficient recursive identification of fir systems with binary-valued observations. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(5), 2687–2700.