Recursive Binary Identification under Data Tampering and Non-Persistent Excitation with Application to Emission Control

Jian Guo, Lihong Pei, Wenchao Xue, *Member, IEEE*, Yanlong Zhao, *Senior Member, IEEE*, and Ji-Feng Zhang, *Fellow, IEEE*

***Abstract*— This paper studies the problem of online parameter estimation for cyber-physical systems with binary outputs that may be subject to adversarial data tampering. Existing methods are primarily offline and unsuitable for real-time learning. To address this issue, we first develop a first-order gradient-based algorithm that updates parameter estimates recursively using incoming data. Considering that persistent excitation (PE) conditions are difficult to satisfy in feedback control scenarios, a second-order quasi-Newton algorithm is proposed to achieve faster convergence without requiring the PE condition. For both algorithms, corresponding versions are developed to handle known and unknown tampering strategies, and their parameter estimates are proven to converge almost surely over time. In particular, the second-order algorithm ensures convergence under a signal condition that matches the minimal excitation required by classical least-squares estimation in stochastic regression models. The second-order algorithm is also extended to an adaptive control framework, providing an explicit upper bound on the tracking error for binary-output FIR systems under unknown tampering. Three numerical simulations verify the theoretical results and show that the proposed methods are robust against data tampering. Finally, the approach is validated via a vehicle emission control problem, where it effectively improves the detection accuracy of excess-emission events.**

***Index Terms*— Binary-valued observations, tampering attack, system identification, convergence rate, adaptive control, vehicle emission control**

This work was supported by the National Natural Science Foundation of China under Grants 62122083, 62433020 and T2293770, Youth Innovation Promotion Association, CAS, and CAS Project for Young Scientists in Basic Research under Grant YSBR-008.

Jian Guo is with the CAS AMSS-PolyU Joint Laboratory of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China (e-mail: j.guo@amss.ac.cn).

Lihong Pei, Wenchao Xue, and Yanlong Zhao are with the State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: plh1225@amss.ac.cn; wenchaoxue@amss.ac.cn; ylzhao@amss.ac.cn).

Ji-Feng Zhang is with the School of Automation and Electrical Engineering, Zhongyuan University of Technology, Zhengzhou 450007, China, the State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, and the School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China (jif@iss.ac.cn)

I. Introduction

Cyber–Physical Systems (CPS) connect people, machines, environments, and computational components, enabling close interaction between the physical and digital worlds [1], [2]. By combining sensing, communication, computing, and control, CPS can respond to real-time changes, adapt to dynamic environments, and continuously improve system performance. These features make CPS a core technology in many industrial applications, such as automated manufacturing, intelligent transportation, healthcare, energy systems, and smart infrastructure [3], [4].

Despite their advantages, CPS often face cybersecurity risks. These vulnerabilities mainly stem from their distributed structures and use of inexpensive sensors and communication devices that frequently lack strong protection [5]–[7]. Such weaknesses allow attackers to disrupt operations, making security crucial in CPS design [8]. Among various cyber threats, Denial-of-Service (DoS) and Data Tampering attacks are particularly common and harmful [5]. DoS attacks block or delay data transmission [9], while Data Tampering attacks subtly alter transmitted data, making detection difficult [10]. These attacks can lead to wrong decisions, reduced performance, or system failures, potentially causing production delays, financial losses, and safety incidents [11], [12].

Several methods have been proposed to strengthen CPS security. For instance, Bayesian-based defense methods detect deceptive attacks by continuously updating the system's belief about potential threats [13], [14]. Other studies have developed proactive schemes to protect critical infrastructures, such as power grids, against dynamic attacks [15]. Research on data-injection attacks has considered trade-offs between attack detection and control performance [16], and also examined balancing security with privacy in connected systems [17]. Adaptive control techniques have been widely explored to maintain stable system operation even if sensors or actuators are compromised [18]–[20]. Additionally, secure-by-construction approaches have been developed to design CPS with built-in security measures, reducing vulnerabilities from the start [21], [22].

In addition to security risks, CPS also commonly face challenges related to quantization. Due to cost constraints and limited bandwidth, sensors in many CPS often send quantized measurements rather than exact values. Binary outputs from switching sensors or optical detectors are typical examples [23]. While quantization helps reduce data size and bandwidth usage, it also introduces nonlinearities that complicate system identification and control [24]. To address this challenge, researchers have developed a variety of identification methods tailored for quantized systems [25], including studies on binary sensor applications [26]–[28], worst-case estimation techniques [29], and methods aimed at achieving asymptotic efficiency [30].

Beyond quantization, Data Tampering attacks also further complicate the identification problem in CPS. Researchers have proposed various methods to address such attacks in binary-output systems. For instance, [31] developed a compensation-based algorithm that ensures strong consistency and asymptotic normality under tampering, later extending it to multi-dimensional systems [32]. To handle packet loss during transmission, [33] proposed algorithms for both known and unknown loss rates, integrating compensation mechanisms to reduce communication costs. More recently, [34] proposed a maximum likelihood estimation (MLE) approach using an iterative Expectation-Maximization algorithm for parameter identification under tampering. Furthermore, [35] explored finite impulse response (FIR) systems under random replay attacks, presenting consistent estimation algorithms and asymptotically optimal defense strategies.

Despite these advances, many existing methods rely heavily on offline processes such as empirical measure estimation or MLE, which may not be suitable for real-time CPS that continuously generate data. This gap highlights the need for online identification algorithms that can update estimates dynamically while maintaining robustness against tampering. To address this challenge, this paper extends our previous work [36] by presenting a unified identification framework. We propose both first-order and second-order recursive algorithms that handle cases with known attack strategies and scenarios requiring online estimation of tampering statistics.

The main contributions of this paper are summarized as follows:

- This paper presents a first-order gradient-based algorithm for parameter estimation under data tampering attacks. The convergence properties of the algorithm, including almost sure convergence and mean-square convergence, are established. To handle more realistic scenarios, a periodic extra-insertion defense mechanism is introduced to estimate tampering statistics during the identification process and ensure convergence under unknown attacks.
- A projected second-order quasi-Newton algorithm is proposed to handle both known and unknown tampering strategies, achieving faster convergence. Integrated into an adaptive control framework, the algorithm provides explicit asymptotic bounds on the estimation error, tracking error, and cumulative regret. These bounds are established under the weakest excitation condition in [37], without requiring persistent excitation (PE) condition.
- The proposed methods are applied to vehicle emission control, using a dataset of diesel vehicle emissions collected in Hefei, China. Transformer-based features are used as regressors in the set-valued model, where "over-limit" and "non-over-limit" cases are modeled as binary outcomes. Data tampering, such as underreporting excessive emissions, is captured via probabilistic label manipulation. The results show improved accuracy in identifying compliant and non-compliant vehicles, even under adversarial conditions.

The paper is organized as follows. Section 2 introduces the system and examines identifiability under tampering attacks. Section 3 presents a first-order gradient-based identification algorithm, establishes its convergence, and proposes a periodic insertion scheme to handle unknown attack strategies. Section 4 develops a second-order Newton-type algorithm and provides theoretical bounds on estimation error, cumulative regret, and tracking performance. Section 5 offers numerical simulations and a real-world case study on vehicle emission control. Section 6 concludes the paper and outlines future directions.

Notation: Let $\mathbb{Z}_+$ be the set of positive integers. For a positive integer k, define $[k] = \{1, 2, \ldots, k\}$. For a set $\mathcal{S} \subset \mathbb{Z}_+$ and an integer l, define $l\mathcal{S} = \{i : i = l \times j, j \in \mathcal{S}\}$. For a sequence of sets $\{\mathcal{S}_k\}_{k\geq 0}$, if $l > j$, then $\bigcup_{k=l}^{j} \mathcal{S}_k = \emptyset$.

II. Model Formulation and Identifiability

A. Model Description Under Tampering Attack

This subsection introduces the identification problem of stochastic FIR systems under tampered binary observations. Consider the following system:

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

where $\theta \in \Theta \subseteq \mathbb{R}^p$ is an unknown but time-invariant parameter vector of known dimension p, $\varphi_k \in \mathbb{R}^p$ is the regressor vector composed of current and past input signals, and w_{k+1} is a stochastic noise sequence with zero mean. Let $\mathcal{F}_k$ denote the natural filtration defined as

$$\mathcal{F}_k \triangleq \sigma\left\{\varphi_0, \ldots, \varphi_k, w_0, \ldots, w_k\right\}, \quad k \geq 0. \tag{2}$$

In this setting, the system output y_{k+1} is not directly observable. Instead, it is accessed through a binary sensor with a known threshold $C \in \mathbb{R}$, which generates the binary-valued signal:

$$s_{k+1}^0 = I_{[y_{k+1} \leq C]} = \begin{cases} 1, & y_{k+1} \leq C, \\ 0, & \text{otherwise}. \end{cases} \tag{3}$$

However, this binary signal is transmitted over a potentially compromised communication channel. Due to data tampering attacks, the received signal s_{k+1} may differ from the original s_{k+1}^0, and is modeled by the following flipping probabilities:

$$\begin{cases} \Pr\{s_k = 0 \mid s_k^0 = 1\} = p, \\ \Pr\{s_k = 1 \mid s_k^0 = 0\} = q, \end{cases} \tag{4}$$

where $p, q \in [0, 1)$ represent the attacker's flipping strategy.

Identification Objective. The objective of this paper is to design a recursive algorithm to estimate the unknown parameter θ based on the sequence of regressors $\{\varphi_k\}_{k\geq 0}$ and the possibly tampered binary observations $\{s_k\}_{k\geq}$. Further,

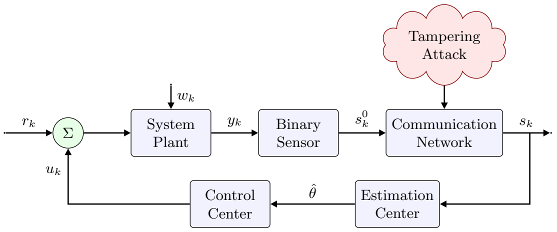


Fig. 1. Closed-loop control system flowchart with tampering attack and estimation/control centers

if control actions are required, the estimation result $\hat{\theta}$ from the estimation center is transmitted to the control center. The control center then utilizes this estimated parameter to design the controller input u_k, forming a closed-loop feedback system. The flow of information and control within this closed-loop system, incorporating potential data tampering attacks, is illustrated in Fig. 1.

Remark 2.1: In many secure control architectures, the communication network is logically divided into an "downlink" (from the sensor to the estimation center) and a "uplink" (from the control center to the actuator/execution unit). By implementing the uplink over a protected, dedicated channel—such as a private optical-fiber link, a local bus, or an encrypted virtual tunnel—one can safely assume that the control signal u_k is not subject to tampering during transmission. As a result, the integrity risk is primarily concentrated on the downlink path that conveys the binary sensor measurements s_k^0 through a potentially untrusted network [38], [39].

B. Identifiability Analysis

Let $F_k(\cdot)$ denote the conditional distribution function of the noise w_{k+1} given the filtration $\mathcal{F}_k$. Based on the law of total probability and the tampering model in (4), the conditional probability of observing $s_{k+1} = 0$ can be computed as

$$\begin{aligned} & P(s_{k+1} = 0 \mid \mathcal{F}_k) \\ =& P(s_{k+1}^0 = 1 \mid \mathcal{F}_k) \cdot P(s_{k+1} = 0 \mid s_{k+1}^0 = 1, \mathcal{F}_k) \\ & + P(s_{k+1}^0 = 0 \mid \mathcal{F}_k) \cdot P(s_{k+1} = 0 \mid s_{k+1}^0 = 0, \mathcal{F}_k) \\ =& pF_k(C - \theta^T \varphi_k) + (1-q)\left[1 - F_k(C - \theta^T \varphi_k)\right] \\ =& (p+q-1)F_k(C - \theta^T \varphi_k) + 1 - q. \end{aligned} \tag{5}$$

Similarly, the conditional probability of $s_{k+1} = 1$ is given by

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

Identifiability. This subsection discusses the identifiability of the binary-output system defined in (1)–(4). From equations (5)–(6), it can be observed that when $p + q = 1$, the dependence on the conditional distribution $F_k(\cdot)$ vanishes. In this degenerate case, the distribution of the observations becomes independent of θ, making parameter identification impossible. Therefore, a necessary condition for identifiability is: $p + q \neq 1$.

In addition, to recover the unknown parameter θ from the sequence of binary observations, the regressor sequence $\{\varphi_k\}$ must be sufficiently informative [40]. A formal excitation condition will be specified in subsequent sections, depending on the structure of the identification algorithm.

III. First-order Gradient Identification Algorithm

A. Known Attack Strategy

A first-order recursive identification algorithm for FIR systems under binary observations with data tampering was introduced in earlier work [36]. This subsection briefly reviews the algorithm, including its motivation, formulation, and main convergence results. Complete proofs are presented in the Appendix.

1) Motivation and Algorithm: Classical identification methods, such as stochastic approximation (SA), yield consistent parameter estimates as the data length increases. However, when only binary observations are available and such data may be tampered, traditional approaches are no longer applicable. To compensate for the limited information, additional assumptions-such as knowledge of the noise distribution-are required.

The key idea is based on the fact that, in system identification, the conditional variance of w_{k+1} is minimized when using the true parameter θ. For the system (1), we have:

$$\begin{aligned} \mathrm{Var}(w_{k+1} | \mathcal{F}_k) &= \mathbb{E}\left[(y_{k+1} - \theta^T \varphi_k)^2 \mid \mathcal{F}_k\right] \\ &\leq \mathbb{E}\left[(y_{k+1} - \beta^T \varphi_k)^2 \mid \mathcal{F}_k\right], \ \forall \beta \in \Theta. \end{aligned}$$

Under appropriate excitation conditions on $\{\varphi_k\}$, the equality can be held if and only if $\beta = \theta$. This observation allows us to reformulate the identification problem as the following stochastic optimization:

$$\min_{\beta \in \Theta} I(\beta) := \tfrac{1}{2}\mathbb{E}\left[(y_{k+1} - \beta^T \varphi_k)^2 \mid \mathcal{F}_k\right]. \tag{7}$$

To solve the above problem, stochastic gradient descent (SGD) can be applied by using a single sample at each step. Specifically, the parameter is updated based on the observed pair (y_{k+1}, φ_k) as:

$$\begin{aligned} \theta_{k+1} &= \theta_k - b_k \nabla_\beta I(\beta, y_{k+1}, \varphi_k)\big|_{\beta = \theta_k} \\ &= \theta_k + b_k (y_{k+1} - \theta_k^T \varphi_k) \varphi_k, \end{aligned}$$

where $\{b_k\}$ is a sequence of diminishing step sizes. This recursion corresponds to the classical first-order stochastic approximation algorithm, which is widely used in system identification [41].

In the presence of tampered binary outputs, as described by the system (1) with binary observations (3) and potential tampering (4), a similar strategy can be employed to develop a recursive identification algorithm. At each time step k, the conditional expectation of the observed binary output s_{k+1} is given by

$$\mathbb{E}(s_{k+1} \mid \mathcal{F}_k) = (1-(p+q))F_k(C - \theta^T \varphi_k) + q. \tag{8}$$

As in the case of y_{k+1}, the conditional variance of s_{k+1} is minimized when $\beta = \theta$. This also allows the identification problem to be formulated as a stochastic optimization task, using a tampered binary loss function:

$$\begin{aligned}&\min_{\beta\in\Theta} I_{TB}(\beta, s_{k+1}, \varphi_k) \qquad (9)\\ := &\tfrac{1}{2}\mathbb{E}\left(\left(s_{k+1} - [(1-(p+q))F_k(C-\beta^T\varphi_k)+q]\right)^2 \mid \mathcal{F}_k\right).\end{aligned}$$

The gradient of the loss function with respect to β is

$$\begin{aligned}&\nabla_\beta I_{TB}(\beta, s_{k+1}, \varphi_k) = -(1-(p+q))f_k(C-\beta^T\varphi_k)\\ &\times\left((1-(p+q))F_k(C-\beta^T\varphi_k)+q-s_{k+1}\right)\varphi_k.\end{aligned}$$

To ensure that $f_k(C-\beta^T\varphi_k)$ does not approach zero, it is natural to constrain β within a compact parameter set Θ. This motivates the use of a projection operator, defined as follows.

Definition 3.1: Let $\Omega \subseteq \mathbb{R}^n$ be a convex compact set. The projection operator $\Pi_\Omega(\cdot)$ is defined by

$$\Pi_\Omega(x) = \arg\min_{\omega\in\Omega}\|x-\omega\|, \quad \forall x\in\mathbb{R}^n.$$

Based on the above analysis, we propose a first-order recursive projection algorithm, referred to as the *Gradient Recursive Projection Algorithm for Tampered Binary Observations with Known Probabilities (GRP-TB-KP).*

Remark 3.1: The loss functions in (7) and (9) are not limited to variance-type or squared error forms. In general, any differentiable loss function $L(\beta, y_{k+1}, \varphi_k)$ can be used, provided it satisfies $L(\beta, y_{k+1}, \varphi_k) \geq L(\theta, y_{k+1}, \varphi_k)$, which guarantees that the minimum is achieved at the true parameter θ. Under this condition, stochastic gradient descent can still be employed to obtain consistent estimates. An important example is the negative log-likelihood function.

Algorithm 1 GRP-TB-KP

```
Require: Initial estimate $\hat{\theta}_1 \in \Theta$, step-size sequence $\{b_k\}_{k\geq 1}$,
    constant $\beta > 0$
for $k = 1, 2, \ldots$ do
    Compute:
        $\tilde{s}_{k+1} = \beta(1-(p+q))$
        $\times\left((1-(p+q))F_k(C-\hat{\theta}_k^T\varphi_k)+q-s_{k+1}\right).$   (10)
    Update parameter estimate:
        $\hat{\theta}_{k+1} = \Pi_\Theta\left\{\hat{\theta}_k + b_k\varphi_k\tilde{s}_{k+1}\right\}.$   (11)
end for
```

2) Convergence Properties: The convergence of the proposed algorithm is established under the following assumptions.

Assumption A1. The unknown parameter satisfies $\theta \in \Theta \subseteq \mathbb{R}^n$, where Θ is a convex compact set. Let $B := \sup_{\nu\in\Theta}\|\nu\|$, where $\|\cdot\|$ denotes the Euclidean norm.

Assumption A2. The conditional distribution and density functions of the observation noise sequence $\{w_k\}$, given $\mathcal{F}_k$, are known and denoted by $F_k(\cdot)$ and $f_k(\cdot)$, respectively.

Assumption A3. (a) The regressor sequence $\{\varphi_k\}$ satisfies

$$\sup_{k\geq 1}\|\varphi_k\| \leq M < \infty. \tag{12}$$

(b) There exist an integer $h \geq p$ and a constant $\delta > 0$ such that

$$\frac{1}{h}\mathbb{E}\left[\sum_{l=k}^{k+h-1}\varphi_l\varphi_l^T \middle| \mathcal{F}_{k-1}\right] \geq \delta I, \quad \forall k\geq 1, \tag{13}$$

where I is the $p\times p$ identity matrix.

Assumption A4. The step-size sequence $\{b_k\}$ satisfies:

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

Assumption A5. The conditional density functions $\{f_k(\cdot)\}$ of the noise $\{w_k\}$, given $\mathcal{F}_{k-1}$, satisfy

$$\underline{f} := \inf_{k\geq 1}\inf_{|x|\leq C+MB} f_k(x) > 0.$$

Remark 3.2: Assumption A2 ensures that the statistical properties of the observation noise are known in advance, which is a standard assumption in the literature on binary identification (e.g., [25], [42]). Assumption A3 places conditions on the input sequence $\{\varphi_k\}$. Condition (13), known as the "conditionally expected sufficiently rich condition," ensures that the regressor sequence contains enough variability to allow parameter identifiability. Compared to the classical PE condition, it is weaker but still sufficient for convergence.

Remark 3.3: Assumption A4 is standard in stochastic approximation and online optimization. These conditions ensure that the step size remains effective over time, decays gradually, and avoids premature vanishing. They strike a balance between convergence stability and parameter adaptation. Assumption A5 guarantees that the noise density does not vanish within the relevant domain, preventing degenerate cases. The constant $\underline{f}$ provides a uniform lower bound, ensuring that the noise remains sufficiently dispersed in the estimation region.

Denote the estimation error by $\tilde{\theta}_k = \hat{\theta}_k - \theta,\ k = 1, 2, \ldots$. The following theorems establish the almost sure convergence, mean-square convergence, and convergence rate of the proposed GRP-TB-KP algorithm.

Theorem 3.1 (Convergence): Consider system (1) with binary-valued observations (3) and tampering attacks (4). Under Assumptions A1–A5, the parameter estimate generated by the algorithm (10)–(11) satisfies

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

Moreover, if $\sum_{k=1}^{\infty} b_k^2 < \infty$, then the estimate $\hat{\theta}_k$ also converges almost surely to the true parameter:

$$\lim_{k\to\infty}\tilde{\theta}_k = 0, \quad \text{a.s.}$$

Theorem 3.2 (Convergence Rate): Under Assumptions (A1)-(A5), the algorithm (10)–(11) achieves the following mean-square convergence rates:

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

Proof: See Appendix I. ■

B. The Attack Strategy is Unknown

In practice, the tampering probabilities (p, q), which characterize the attack strategy, are typically unknown to the estimator, making existing methods that require their values inapplicable. To address this, we adopt a more practical approach where (p, q) are estimated online along with the system parameters. Following Theorem 3.1 and the idea in [31], we substitute the real-time estimates of (p, q) into the identification algorithm.

1) Algorithm Design: We now describe the design of the algorithm, starting with its core idea.

Main idea. The approach introduces a known binary sequence (consisting of 0s and 1s) into the transmission channel. By comparing the received sequence with the original one, the estimation center can estimate the tampering probabilities (p, q) using the law of large numbers.

To implement this, we propose a periodic extra-insertion scheme. Known binary signals are inserted at fixed locations, and the corresponding received signals are used to estimate the tampering behavior before proceeding with identification. Specifically, given a period T, define two non-empty, disjoint subsets $\mathcal{T}_0$ and $\mathcal{T}_1$ of the index set $\mathcal{T} = \{1, 2, \ldots, T\}$.

During the l-th period, for each time $k \in \{(l-1)T + 1, \ldots, lT\}$, the binary output s_k^0 is first transmitted and possibly tampered as described in (4). In addition, if $k - lT \in \mathcal{T}_0$, a known signal $s^0_{k+1/2} = 0$ is transmitted; if $k - lT \in \mathcal{T}_1$, then $s^0_{k+1/2} = 1$ is sent. The received signal $s_{k+1/2}$ then follows the tampering model:

$$\begin{cases} \Pr\{s_{k+1/2} = 0 \mid s^0_{k+1/2} = 1\} = p, \\ \Pr\{s_{k+1/2} = 1 \mid s^0_{k+1/2} = 0\} = q. \end{cases} \tag{14}$$

At each time k, define the index sets

$$\begin{aligned} \mathcal{S}_k^0 &= \left(\bigcup_{i=1}^{l-1} i\mathcal{T}_0\right) \bigcup \left([k-(l-1)T] \cap \mathcal{T}_0\right), \\ \mathcal{S}_k^1 &= \left(\bigcup_{i=1}^{l-1} i\mathcal{T}_1\right) \bigcup \left([k-(l-1)T] \cap \mathcal{T}_1\right). \end{aligned} \tag{15}$$

The tampering probabilities can then be estimated as

$$\begin{aligned} \hat{q}_k &= \frac{1}{|\mathcal{S}_k^0|} \sum_{i \in \mathcal{S}_k^0} s_{i+1/2} \\ &= \frac{1}{l|\mathcal{T}_0| + |[k-(l-1)T] \cap \mathcal{T}_0|} \sum_{i \in \mathcal{S}_k^0} s_{i+1/2}, \\ \hat{p}_k &= \frac{1}{|\mathcal{S}_k^1|} \sum_{i \in \mathcal{S}_k^1} (1 - s_{i+1/2}) \\ &= \frac{1}{l|\mathcal{T}_1| + |[k-(l-1)T] \cap \mathcal{T}_1|} \sum_{i \in \mathcal{S}_k^1} (1 - s_{i+1/2}). \end{aligned} \tag{16}$$

These estimates $\hat{p}_k$ and $\hat{q}_k$ are then used in the identification algorithm to replace the unknown true values at time k.

Based on the above discussion, we propose the *Gradient Recursive Projection Algorithm for Tampered Binary Observations with Unknown Probabilities (GRP-TB-UP)*.

Remark 3.4: In Algorithm 2, the sets $\mathcal{T}_0$ and $\mathcal{T}_1$ are designed for online estimation of the attack probabilities p and q, and are predefined and known to both the system and the estimation center. The sets $\mathcal{S}^0$ and $\mathcal{S}^1$ correspond to $\mathcal{S}_k^0$ and $\mathcal{S}_k^1$ as defined in (15). The estimation center incrementally records the positions of inserted signals and their corresponding received values, which are then used to estimate the tampering probabilities. The entire algorithm is recursive and operates in an online manner.

Remark 3.5: In practice, the fixed insertion sets $\mathcal{T}_0$ and $\mathcal{T}_1$ may become known to the attacker. To enhance robustness, one may adopt varying insertion positions over time, denoted by $\mathcal{T}_{0l}$ and $\mathcal{T}_{1l}$ for the l-th period. As long as the condition

$$\sum_{j=1}^{l} |\mathcal{T}_{0j}| = O(l), \quad \sum_{j=1}^{l} |\mathcal{T}_{1j}| = O(l)$$

is satisfied, the theoretical guarantees of the algorithm remain valid.

Algorithm 2 GRP-TB-UP

```
Require: Period $T$; disjoint subsets $\mathcal{T}_0, \mathcal{T}_1 \subseteq \mathcal{T} = \{1, 2, \ldots, T\}$; initialization: $\mathcal{S}^0 = \mathcal{S}^1 = \emptyset$, parameter estimate $\hat{\theta}_1^u \in \Theta$.
for $l = 1, 2, \ldots$ do
    for $k = (l-1)T + 1, \ldots, lT$ do
        Step 1: Data reception
        Transmit the observed data $s_k^0$
        if $k - (l-1)T \in \mathcal{T}_0$ then
            Transmit $s^0_{k+1/2} = 0$; receive $s_{k+1/2}$; update $\mathcal{S}_0 = \mathcal{S}_0 \cup \{k\}$
        else if $k - (l-1)T \in \mathcal{T}_1$ then
            Transmit $s^0_{k+1/2} = 1$; receive $s_{k+1/2}$; update $\mathcal{S}_1 = \mathcal{S}_1 \cup \{k\}$
        end if
        Step 2: Attack probability estimation
        Compute the estimates of $p$ and $q$:
            $\hat{q}_k = \frac{1}{|\mathcal{S}^0|} \sum_{i \in \mathcal{S}^0} s_{i+1/2}, \quad \hat{p}_k = \frac{1}{|\mathcal{S}^1|} \sum_{i \in \mathcal{S}^1} (1 - s_{i+1/2})$
        Step 3: Recursive projected identification
        Compute:
            $\tilde{s}^u_{k+1} = \beta(1 - (\hat{p}_k + \hat{q}_k)) \times \Big((1 - (\hat{p}_k + \hat{q}_k)) F_k(C - \hat{\theta}_k^{uT} \varphi_k) + \hat{q}_k - s_{k+1}\Big)$   (17)
        Update the parameter estimate:
            $\hat{\theta}^u_{k+1} = \Pi_\Theta \left\{ \hat{\theta}^u_k + b_k \varphi_k \tilde{s}^u_{k+1} \right\}$   (18)
    end for
end for
```

2) Convergence analysis: Denote the estimation error by $\tilde{\theta}_k^u = \hat{\theta}_k^u - \theta,\ k = 1, \ldots$. Due to the additional transmission of inserted data, the filtration $\mathcal{F}_k$ in (2) is modified as

$$\mathcal{F}_k \triangleq \sigma\left\{\varphi_i, w_i, s_i\right\}_{i<k+1}, \quad k \geq 0,$$

which ensures that $s_{k+1/2}$ is $\mathcal{F}_k$-measurable.

The following result establishes the convergence of Algorithm 2.

Theorem 3.3: Consider the system (1) with binary-valued observation (3), under the defense scheme (14)–(15) and the data tampering model (4). Let the step size be $b_k = \frac{1}{k^\gamma}$ with $1/2 < \gamma \leq 1$. If Assumptions A1–A3 and A5 hold, then Algorithm 2 converges almost surely.

Proof: From (14), we have

$$P\{s_{k+1/2} = 1\} = \begin{cases} q, & k \mod T \in \mathcal{T}_0, \\ 1-p, & k \mod T \in \mathcal{T}_1. \end{cases} \tag{19}$$

Thus, by the law of large numbers and (15)–(16), it follows that

$$\hat{q}_k \to q, \quad \hat{p}_k \to p, \quad \text{a.s. as } k \to \infty. \tag{20}$$

In addition, by the law of the iterated logarithm,

$$\begin{aligned} |q_k - q| &= O\left(\sqrt{\frac{\log\log\lfloor\frac{k}{T}\rfloor}{\lfloor\frac{k}{T}\rfloor}}\right) = O\left(\sqrt{\frac{\log\log k}{k}}\right), \\ |p_k - p| &= O\left(\sqrt{\frac{\log\log k}{k}}\right). \end{aligned} \tag{21}$$

Define the error term $\epsilon_{k+1} = \tilde{s}^u_{k+1} - \tilde{s}_{k+1}$. Then from (10), (17), and the above convergence rates,

$$\epsilon_{k+1} \to 0, \quad |\epsilon_{k+1}| = O\left(\sqrt{\frac{\log\log k}{k}}\right). \tag{22}$$

Now consider the update rule in (18). Using the non-expansiveness of projection operators, and Proposition I.1 in Appendix I, we get

$$\begin{aligned} \|\tilde{\theta}^u_{k+1}\|^2 &= \left\|\Pi_\Theta\left\{\hat{\theta}^u_k + b_k\varphi_k\tilde{s}_{k+1} + b_k\varphi_k\epsilon_{k+1}\right\} - \theta\right\|^2 \\ &\leq \left\|\tilde{\theta}^u_k + b_k\varphi_k\tilde{s}_{k+1} + b_k\varphi_k\epsilon_{k+1}\right\|^2 \\ &= \|\tilde{\theta}^u_k\|^2 + 2b_k\tilde{s}_{k+1}\varphi_k^T\tilde{\theta}^u_k + O(b_k^2) + O(b_k\epsilon_{k+1}). \end{aligned} \tag{23}$$

Substituting (22) and noting that $b_k = \frac{1}{k^\gamma}$ with $\gamma > \frac{1}{2}$, we obtain

$$\begin{aligned} &\|\tilde{\theta}^u_{k+1}\|^2 \\ &\leq \|\tilde{\theta}^u_k\|^2 + 2b_k\tilde{s}_{k+1}\varphi_k^T\tilde{\theta}^u_k + O\left(\frac{1}{k^{2\gamma}}\right) + O\left(\frac{\sqrt{\log\log k}}{k^{\gamma+1/2}}\right) \\ &= \|\tilde{\theta}^u_k\|^2 + 2b_k\tilde{s}_{k+1}\varphi_k^T\tilde{\theta}^u_k + O\left(\frac{\sqrt{\log\log k}}{k^{\gamma+1/2}}\right). \end{aligned} \tag{24}$$

Let $B_k := \frac{\sqrt{\log\log k}}{k^{\gamma+1/2}}$. Clearly, $B_k = O(B_{k+1})$ and $b_k^2 = O(B_k)$. From (17), we have

$$\mathbb{E}[\tilde{s}_{k+1} \mid \mathcal{F}_k] = \beta(1-p-q)^2\left(F(C - \varphi_k^T\hat{\theta}^u_k) - F(C - \varphi_k^T\theta)\right). \tag{25}$$

Proceeding as in Theorem 3.1, we derive

$$\begin{aligned} &\mathbb{E}[\|\tilde{\theta}^u_{k+h}\|^2] \\ &= \left(1 - 2\beta(1-p-q)^2\underline{f}\delta\sum_{l=k}^{k+h-1} b_l\right)\mathbb{E}[\|\tilde{\theta}^u_k\|^2] + O(B_{k+h}). \end{aligned} \tag{26}$$

Now observe that $\sum_{k=1}^\infty b_k = \infty$ and

$$\lim_{k\to\infty}\frac{B_k}{\sum_{l=k-h}^{k-1} b_{l+1}} \leq \lim_{k\to\infty}\frac{B_k}{hb_k} = \lim_{k\to\infty}\frac{\sqrt{\log\log k}}{hk^{1/2}} = 0.$$

Then, by Lemma I.1, we conclude $\lim_{k\to\infty}\mathbb{E}[\|\tilde{\theta}^u_k\|^2] = 0$. Furthermore, using (24) and [41, Lemma 1.2.2], since

$$\mathbb{E}[\|\tilde{\theta}^u_{k+1}\|^2 \mid \mathcal{F}_k] \leq \|\tilde{\theta}^u_k\|^2 + O(B_k), \ \sum_{k=1}^\infty B_k < \infty,$$

we obtain that $\|\tilde{\theta}^u_k\|$ converges almost surely to a finite limit. Combined with the fact that $\mathbb{E}[\|\tilde{\theta}^u_k\|^2] \to 0$, we conclude that $\tilde{\theta}^u_k \to 0$ almost surely. ■

IV. Second-order Newton Identification Algorithm

In the previous sections, we proposed gradient-based recursive projection algorithms for both known and unknown attack strategies. These methods adopt scalar step sizes and rely on first-order information, which facilitates implementation. However, as pointed out in [43], such first-order methods often suffer from slow convergence, especially in high-dimensional problems or when parameter sensitivities vary significantly. This limitation arises from their inability to utilize curvature information in the cost function.

To address this, we introduce second-order Newton-type algorithms that employ matrix-based gains to adaptively scale the update direction. These algorithms incorporate curvature information and typically achieve faster convergence and improved estimation accuracy.

A. The Attack Strategy is Known

We first consider the case where the tampering probabilities (p, q) are known. To estimate the unknown parameter θ in the binary-valued system (1)–(3) under the attack model (4), we develop a quasi-Newton recursive identification algorithm. A projection step is included to ensure the boundedness of the estimates and to guarantee convergence. Before presenting the algorithm, we introduce the following definitions.

Definition 4.1: For the Euclidean space $\mathbb{R}^p$ ($p \geq 1$), the weighted norm $\|\cdot\|_Q$ associated with a positive definite matrix Q is defined as

$$\|x\|_Q^2 = x^TQx, \quad \forall x \in \mathbb{R}^p. \tag{27}$$

Definition 4.2: For a convex compact set $\mathcal{D} \subseteq \mathbb{R}^p$ and a positive definite matrix Q, the projection operator $\Pi_Q(\cdot)$ is defined as

$$\Pi_Q(x) = \arg\min_{\omega\in\mathcal{D}}\|x - \omega\|_Q, \quad \forall x \in \mathbb{R}^p. \tag{28}$$

Remark 4.1: The projection operator $\Pi_Q(\cdot)$ is non-expansive in the Euclidean norm:

$$\|\Pi_Q(x) - \Pi_Q(y)\| \leq \|x - y\|, \quad \forall x, y \in \mathbb{R}^p. \tag{29}$$

We now present the proposed *Newton-type Recursive Projection Algorithm for Tampered Binary observations with Known Probabilities (NRP-TB-KP)*, as summarized in Algorithm 3.

Algorithm 3 NRP-TB-KP

```
Initialize: $\hat{\theta}_1 \in \Theta$, $P_1 > 0$, and $\beta_0 = \operatorname{sign}(1-(p+q)) \min\left\{1, \inf_{|x|\le LM+C} |1-(p+q)|f_1(x)\right\}$,
for $k = 1, 2, \dots$ do
    $\beta_k = \operatorname{sign}(1-(p+q))\times \min\left\{|\beta_{k-1}|, \inf_{|x|\le LM+C} |1-(p+q)|f_{k+1}(x)\right\}$
    $a_k = \frac{1}{1+\beta_k^2\varphi_k^T P_k \varphi_k}$
    $\tilde{s}_{k+1} = (1-(p+q))F_k(C - \hat{\theta}_k^T\varphi_k) + q - s_{k+1}$
    $P_{k+1} = P_k - \beta_k^2 a_k P_k \varphi_k \varphi_k^T P_k$
    $\hat{\theta}_{k+1} = \Pi_{P_{k+1}^{-1}}\left(\hat{\theta}_k + a_k\beta_k P_k\varphi_k\tilde{s}_{k+1}\right)$
end for
```

Remark 4.2: By the matrix inversion formula [41], the inverse of P_k can be updated recursively as

$$P_{k+1}^{-1} = P_k^{-1} + \beta_k^2 \varphi_k \varphi_k^T. \tag{30}$$

Since P_1 is positive definite, P_k^{-1} remains positive definite for all k, which ensures that the projection operator $\Pi_{P_k^{-1}}$ in Algorithm 3 is well-defined.

Since φ_k is $\mathcal{F}_k$-measurable, the conditional expectation of y_{k+1} given $\mathcal{F}_k$ is

$$\mathbb{E}(y_{k+1} \mid \mathcal{F}_k) = \theta^T\varphi_k + \mathbb{E}(w_{k+1} \mid \mathcal{F}_k), \tag{31}$$

which serves as the optimal predictor in the mean-square sense. Replacing the true parameter with the estimate $\hat{\theta}_k$, the adaptive predictor becomes

$$\hat{y}_{k+1} = \hat{\theta}_k^T\varphi_k + \mathbb{E}(w_{k+1} \mid \mathcal{F}_k). \tag{32}$$

Letting $\tilde{\theta}_k = \theta - \hat{\theta}_k$, the instantaneous regret, defined as the squared deviation between the optimal and adaptive predictors, is

$$R_k = \left(\mathbb{E}(y_{k+1} \mid \mathcal{F}_k) - \hat{y}_{k+1}\right)^2 = \left(\tilde{\theta}_k^T\varphi_k\right)^2. \tag{33}$$

A small value of R_k is desirable and plays a crucial role in evaluating the performance of adaptive control algorithms.

The following three theorems summarize the main theoretical results of this subsection. Without requiring PE condition, we establish asymptotic bounds for the estimation error, cumulative regret, and tracking performance under Algorithm 3.

Theorem 4.1 (Estimation Error Bound): Under Assumptions A1, A2, A3(a), and A5, the estimation error generated by Algorithm 3 satisfies

$$\|\tilde{\theta}_{n+1}\|^2 = O\left(\frac{\log \lambda_{\max}(P_{n+1}^{-1})}{\lambda_{\min}(P_{n+1}^{-1})}\right), \quad \text{a.s.} \tag{34}$$

where $\tilde{\theta}_k = \theta - \hat{\theta}_k$.

Theorem 4.2 (Regret Bound): Let Assumptions A1, A2, A3(a), and A5 hold. Then the cumulative regret, defined by $R_k = (\tilde{\theta}_k^T\varphi_k)^2$, satisfies

$$\sum_{k=0}^{n} R_k = O\left(\frac{\log |P_{n+1}^{-1}|}{\beta_n^2}\right), \quad \text{a.s.}$$

Remark 4.3: According to the update rule of β_n in Algorithm 3, the regret bound becomes unbounded as $1-(p+q) \to 0$. This is expected, since when $p+q \to 1$, it becomes increasingly difficult to distinguish between tampered and genuine signals, which severely undermines the identifiability of the system.

Theorem 4.3 (Tracking Error in Adaptive Control): Under the assumptions of Theorem 4.2, suppose that the noise densities $\{f_k(x)\}$ satisfy

$$\sup_k \mathbb{E}\left[|w_k|^\alpha \mid \mathcal{F}_{k-1}\right] < \infty, \quad \text{a.s., for some } \alpha > 4, \tag{35}$$

and the regressor φ_k is designed such that

$$\hat{\theta}_k^T\varphi_k + \int_{-\infty}^{\infty} x f_k(x)\,dx = y_{k+1}^*, \tag{36}$$

for any bounded reference signal $\{y_{k+1}^*\}$. Then the average tracking error

$$J_n = \frac{1}{n}\sum_{k=0}^{n-1}(y_{k+1} - y_{k+1}^*)^2 \tag{37}$$

satisfies

$$\left|J_n - \frac{1}{n}\sum_{k=1}^{n}\sigma_k^2\right| = O\left(\sqrt{\frac{\log\log n}{n}}\right), \quad \text{a.s.,} \tag{38}$$

where $\sigma_k^2 = \mathbb{E}\left[\left(w_k - \mathbb{E}(w_k \mid \mathcal{F}_{k-1})\right)^2 \mid \mathcal{F}_{k-1}\right]$.

Proof: The proofs of Theorems 4.1–4.3 are provided in Appendix II. ■

B. The Attack Strategy is Unknown

We now turn to the case where the tampering parameters (p, q) are unknown. Following a similar framework to the first-order case, we estimate p and q online in parallel with the system identification process. The proposed *Newton-type Recursive Projection Algorithm for Tampered Binary observations with Unknown Probabilities (NRP-TB-UP)* is presented in Algorithm 4. This algorithm offers similar theoretical guarantees to the case with known tampering parameters.

Theorem 4.4: Consider system (1) with binary-valued observations (3), operating under the defense scheme (14)–(15) and subjected to the data tampering attack (4). Suppose that Assumptions A1, A2, A3(a), and A5 hold. Then, the estimation error produced by Algorithm 4 satisfies

$$\|\tilde{\theta}_{n+1}^u\|^2 = O\left(\frac{\log \lambda_{\max}(P_{n+1}^{-1})}{\lambda_{\min}(P_{n+1}^{-1})}\right), \text{ a.s.,} \tag{39}$$

where $\tilde{\theta}_k^u = \theta - \hat{\theta}_k^u$. Moreover, the cumulative regret, defined as $R_k = (\tilde{\theta}_k^{uT}\varphi_k)^2$, admits the bound

$$\sum_{k=0}^{n} R_k = O\left(\frac{\log |P_{n+1}^{-1}|}{\beta_n^2}\right), \quad \text{a.s.} \tag{40}$$

Proof: Define the stochastic Lyapunov function:

$$V_k^u = \tilde{\theta}_k^{uT} P_k^{-1} \tilde{\theta}_k^u.$$

Algorithm 4 NRP-TB-UP

```
Require: Period $T$, disjoint subsets $\mathcal{T}_0, \mathcal{T}_1 \subseteq \mathcal{T} = \{1, 2, \dots, T\}$, $\mathcal{S}^0 = \mathcal{S}^1 = \emptyset$, initial estimate $\hat{\theta}_1 \in \Theta$, $P_1 > 0$.
for $l = 1, 2, \dots$ do
    for $k = (l-1)T + 1, \dots, lT$ do
        Step 1: Data reception
        Transmit the observed data $s_k^0$
        if $k - (l-1)T \in \mathcal{T}_0$ then
            Transmit $s_{k+1/2}^0 = 0$, receive $s_{k+1/2}$, and update $\mathcal{S}^0 = \mathcal{S}^0 \cup \{k\}$
        else if $k - (l-1)T \in \mathcal{T}_1$ then
            Transmit $s_{k+1/2}^0 = 1$, receive $s_{k+1/2}$, and update $\mathcal{S}^1 = \mathcal{S}^1 \cup \{k\}$
        end if
        Step 2: Online estimation of $p, q$
            $\hat{q}_k = \frac{1}{|\mathcal{S}^0|} \sum_{i \in \mathcal{S}^0} s_{i+1/2}, \quad \hat{p}_k = \frac{1}{|\mathcal{S}^1|} \sum_{i \in \mathcal{S}^1} (1 - s_{i+1/2})$
        Step 3: Recursive parameter update
            $\beta_k = \operatorname{sign}(1 - (\hat{p}_k + \hat{q}_k))$
            $\cdot \min\left\{ |\beta_{k-1}|, \inf_{|x| \le LM + C} f_{k+1}(x)(1 - (\hat{p}_k + \hat{q}_k)) \right\}$
            $a_k = \frac{1}{1 + \beta_k^2 \varphi_k^T P_k \varphi_k}$
            $\tilde{s}_{k+1}^u = (1 - (\hat{p}_k + \hat{q}_k)) F_k(C - \hat{\theta}_k^{uT} \varphi_k) + \hat{q}_k - s_{k+1}$
            $P_{k+1} = P_k - \beta_k^2 a_k P_k \varphi_k \varphi_k^T P_k$
            $\hat{\theta}_{k+1}^u = \Pi_{P_{k+1}^{-1}} \left( \hat{\theta}_k^u + a_k \beta_k P_k \varphi_k \tilde{s}_{k+1}^u \right)$
    end for
end for
```

Let $\epsilon_{k+1} = \tilde{s}_{k+1} - \tilde{s}_{k+1}^u$, where $\tilde{s}_{k+1}$ is given in Algorithm 3. Define

$$
\begin{aligned}
\varepsilon_{k+1} &= (1 - (p+q)) F_k(C - \theta^T \varphi_k) + q - s_{k+1}, \quad (41) \\
\psi_k &= (1 - (p+q)) \\
&\quad \cdot \left(F_{k+1}(C - \hat{\theta}_k^T \varphi_k) - F_{k+1}(C - \theta^T \varphi_k) \right). \quad (42)
\end{aligned}
$$

Following the argument in (59), we obtain

$$
\begin{aligned}
V_{k+1}^u \le\ & \tilde{\theta}_k^{uT} P_k^{-1} \tilde{\theta}_k^u - 2\beta_k \tilde{\theta}_k^{uT} \varphi_k \psi_k + \beta_k^2 (\tilde{\theta}_k^{uT} \varphi_k)^2 \\
& + 2a_k \beta_k^2 (\psi_k + \epsilon_{k+1}) \varphi_k^T P_k \varphi_k \varepsilon_{k+1} \\
& - 2\beta_k \varphi_k^T \tilde{\theta}_k^u (\varepsilon_{k+1} - \epsilon_{k+1}) + a_k \beta_k^2 \varphi_k^T P_k \varphi_k \varepsilon_{k+1}^2 \\
& + a_k \beta_k^2 \varphi_k^T P_k \varphi_k \epsilon_{k+1}^2 - 2a_k \beta_k^2 \varphi_k^T P_k \varphi_k \psi_k \epsilon_{k+1} \\
& + a_k \beta_k^2 \varphi_k^T P_k \varphi_k. \qquad (43)
\end{aligned}
$$

According to (21), there exists $K \in \mathbb{N}_+$ such that for all $k > K$, we have $\operatorname{sign}(1 - (p_k + q_k)) = \operatorname{sign}(1 - (p+q))$, and $|(1 - (p_k + q_k))| < \frac{4}{3}|1 - (p+q)|$. Based on the definition of β_k in Algorithm 4, and using (57) together with the mean value theorem, it follows that

$$
\begin{aligned}
& 2\beta_k \tilde{\theta}_k^{uT} \varphi_k \psi_k = 2\beta_k (\tilde{\theta}_k^{uT} \varphi_k)^2 f_k(\xi_k)(1 - (p+q)) \\
\ge & \frac{2|p+q-1|}{|(1 - (p_k + q_k))|} \beta_k^2 (\tilde{\theta}_k^{uT} \varphi_k)^2 \ge \frac{3}{2} \beta_k^2 (\tilde{\theta}_k^{uT} \varphi_k)^2, \qquad (44)
\end{aligned}
$$

where ξ_k lies between $C - \theta^T \varphi_k$ and $C - \hat{\theta}_k^T \varphi_k$.

Furthermore, by (22) and the fact that $|\psi_k| < 1$, for sufficiently large n, we have

$$
\begin{aligned}
& a_k \beta_k^2 \varphi_k^T P_k \varphi_k \epsilon_{k+1}^2 - 2a_k \beta_k^2 \varphi_k^T P_k \varphi_k \psi_k \epsilon_{k+1} + a_k \beta_k^2 \varphi_k^T P_k \varphi_k \\
& = O(a_k \beta_k^2 \varphi_k^T P_k \varphi_k), \qquad (45) \\
& a_k \beta_k^2 (\psi_k + \epsilon_{k+1}) \varphi_k^T P_k \varphi_k \varepsilon_{k+1} = O(a_k \beta_k^2 \psi_k \varphi_k^T P_k \varphi_k \varepsilon_{k+1}).
\end{aligned}
$$

Summing both sides of (43) and applying (44)–(45), we obtain

$$
\begin{aligned}
V_{n+1}^u \le\ & V_0^u - \frac{1}{2} \sum_{k=0}^{n} \beta_k^2 (\tilde{\theta}_k^{uT} \varphi_k)^2 + 2 \sum_{k=0}^{n} \beta_k \varphi_k^T \tilde{\theta}_k^u \epsilon_{k+1} \\
& \underbrace{+ 2O\left(\sum_{k=0}^{n} a_k \beta_k^2 \psi_k \varphi_k^T P_k \varphi_k \varepsilon_{k+1} \right) - 2 \sum_{k=0}^{n} \beta_k \varphi_k^T \tilde{\theta}_k^u \varepsilon_{k+1}}_{\mathrm{I}} \\
& \underbrace{+ O\left(\sum_{k=0}^{n} a_k \beta_k^2 \varphi_k^T P_k \varphi_k \right) + \sum_{k=0}^{n} a_k \beta_k^2 \varphi_k^T P_k \varphi_k \varepsilon_{k+1}^2}_{\mathrm{II}}. \qquad (46)
\end{aligned}
$$

From (41) and (8), we know that

$$
\mathbb{E}(\varepsilon_{k+1} \mid \mathcal{F}_k) = 0, \ \sup_k \mathbb{E}\left[|\omega_{k+1}|^2 \mid \mathcal{F}_k \right] < \infty, \quad \text{a.s.,}
$$

which implies that $\{\varepsilon_k, \mathcal{F}_k\}$ forms a martingale difference sequence with finite second moment.

Following the arguments in [44], and using Assumption A3(a) $\sup_{k \ge 1} \|\varphi_k\| \le M < \infty$ along with (22), we have

$$
\begin{aligned}
& 2 \sum_{k=0}^{n} \beta_k \varphi_k^T \tilde{\theta}_k^u \epsilon_{k+1} = o\left(\sum_{k=0}^{n} \beta_k^2 (\tilde{\theta}_k^{uT} \varphi_k)^2 \right) + O(1), \\
& \mathrm{I} = o\left(\sum_{k=0}^{n} \beta_k^2 (\tilde{\theta}_k^{uT} \varphi_k)^2 \right) + O(1), \\
& \mathrm{II} = O\left(\log |P_{n+1}^{-1}| \right).
\end{aligned}
$$

Combining the above estimates yields

$$
\tilde{\theta}_{n+1}^{uT} P_{n+1}^{-1} \tilde{\theta}_{n+1}^u + \sum_{k=0}^{n} \beta_k^2 (\tilde{\theta}_k^{uT} \varphi_k)^2 = O(\log |P_{n+1}^{-1}|), \quad \text{a.s.}
$$

Finally, since $\{\beta_k\}$ is non-increasing, the results in (39) and (40) follow. ■

Similar to Theorem 4.3, Theorem 4.4 also implies the following result on the tracking error in adaptive control of binary FIR systems under unknown tampering attacks.

Theorem 4.5: Consider system (1) with binary-valued observations (3), operating under the defense scheme (14)–(15) and subjected to the data tampering attack (4). Suppose the conditions of Theorem 4.3 hold, and that the regressor φ_k is constructed as in Theorem 4.3 for a bounded reference signal $\{y_{k+1}^*\}$. Then, the average tracking error satisfies

$$
\left| J_n - \frac{1}{n} \sum_{k=1}^{n} \sigma_k^2 \right| = O\left(\sqrt{\frac{\log \log n}{n}} \right), \quad \text{a.s.}
$$

where $\sigma_k^2 = \mathbb{E}\left[(w_k - \mathbb{E}(w_k \mid \mathcal{F}_{k-1}))^2 \mid \mathcal{F}_{k-1} \right]$.

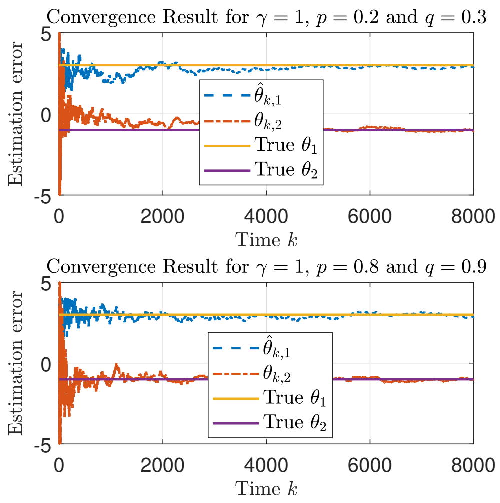


Fig. 2. Convergence of the estimation shown by a trajectory of $\hat{\boldsymbol{\theta}}_{n+1}$.

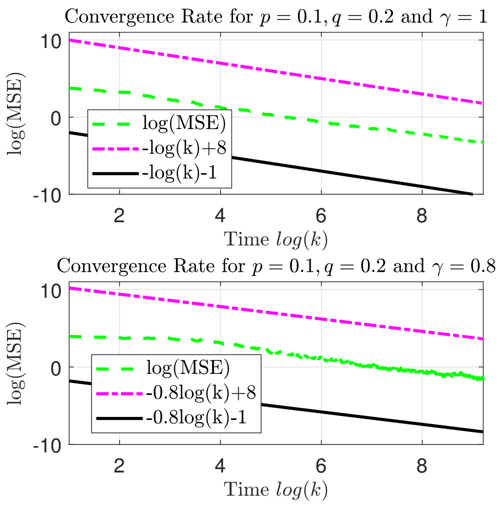


Fig. 3. Convergence rate of the estimation shown by a trajectory of $\boldsymbol{k\tilde\theta_k^T\tilde\theta_k/\ln k}$.

V. Numerical Simulations and Practical Case Study

A. Numerical simulations

This section presents three numerical simulations to validate the convergence of the first-order and second-order algorithms, as well as the tracking error bounds in adaptive control.

Example 1. 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 (10)-(11) 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 2 and 3 present the estimation results of the algorithm for attack strategy shown in (4) as $p = 0.2, q = 0.3$ and $p = 0.8, q = 0.9$, respectively. From Figure 2, 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 (Theorem 3.1). Figure 3 shows the convergence rate for $\gamma = 1$ and $\gamma = 0.8$, validating the results derived in Theorem 3.2. We also consider the case where the attack strategy is unknown, and the attack probabilities p and q are estimated online. The estimator periodically updates p and q using auxiliary data collected at the time indices $\mathcal{T}_0 = \{1, 3, 5, 7, 9\}$ and $\mathcal{T}_1 = \{2, 4, 6, 8, 10\}$ within each period of length $T = 20$. In the simulation, two attack scenarios are tested: $(p, q) = (0.2, 0.3)$ and $(p, q) = (0.8, 0.9)$. The simulation results are shown in Figure 4, where the parameter estimates $\hat\theta_k$ converge to the true values, and the online estimates of p and q gradually approach their true values as the sample size increases. These results validate the Theorem 3.3.

Example 2. This example is to validate the theoretical analysis for Algorithm 4. We conduct numerical simulations with the same setting as Example 1. The regression vector is defined as $\varphi_k = [u_k, u_{k-1}]^\top$, with $u_k \sim \mathcal{N}(0, \sigma_k^2)$ and $\sigma_k = k^{-1/8}$. At each time step, the binary observation s_k is possibly flipped by an adversary with known probabilities $p = 0.1$ and $q = 0.2$. The online estimates of p and q are computed based on partial feedback under a periodic schedule. Figure 5 presents the simulation results over $N = 20000$ time steps. Subplot (a) shows the parameter estimates $\hat\theta_1$ and $\hat\theta_2$ converging to their true values. Subplot (b) presents the squared estimation error $\|\tilde\theta_k\|^2$ along with its theoretical bound $\mathcal{O}(\log k/k3/4)$, validating Theorem 4.4. Subplot (c) depicts the cumulative regret $\sum_{k=0}^n R_k$ and confirms the bound $\mathcal{O}(\log k/\beta_n^2)$ as stated in Theorem 4.4. Subplot (d) demonstrates the convergence of the online estimates of the attack probabilities p and q, verifying the effectiveness of the periodic extra-insertion scheme.

Example 3. We also consider the binary observation model with unknown parameter $\theta = [3, -1]^\top$, Gaussian noise $w_k \sim \mathcal{N}(0, 1)$, and regressor $\varphi_k = [u_k,\ u_{k-1}]^\top$. We implement the NRP-TB-UP algorithm to estimate θ online using only binary feedback, while simultaneously designing a control input u_k such that the system output y_k tracks a reference signal $y_k^* = 4\sin(2\pi k/18000)$. The control law solves

$$\hat\theta_k^\top\varphi_k + \int x f_k(x)\,dx = y_{k+1}^*,$$

where f_k is the density of w_k. According to Theorem 4.5, the average tracking error $J_n = \frac{1}{n}\sum_{k=1}^n (y_k - y_k^*)^2$ is expected to satisfy $|J_n - \sigma^2| = \mathcal{O}\left(\sqrt{\frac{\log\log n}{n}}\right)$, a.s. To verify this, we test two attack settings: $(p, q) = (0.2, 0.3)$ and $(p, q) = (0.8, 0.9)$. Each scenario uses $N = 20000$ steps, with initialization $\hat\theta_1 = [1, 1]^\top$ and random u_0. Results are shown in Figure 6.

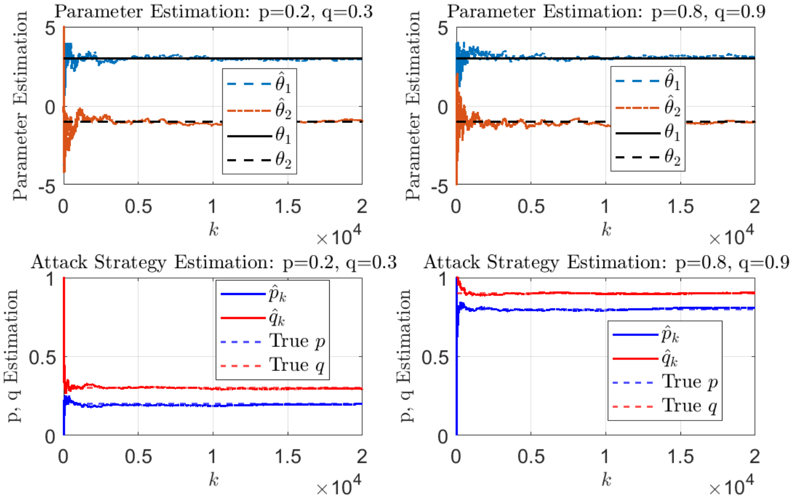


Fig. 4. Convergence of parameter estimates $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ and attack probability estimates $(\hat{\boldsymbol{p}}_{\boldsymbol{k}}, \hat{\boldsymbol{q}}_{\boldsymbol{k}})$ under unknown attack strategies with $(\boldsymbol{p}, \boldsymbol{q}) = (\mathbf{0.2}, \mathbf{0.3})$ and $(\boldsymbol{p}, \boldsymbol{q}) = (\mathbf{0.8}, \mathbf{0.9})$.

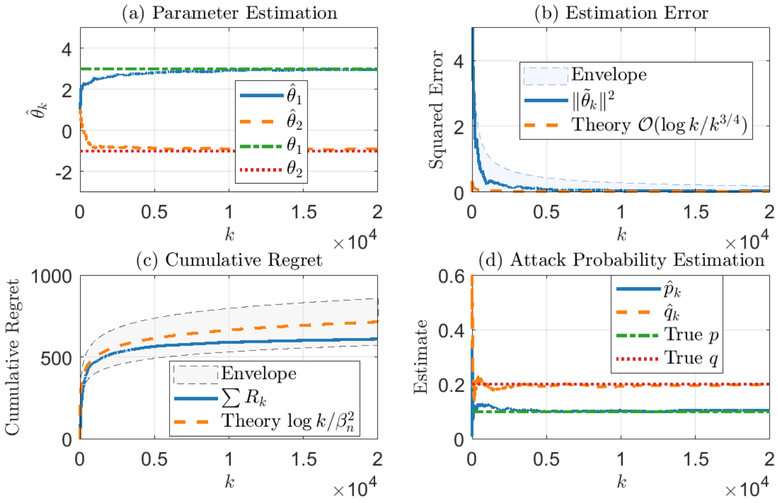


Fig. 5. Simulation results of Algorithm 4. (a) Parameter estimates; (b) Estimation error and its theoretical bound; (c) Cumulative regret vs theoretical rate; (d) Online estimation of attack probabilities.

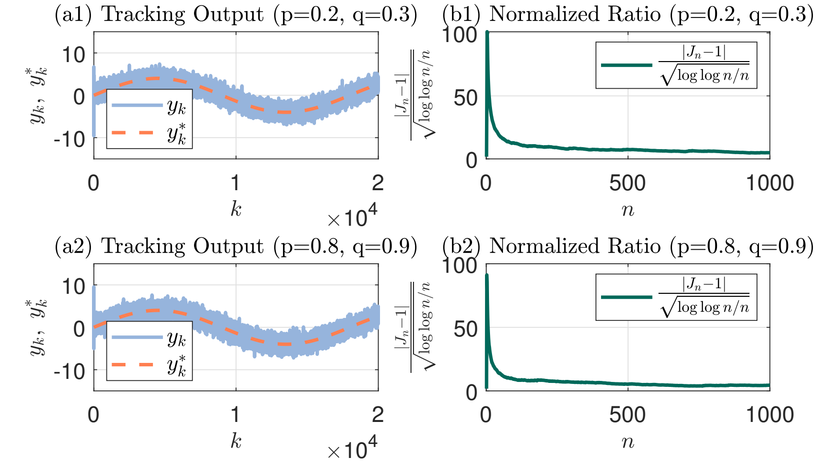


Fig. 6. Adaptive tracking with binary feedback under two attack levels. (a1,a2) Output $\boldsymbol{y}_{\boldsymbol{k}}$ tracks reference $\boldsymbol{y}_{\boldsymbol{k}}^*$ under $(\boldsymbol{p}, \boldsymbol{q}) = (\mathbf{0.2}, \mathbf{0.3})$ and $(\mathbf{0.8}, \mathbf{0.9})$. (b1,b2) Normalized error $\frac{|\boldsymbol{J_n} - \boldsymbol{\sigma}^2|}{\sqrt{\log\log \boldsymbol{n}/\boldsymbol{n}}}$.

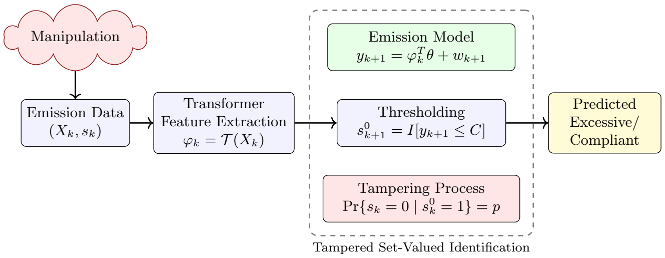

Manipulation
Emission Data (X_k, s_k)
Transformer Feature Extraction $\varphi_k = \mathcal{T}(X_k)$
Emission Model $y_{k+1} = \varphi_k^T\theta + w_{k+1}$
Thresholding $s_{k+1}^0 = I[y_{k+1} \le C]$
Tampering Process $\Pr\{s_k = 0 \mid s_k^0 = 1\} = p$
Tampered Set-Valued Identification
Predicted Excessive/ Compliant

Fig. 7. Modeling pipeline for tampered emission data and binary classification.

B. Practical Case: Vehicle Emissions Inspection Fraud

Vehicle emission control, especially for heavy-duty diesel vehicles, plays a critical role in reducing air pollution. However, fraudulent practices in emissions testing—such as using cheat devices to manipulate onboard diagnostic (OBD) systems and falsify excessive emission data—pose serious challenges to environmental monitoring and regulatory enforcement [45], [46]. To illustrate the practical relevance of our proposed tampering detection algorithm, we examine a real-world case involving in-use vehicle emissions data. This problem can be naturally formulated within our tampering framework, as falsified emission reports are essentially misclassified compliance labels. We use OBD monitoring data collected in *Hefei, China*, from *August 8 to November 24, 2020*. The dataset includes more than *140,000 records* from over *100 heavy-duty diesel vehicles*, containing both emission measurements and GPS information. Each vehicle was observed for approximately *3 hours*, with data recorded every *5 seconds*, covering a geographic range between *latitudes 31.754772–31.785486* and *longitudes 117.192519–117.245141*.

Given the dynamic driving conditions and the complex temporal structure of emission signals, traditional static models often fail to detect anomalies effectively. To address this, we first process the raw OBD signals using a Transformer-based neural network [47], which captures long-range dependencies and nonlinear patterns. This yields a sequence of feature vectors φ_k, each representing the system state at time k.

These features are then used in the model introduced in Section 2.1. We apply the regression model (1) to estimate the predicted emission y_{k+1}, followed by a thresholding operation (3) to assign binary compliance labels s_k^0. A label of $s_k^0 = 1$ indicates an excessive emission event, i.e., predicted emissions exceeding the regulatory threshold C.

However, in real-world inspection scenarios, such binary labels may be deliberately falsified. To model this, we adopt the tampering mechanism (4), where p denotes the probability of misreporting excessive emissions as compliant, and q the reverse. This completes our pipeline—from raw OBD signals to tampering-aware binary classification—as shown in Figure 7.

To assess the effectiveness of the proposed tampering-aware algorithms, we present two sets of simulation results. Figure 8 shows confusion matrices for five classification algorithms: support vector machine (SVM), logistic regression (Logistic), decision tree (Tree), NRP-TB-KP, and GRP-TB-KP at tampering probabilities $p = 0.3$. The top row results are based on original raw data, while the bottom row represents results after extracting features with a Transformer. Clearly, classification using raw data alone yields relatively low accuracy, emphasizing the need for effective feature extraction methods. After Transformer-based feature extraction, classification accuracy improves across all algorithms. Specifically, the proposed methods based on Transformer-extracted features, NRP-TB-KP and GRP-TB-KP, outperform the baseline classifiers. Among these, the NRP-TB-KP algorithm achieves the highest accuracy overall, highlighting its effectiveness in addressing label tampering and enhancing classification robustness.

Figure 9 presents scatter plots corresponding to the confusion matrices in Figure 8. Clearly, the proposed NRP-TB-KP algorithm achieves more accurate classification than the compared methods. The points labeled “Tam” represent cases with falsified excessive emissions, which are effectively identified and corrected by our proposed approach.

Figure 10 shows the average accuracy and computation time of different methods under varying tampering probabilities ($p = 0$ to 0.9). Observations include: (i) Newton-based methods generally achieve higher accuracy, whereas gradient-based methods have shorter computation times; (ii) ignoring tampering consistently leads to decreased performance; and (iii) accuracy typically declines as the tampering probability increases, except the case $p = 0.1$. Nonetheless, methods explicitly incorporating tampering probabilities consistently outperform those that do not. These results highlight the importance of considering tampering in emission detection tasks and illustrate the trade-off between accuracy and computational efficiency in practical scenarios.

VI. Concluding remarks

This paper investigated the problem of parameter estimation and adaptive control for systems with binary observations under data tampering attacks. We developed both gradient-based and second-order Quasi-Newton identification algorithms that are applicable when the attack strategy is either known or unknown. The proposed methods ensure asymptotic convergence of parameter estimates and do not rely on PE condition. In addition, the second-order algorithm was integrated into an adaptive control framework, allowing for explicit tracking error bounds in binary FIR systems even under unknown attacks. Simulation results show the robustness and efficiency of the algorithms, and a vehicle emission control task is used to test their ability.

Future work may extend the framework to multi-agent systems, consider time-varying attack models, or apply it to networked control systems with quantized or event-triggered communication.

[1] Z. Kazemi, A. A. Safavi, M. M. Arefi, and F. Naseri, “Finite-time secure dynamic state estimation for cyber–physical systems under unknown inputs and sensor attacks,” *IEEE transactions on systems, man, and cybernetics: systems*, vol. 52, no. 8, pp. 4950–4959, 2021.

[2] C. Ma, N. Xi, D. Lu, Y. Feng, and J. Ma, “Ctomp: a cycle-task-oriented memory protection scheme for unmanned systems,” *Science China Information Sciences*, vol. 67, no. 6, p. 162305, 2024.

[3] J. Lee, B. Bagheri, and H.-A. Kao, “A cyber-physical systems architecture for industry 4.0-based manufacturing systems,” *Manufacturing letters*, vol. 3, pp. 18–23, 2015.

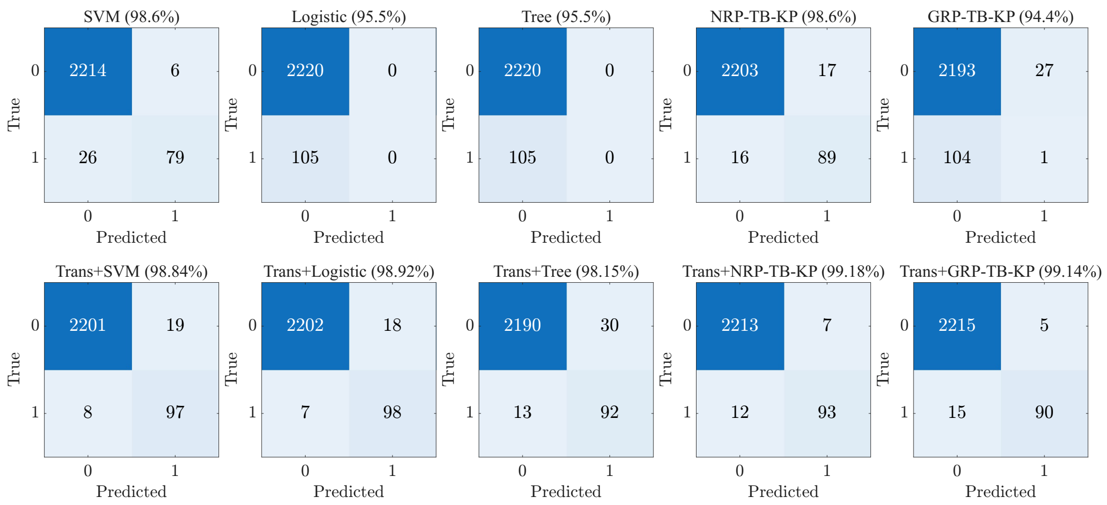
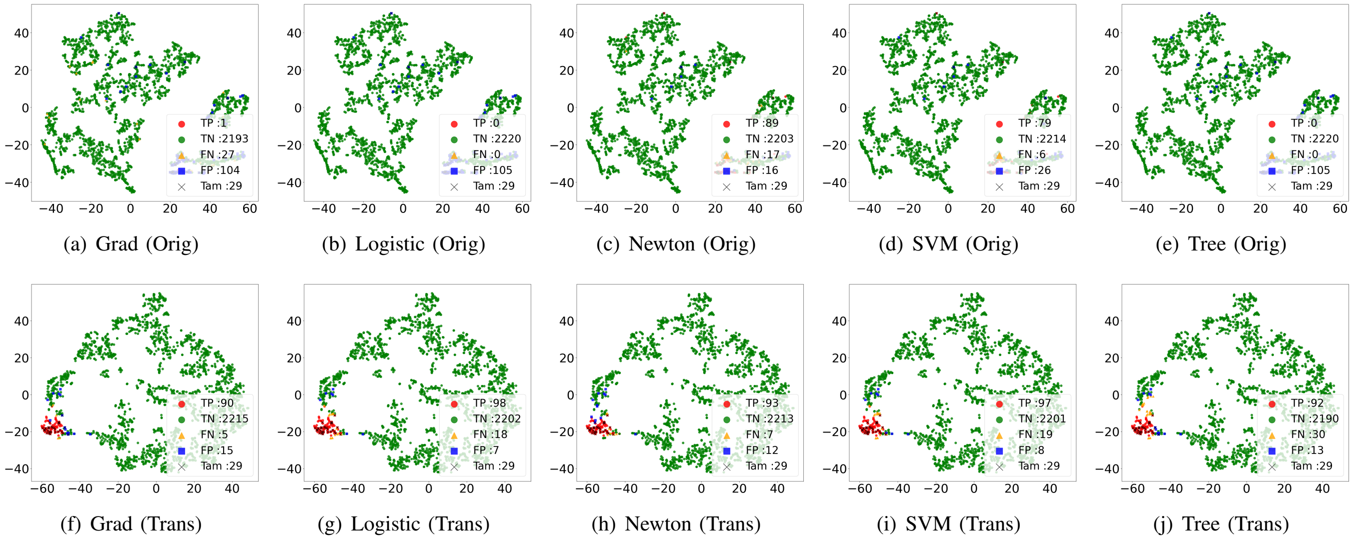


Fig. 8. Confusion matrices of five classifiers using Original (top) and Transformer-based features (bottom).



(a) Grad (Orig) (b) Logistic (Orig) (c) Newton (Orig) (d) SVM (Orig) (e) Tree (Orig)

(f) Grad (Trans) (g) Logistic (Trans) (h) Newton (Trans) (i) SVM (Trans) (j) Tree (Trans)

Fig. 9. Detection results under different methods. Top: original data; bottom: Transformer-enhanced data. Labels: TP (True Positive), FP (False Positive), TN (True Negative), FN (False Negative), Tam (Tampered).

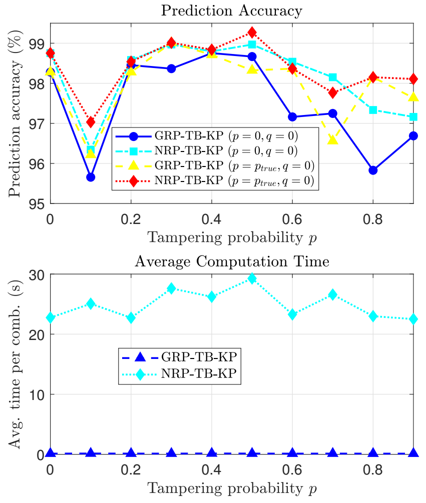


Fig. 10. Prediction accuracy and computation time under varying tampering probabilities.

[4] A. Hahn, A. Ashok, S. Sridhar, and M. Govindarasu, "Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 847–855, 2013.

[5] L. Peng, L. Shi, X. Cao, and C. Sun, "Optimal attack energy allocation against remote state estimation," *IEEE Transactions on Automatic Control*, vol. 63, no. 7, pp. 2199–2205, 2017.

[6] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, Nov. 2013.

[7] C. Kwon and I. Hwang, "Reachability analysis for safety assurance of cyber-physical systems against cyber attacks," *IEEE Transactions on Automatic Control*, vol. 63, no. 7, pp. 2272–2279, 2018.

[8] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Transactions on Automatic control*, vol. 59, no. 6, pp. 1454–1467, 2014.

[9] K. Pelechrinis, M. Iliofotou, and S. V. Krishnamurthy, "Denial of service attacks in wireless networks: The case of jammers," *IEEE Communications surveys & tutorials*, vol. 13, no. 2, pp. 245–257, 2010.

[10] I. Jovanov and M. Pajic, "Relaxing integrity requirements for attack-resilient cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 12, pp. 4843–4858, 2019.

[11] G. Oliva, S. Cioabă, and C. N. Hadjicostis, "Distributed calculation of edge-disjoint spanning trees for robustifying distributed algorithms

against man-in-the-middle attacks," *IEEE Transactions on Control of Networked Systems*, vol. 5, no. 4, pp. 1646–1656, Dec. 2018.

[12] K. Yamada, J. Hoshino, and R. Kubo, "Detection of data tampering attacks using redundant network paths with different delays for networked control systems," *Nonlinear Theory and Its Applications*, vol. 10, no. 2, pp. 140–156, 2019.

[13] H. Sasahara and H. Sandberg, "Asymptotic security using bayesian defense mechanism with application to cyber deception," *IEEE Transactions on Automatic Control*, vol. 69, no. 8, pp. 5004–5019, 2023.

[14] D. Umsonst, S. Sarıtaş, G. Dán, and H. Sandberg, "A bayesian nash equilibrium-based moving target defense against stealthy sensor attacks," *IEEE Transactions on Automatic Control*, vol. 69, no. 3, pp. 1659–1674, 2023.

[15] S. Amini, F. Pasqualetti, and H. Mohsenian-Rad, "Dynamic load altering attacks against power system stability: Attack models and protection schemes," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2862–2872, 2016.

[16] C.-Z. Bai, F. Pasqualetti, and V. Gupta, "Data-injection attacks in stochastic control systems: Detectability and performance tradeoffs," *Automatica*, vol. 82, pp. 251–260, 2017.

[17] V. Katewa, R. Anguluri, and F. Pasqualetti, "On a security vs privacy trade-off in interconnected dynamical systems," *Automatica*, vol. 125, p. 109426, 2021.

[18] X. Jin, W. M. Haddad, and T. Yucelen, "An adaptive control architecture for mitigating sensor and actuator attacks in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 6058–6064, 2017.

[19] T. Yucelen, W. M. Haddad, and E. M. Feron, "Adaptive control architectures for mitigating sensor attacks in cyber-physical systems," *Cyber-Physical Systems*, vol. 2, no. 1-4, pp. 24–52, 2016.

[20] W. M. Haddad, D. Venkat, T. Yucelen, and M. S. Whorton, "Adaptive control for nonlinear cyber–physical systems in the presence of actuator attacks," *Nonlinear Analysis: Real World Applications*, vol. 84, p. 104302, 2025.

[21] S. Liu, A. Trivedi, X. Yin, and M. Zamani, "Secure-by-construction synthesis of cyber-physical systems," *Annual Reviews in Control*, vol. 53, pp. 30–50, 2022.

[22] B. Zhong, S. Liu, M. Caccamo, and M. Zamani, "Secure-by-construction synthesis for control systems," *IEEE Transactions on Automatic Control*, 2025.

[23] L. Y. Wang, G. G. Yin, J. F. Zhang, and Y. L. Zhao, *System Identification with Quantized Observations*. Boston, MA, USA: Birkhäuser, 2010.

[24] J. Wu, Q. Jia, K. Johansson, and L. Shi, "Event-based sensor data scheduling: Trade-off between sensor communication rate and estimation quality," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 1041–1046, Apr. 2013.

[25] L. Y. Wang, J. F. Zhang, and G. Yin, "System identification using binary sensors," *IEEE Transactions on Automatic Control*, vol. 48, no. 11, pp. 1892–1907, Nov. 2003.

[26] J. Guo, L. Y. Wang, G. Yin, Y. L. Zhao, and J. F. Zhang, "Asymptotically efficient identification of fir systems with quantized observations and general quantized inputs," *Automatica*, vol. 57, pp. 113–122, 2015.

[27] G. Bottegal, H. Hjalmarsson, and G. Pillonetto, "A new kernel-based approach to system identification with quantized output data," *Automatica*, vol. 85, pp. 145–152, 2017.

[28] M. Pouliquen, E. Pigeon, O. Gehan, and A. Goudjil, "Identification using binary measurements for iir systems," *IEEE Transactions on Automatic Control*, vol. 65, no. 2, pp. 786–793, Feb. 2020.

[29] M. Casini, A. Garulli, and A. Vicino, "Input design in worst-case system identification using binary sensors," *IEEE Transactions on Automatic Control*, vol. 56, no. 5, pp. 1186–1191, May. 2011.

[30] Y. Wang, Y. Zhao, and J.-F. Zhang, "Asymptotically efficient quasi-newton type identification with quantized observations under bounded persistent excitations," *Automatica*, vol. 166, p. 111722, 2024.

[31] J. Guo, X. Wang, W. Xue, and Y. Zhao, "System identification with binary-valued observations under data tampering attacks," *IEEE Transactions on Automatic Control*, vol. 66, no. 8, pp. 3825–3832, 2020.

[32] J. Guo, R. Jia, R. Su, and Y. Zhao, "Identification of fir systems with binary-valued observations against data tampering attacks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 9, pp. 5861–5873, 2023.

[33] J. Guo, J. Cheng, and J.-D. Diao, "System identification with binary-valued output observations under either-or communication and data packet dropout," *Systems & Control Letters*, vol. 156, p. 105010, 2021.

[34] J. Guo, W. Xue, T. Wang, J.-F. Zhang, and Y. Zhang, "On iterative parameter identification of fir systems with batched possibly incorrect binary-valued observations," in *2023 62nd IEEE Conference on Decision and Control (CDC)*. IEEE, 2023, pp. 4936–4941.

[35] J. Guo, Q. Zhang, and Y. Zhao, "Identification of fir systems with binary-valued observations under replay attacks," *Automatica*, vol. 172, p. 112001, 2025.

[36] J. Guo, W. Xue, and J.-F. Zhang, "Recursive identification of fir systems under data tampering attacks with binary-valued outputs," in *Proceedings of the 13th IFAC Symposium on Nonlinear Control Systems (NOLCOS 2025)*, Reykjavík, Iceland, Jul. 2025.

[37] T. L. Lai and C. Z. Wei, "Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems," *The Annals of Statistics*, pp. 154–166, 1982.

[38] B. Zheng, C. You, W. Mei, and R. Zhang, "A survey on channel estimation and practical passive beamforming design for intelligent reflecting surface aided wireless communications," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 1035–1071, 2022.

[39] H. Lee, H. Choi, H. Kim, S. Kim, C. Jang, Y. Choi, and J. Choi, "Downlink channel reconstruction for spatial multiplexing in massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 9, pp. 6154–6166, 2021.

[40] L. Ljung, *System identification toolbox: User's guide*. Citeseer, 1995.

[41] H.-F. Chen, "Stochastic approximation with applications," in *Encyclopedia of Systems and Control*. Springer, 2021, pp. 2154–2160.

[42] J. Guo and Y. Zhao, "Recursive projection algorithm on fir system identification with binary-valued observations," *Automatica*, vol. 49, pp. 3396–3401, 2013.

[43] L. Ljung and T. Söderström, *Theory and practice of recursive identification*. MIT press, 1983.

[44] L. Guo, "Time-varying stochastic systems, stability and adaptive theory," in *2nd ed.* Science Press, 2020.

[45] W. Zhang, "Various forms of environmental monitoring data falsification severely punished and exposed by ecological environment departments," *Legal Daily*, no. 006, 2023, (In Chinese).

[46] C. Wang and J. Wu, "Data tampering, substitute testing, and other frequent fraudulent practices in vehicle emissions testing," *Xinhua Daily Telegraph*, no. 002, 2025, (In Chinese).

[47] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 30, 2017, pp. 5998–6008.

[48] P. Pedregal, *Introduction to optimization*. Springer, 2004, vol. 46.

[49] J. Wang, J. Ke, and J.-F. Zhang, "Differentially private bipartite consensus over signed networks with time-varying noises," *IEEE Transactions on Automatic Control*, 2024.

[50] H. Zhang, T. Wang, and Y. Zhao, "Asymptotically efficient recursive identification of fir systems with binary-valued observations," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 5, pp. 2687–2700, 2019.

[51] L. Zhang, Y. Zhao, and L. Guo, "Identification and adaptation with binary-valued observations under non-persistent excitation condition," *Automatica*, vol. 138, p. 110158, 2022.

Appendix I
Proofs of Theorem 3.1 and 3.2

Recall the estimate error $\tilde{\theta}_k = \hat{\theta}_k - \theta$. First, we give the following useful lemmas.

Proposition I.1: The projection operator given by Definition 3.1 follows

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

Lemma I.1: [48] 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 I.2: [49] For $0 < b \leq 1, a > 0, k_0 \geq 0$ and sufficiently large l, we have

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

Lemma I.3: [50] For any given positive integer l and $a,b\in\mathbb{R}$, the following results hold

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

Lemma I.4: If Assumptions A1–A3 hold, then

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

Proof: First, we have

$$\begin{aligned}\|\tilde{\theta}_{k+l}-\tilde{\theta}_k\| &= \|\hat{\theta}_{k+l}-\hat{\theta}_k\| = \left\|\sum_{j=1}^{l}(\hat{\theta}_{k+j}-\hat{\theta}_{k+j-1})\right\| \\ &\le \sum_{j=1}^{l}\left\|(\hat{\theta}_{k+j}-\hat{\theta}_{k+j-1})\right\|. \end{aligned}\tag{47}$$

Since $0\le(1-(p+q))F_k(C-\theta^T\varphi_k)+q\le 1$, it follows that $|\tilde{s}_k|\le\beta$. Combining this with Proposition I.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 (47) and Assumption 4, implies the lemma. ■

Proof of Theorem 3.1. By $\tilde{s}_k^2\le\beta^2$, Proposition I.1 and (10), we have

$$\begin{aligned}\|\tilde{\theta}_{k+1}\|^2 &\le \|\tilde{\theta}_k\|^2+2b_k\tilde{s}_{k+1}\phi_k^T\tilde{\theta}_k+b_k^2\|\phi_k\|^2\beta^2 \\ &= \|\tilde{\theta}_k\|^2+2b_k\tilde{s}_{k+1}\phi_k^T\tilde{\theta}_k+O(b_k^2). \end{aligned}\tag{48}$$

From (8) and (11), it follows that

$$\begin{aligned}&\mathbb{E}[\tilde{s}_{k+1}|\mathcal{F}_k] \\ &= \beta(1-p-q)^2\left(F(C-\phi_k^T\hat{\theta}_k)-F(C-\phi_k^T\theta)\right). \end{aligned}\tag{49}$$

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

$$\begin{aligned}&\mathbb{E}\left[2b_k\tilde{s}_{k+1}\phi_k^T\tilde{\theta}_k|\mathcal{F}_k\right]=2b_k\phi_k^T\tilde{\theta}_k\mathbb{E}[\tilde{s}_{k+1}|\mathcal{F}_k] \\ &= 2b_k\phi_k^T\tilde{\theta}_k\beta(1-p-q)^2\left(F_k(C-\phi_k^T\hat{\theta}_k)-F_k(C-\phi_k^T\theta)\right) \\ &= -2b_k\beta(1-p-q)^2f_k(\xi_k)\tilde{\theta}_k^T\phi_k\phi_k^T\tilde{\theta}_k \\ &\le -2b_k\beta(1-p-q)^2\underline{f}\tilde{\theta}_k^T\phi_k\phi_k^T\tilde{\theta}_k, \end{aligned}\tag{50}$$

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 (48) and substituting (50) into it, we can obtain

$$\begin{aligned}&\mathbb{E}\|\tilde{\theta}_{k+1}\|^2\le\mathbb{E}\|\tilde{\theta}_k\|^2+2b_k\mathbb{E}\tilde{s}_{k+1}\phi_k^T\tilde{\theta}_k+O(b_k^2) \\ &= \mathbb{E}\|\tilde{\theta}_k\|^2+\mathbb{E}\left[\mathbb{E}\left[2b_k\tilde{s}_{k+1}\phi_k^T\tilde{\theta}_k|\mathcal{F}_k\right]\right]+O(b_k^2). \\ &\le \mathbb{E}\|\tilde{\theta}_k\|^2-2b_k\beta(1-p-q)^2\underline{f}\mathbb{E}[\tilde{\theta}_k^T\phi_k\phi_k^T\tilde{\theta}_k] \\ &\quad +O(b_k^2). \end{aligned}\tag{51}$$

By iterating (51) h times and noting $b_k=O(b_{k+1})$, we obtain

$$\begin{aligned}&\mathbb{E}\|\tilde{\theta}_{k+h}\|^2 \\ &\le \mathbb{E}\|\tilde{\theta}_k\|^2-2\beta(1-p-q)^2\underline{f}\mathbb{E}\left[\sum_{l=k}^{k+h-1}[b_l\tilde{\theta}_l^T\phi_l\phi_l^T\tilde{\theta}_l]\right] \\ &\quad +O(b_{k+h}^2) \\ &\le \mathbb{E}\|\tilde{\theta}_k\|^2-2\beta(1-p-q)^2\underline{f}\mathbb{E}\left[\sum_{l=k}^{k+h-1}[b_l\tilde{\theta}_k^T\phi_l\phi_l^T\tilde{\theta}_k]\right] \\ &\quad -2\beta(1-p-q)^2\underline{f}\mathbb{E}\left[\sum_{l=k}^{k+h-1}[b_l(\tilde{\theta}_l-\tilde{\theta}_k)^T\phi_l\phi_l^T(\tilde{\theta}_l-\tilde{\theta}_k)]\right] \\ &\quad +O(b_{k+h}^2). \end{aligned}\tag{52}$$

By Lemma I.4 and (12), the last two terms of (51) are of order $O(b_{k+h}^2)$. In addition,

$$\begin{aligned}&\mathbb{E}\left[\sum_{l=k}^{k+h-1}[b_l\tilde{\theta}_k^T\phi_l\phi_l^T\tilde{\theta}_k]\right] \\ &= \mathbb{E}\left[\tilde{\theta}_k^T\mathbb{E}\left[\sum_{l=k}^{k+h-1}b_l\phi_l\phi_l^T\middle|\mathcal{F}_k\right]\tilde{\theta}_k\right]. \end{aligned}\tag{53}$$

By Assumption A3, we have

$$\begin{aligned}&\mathbb{E}\left[\sum_{l=k}^{k+h-1}b_l\phi_l\phi_l^T\middle|\mathcal{F}_k\right] \\ &= \sum_{l=k}^{k+h-1}b_l\frac{1}{h}\mathbb{E}\left[\sum_{l=k}^{k+h-1}\phi_l\phi_l^T\middle|\mathcal{F}_k\right]+O(b_{k+h}^2) \\ &\ge \delta\sum_{l=k}^{k+h-1}b_lI+O(b_{k+h}^2). \end{aligned}\tag{54}$$

Substituting (53) and (54) into (52) yields

$$\begin{aligned}&\mathbb{E}\|\tilde{\theta}_{k+h}\|^2 \\ &\le \mathbb{E}\|\tilde{\theta}_k\|^2-2\beta(1-p-q)^2\underline{f}\delta\sum_{l=k}^{k+h-1}b_l\mathbb{E}\|\tilde{\theta}_k\|^2+O(b_{k+h}^2) \\ &= \left(1-2\beta(1-p-q)^2\underline{f}\delta\sum_{l=k}^{k+h-1}b_l\right)\mathbb{E}\|\tilde{\theta}_k\|^2+O(b_{k+h}^2). \end{aligned}\tag{55}$$

Then, based on Lemma I.1 and Assumption A4, 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 $\lim_{k\to\infty}\mathbb{E}[\|\tilde{\theta}_k\|^2]=0$.

On the other hand, by (51) we have $\mathbb{E}[\|\tilde{\theta}_{k+1}\|^2|\mathcal{F}_k]\le\|\tilde{\theta}_k\|^2+O(b_k^2)$, which together with [41, Lemma 1.2.2] and $\sum\limits_{k=1}^{\infty}b_k^2<\infty$ implies that $\|\tilde{\theta}_k\|$ converges to a bounded limit a.s. Notice that $\lim_{k\to\infty}\mathbb{E}[\tilde{\theta}_k^T\tilde{\theta}_k]=0$. Then, there is a subsequence of $\tilde{\theta}_k$ that converges almost surely to 0. Consequently, $\tilde{\theta}_k$ almost surely converges to 0. □

Proof of Theorem 3.2. 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 (55), we have

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

where $\kappa = \lceil\frac{k-K}{h}\rceil - \lfloor\frac{k-K}{h}\rfloor$. This together with Lemma I.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 (55),

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

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 I.3, we have $\mathbb{E}\left\|\tilde{\theta}_k\right\|^2 = O\left(\frac{1}{k}\right)$. This completes this part's proof. □

Appendix II
Proofs of Theorem 4.1-4.3

For convenience, we introduce the notation

$$
\begin{aligned}
\varepsilon_{k+1} &= (1-(p+q))F_k(C-\theta^T\varphi_k) + q - s_{k+1}, &(56)\\
\psi_k &= (1-(p+q))\\
&\quad\left(F_{k+1}(C-\hat{\theta}_k^T\varphi_k) - F_{k+1}(C-\theta^T\varphi_k)\right). &(57)
\end{aligned}
$$

We then give the following lemma.

Lemma II.1: Let Assumptions A1, A2, A3(a), and A5 be satisfied. Then the parameter estimate given by Algorithm 3 has the following property as $n \to \infty$:

$$
\tilde{\theta}_{n+1}^T P_{n+1}^{-1}\tilde{\theta}_{n+1} + \beta_n^2\sum_{k=0}^{n}\left(\tilde{\theta}_k^T\varphi_k\right)^2 = O\left(\log|P_{n+1}^{-1}|\right), \quad \text{a.s.} \tag{58}
$$

Proof: Recall $\tilde{\theta}_k = \theta - \hat{\theta}_k$ and define the stochastic Lyapunov function:

$$
V_k = \tilde{\theta}_k^T P_k^{-1}\tilde{\theta}_k.
$$

By (29) and noting $a_k P_{k+1}^{-1}P_k\varphi_k = \varphi_k$, $P_{k+1}^{-1} = P_k^{-1} + \beta_k^2\varphi_k\varphi_k^T$ and $|\psi_k| \leq 1$, we get

$$
\begin{aligned}
V_{k+1} \leq\ &\tilde{\theta}_k^T P_k^{-1}\tilde{\theta}_k - 2\beta_k\tilde{\theta}_k^T\varphi_k\psi_k + \beta_k^2(\tilde{\theta}_k^T\varphi_k)^2\\
&+2a_k\beta_k^2\psi_k\varphi_k^T P_k\varphi_k\varepsilon_{k+1} - 2\beta_k\tilde{\theta}_k^{uT}\varphi_k\varepsilon_{k+1}\\
&+a_k\beta_k^2\varphi_k^T P_k\varphi_k + a_k\beta_k^2\varphi_k^T P_k\varphi_k\varepsilon_{k+1}^2.
\end{aligned} \tag{59}
$$

By the definition of β_k in Algorithm 3, (57), and the Mean Value Theorem, it follows

$$
2\beta_k\tilde{\theta}_k^T\varphi_k\psi_k = 2\beta_k(\tilde{\theta}_k^T\varphi_k)^2 f_k(\xi_k)(1-(p+q)) \geq 2\beta_k^2(\tilde{\theta}_k^T\varphi_k)^2
$$

where ξ_k lies between $C - \theta^T\varphi_k$ and $C - \hat{\theta}_k^T\varphi_k$. Then by summing up both sides of (59), it becomes

$$
\begin{aligned}
V_{n+1} \leq\ &V_0 - \sum_{k=0}^{n}\beta_k^2(\tilde{\theta}_k^T\varphi_k)^2\\
&\underbrace{+2\sum_{k=0}^{n}a_k\beta_k^2\psi_k\varphi_k^T P_k\varphi_k\varepsilon_{k+1} - 2\sum_{k=0}^{n}\beta_k\tilde{\theta}_k^{uT}\varphi_k\varepsilon_{k+1}}_{\text{I}}\\
&\underbrace{+\sum_{k=0}^{n}a_k\beta_k^2\varphi_k^T P_k\varphi_k + \sum_{k=0}^{n}a_k\beta_k^2\varphi_k^T P_k\varphi_k\varepsilon_{k+1}^2}_{\text{II}}.
\end{aligned} \tag{60}
$$

From (56) and (8), it follows $\mathbb{E}(\varepsilon_{k+1} \mid \mathcal{F}_k) = 0$, which means $\{\varepsilon_k, \mathcal{F}_k\}$ is a martingale difference sequence. Similar to the proof in [44] and noting Assumption A3(a) $\sup_{k\geq 1}\|\varphi_k\| \leq M < \infty$, we have

$$\begin{aligned}\mathrm{I} &= o\left(\sum_{k=0}^{n}\beta_k^2(\tilde{\theta}_k^T\varphi_k)^2\right)+O(1)\\ \mathrm{II} &= O\left(\log|P_{n+1}^{-1}|\right).\end{aligned} \tag{61}$$

Combining all terms, we obtain

$$\tilde{\theta}_{n+1}^T P_{n+1}^{-1}\tilde{\theta}_{n+1}+\sum_{k=0}^{n}\beta_k^2(\tilde{\theta}_k^T\varphi_k)^2 = O(\log|P_{n+1}^{-1}|), \quad \text{a.s.}$$

Finally, since $\{\beta_k\}$ is a non-increasing sequence, we conclude (58). ■

Proof of Theorem 4.1 and 4.2. Directly obtained from Lemma II.1.

Proof of Theorem 4.3. By Lemma II.1, the proof follows similarly to [51, Theorem 3].

Jian Guo received the B.S. degree in Mathematics from Xi'an Jiaotong University, Xi'an, China, in 2019, and the Ph.D. degree in system analysis and integration from the Academy of Mathematics and Systems Science (AMSS), Chinese Academy of Sciences (CAS), Beijing, China, in 2024. He is currently a postdoctoral fellow at the CAS AMSS–PolyU Joint Laboratory of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China. His research interests include sparse identification, control of multi-agent systems, and dynamic stochastic variational inequalities.

Lihong Pei received the Ph.D. degree in control science and engineering from the Department of Automation, University of Science and Technology of China, Hefei, China, in 2024. She is currently a Post-Doctoral Fellow at the Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences. Her research focuses on the intelligent monitoring and control of carbon emissions in transportation.

Wenchao XUe (Member,IEEE) received the B.S. degree in applied mathematics from Nankai University, Tianjin, China, in 2007, and the Ph.D. degree in control theory from the Academy of Mathematics and Systems Science (AMSS), Chinese Academy of Sciences (CAS) Beijing, China, in 2012. He is currently a Professor with the key Lab of System and Control, AMSS, CAS. His research interests include nonlinear uncertain systems control, nonlinear uncertain systems filter and distributed filter.

Dr. Xue is an Associate Editor for IFAC Journal of Control Engineering practice.

Yanlong Zhao (Senior Member, IEEE) 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 full Professor and a Vice Director of State Key Laboratory of Mathematical Sciences. His research interests include identification and control of quantized systems, networked systems, information theory and modeling of financial systems.

He received the second prize of the State Natural Science Award of China in 2015. He has been a Deputy Editor-in-Chief Journal of Systems and Science and Complexity, an Associate Editor of Automatica, SIAM Journal on Control and Optimization, and IEEE Transactions on Systems, Man and Cybernetics: Systems. He served as a Vice President of Asian Control Association and IEEE CSS Beijing Chapter, and is now a Vice President of Chinese Association of Automation (CAA), Chair of Technical Committee on Control Theory (TCCT), CAA and member of IFAC TC 1.1 Modeling, Identification & Signal Processing.

Ji-Feng Zhang (IEEE Fellow) received the B.S. degree in mathematics from Shandong University, China, in 1985, and the Ph.D. degree from the Institute of Systems Science, Chinese Academy of Sciences (CAS), China, in 1991. Now he is with the School of Automation and Electrical Engineering, Zhongyuan University of Technology; and the State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, CAS. His current re search interests include system modeling, adaptive control, stochastic systems, and multi-agent systems.

He is an IEEE Fellow, IFAC Fellow, CAA Fellow, SIAM 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 was a Vice-President of the Chinese Association of Automation, the Chinese Mathematical Society and the Systems Engineering Society of China. 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. 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