Electronic Research Archive

https://www.aimspress.com/journal/era

ERA, 33(6): 3989–4013.
DOI: 10.3934/era.2025177
Received: 08 March 2025
Revised: 31 May 2025
Accepted: 11 June 2025
Published: 24 June 2025

Research article

Identification of a FIR system with binary-valued observation under data tampering attack and differential privacy preservation

Bochen Li[1] **and Ting Wang**[2,*]

[1] School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

[2] School of Intelligence Science and Technology, University of Science and Technology Beijing, Beijing 100083, China

* **Correspondence:** Email: wangting@ustb.edu.cn.

Abstract: This paper explores the use of differential privacy encryption to protect against data tampering attacks in the context of finite impulse response (FIR) system identification under binary observation conditions. The study begins by introducing the core principles of differential privacy and discussing the current security challenges faced by FIR systems. It highlights the risks of data tampering and privacy leakage during the system identification process. To address these challenges, two distinct differential privacy algorithms are proposed, providing dual encryption protection for the system parameters. By integrating differential privacy mechanisms into the FIR system, the proposed approach ensures the security and privacy of both data and parameters during transmission and processing. Experimental results demonstrate that the dual differential privacy protection effectively safeguards data while providing accurate parameter estimation, validating the effectiveness of the proposed scheme.

Keywords: differential privacy encryption scheme; finite impulse response (FIR) system; multiparameter system protection

1. Introduction

In today's highly interconnected world, data privacy protection has become a critical concern across various application domains. Personal medical records, industrial control parameters, and even trade secrets are increasingly vulnerable to attacks or misuse [1–4]. When sensitive information is maliciously accessed or tampered with, it can lead not only to significant economic and societal losses but also to a destabilization of system security and stability. Therefore, the effective protection of data privacy and the prevention of unauthorized inferences and misuse of both individual and core

information have become central issues of widespread interest in both academic and industrial circles [5–7].

The primary benefit of data privacy protection lies in its ability to intelligently mask, encrypt, or perturb personal and system-critical data, thereby preventing external attackers from reverse-engineering sensitive information. By implementing robust privacy-preserving techniques such as differential privacy, homomorphic encryption, and secure multi-party computation, organizations can effectively minimize security risks while maintaining the integrity and utility of the data. This ensures that individuals' personal privacy, corporate trade secrets, and other confidential information remain safeguarded during data collection, processing, analysis, and sharing. Moreover, a well-designed privacy protection framework fosters user trust, strengthens regulatory compliance, and enhances the overall resilience of digital systems against cyber threats, contributing to a more secure and sustainable data-driven ecosystem.

Over the years, researchers have developed various strategies for data security and privacy protection from both theoretical and practical perspectives. For instance, Rawat et al., in their exploration of cyber-physical systems (CPS) in smart grids and healthcare systems, emphasized the importance of security measures at both the network and physical layers for ensuring data confidentiality [8]. Ljung's systematic research on system identification laid the foundation for subsequent work on modeling and algorithmic optimization related to privacy protection and assessment [9]. Pouliquen et al. proposed a robust parameter estimation framework for scenarios with limited information based on binary measurements [10]. In a related study, Guo et al. developed an adaptive tracking approach for first-order systems using binary-valued observations with fixed thresholds, enhancing the accuracy of parameter estimation in constrained data environments [11]. The framework of local differential privacy had widely adopted in data statistics and analysis tasks, with comprehensive surveys that of Wang et al. demonstrating its crucial role in frequency estimation and machine learning applications under distributed settings [12]. Moreover, Ding et al. and Liu have each proposed methods for distributed control and networked predictive control in industrial CPS, embedding security and privacy requirements into real-time collaborative control processes [13, 14]. In addition, Mahmoud et al. introduced a multi-layer defense approach for network attack modeling and vulnerability assessment, offering innovative solutions for maintaining data confidentiality against diverse threats [15]. Recent work by Taheri et al. explored differential privacy-driven adversarial attacks, like noise injection, to degrade deep learning-based malware detection, proposing defenses to enhance robustness [16]. In contrast, our differential privacy-based FIR system identification algorithm uses Laplace and randomized response mechanisms to encrypt data, ensuring parameter estimation accuracy in cyber-physical systems under privacy and security constraints.

In the realm of federated learning, data and model poisoning attacks pose significant challenges to the integrity and reliability of distributed systems, particularly impacting the accuracy of system identification processes. Nabavirazavi et al. proposed a randomization and mixture approach to enhance the robustness of federated learning against such attacks, demonstrating improved resilience in parameter estimation under adversarial conditions [17]. Similarly, Taheri et al. introduced robust aggregation functions to mitigate the effects of poisoned data, ensuring more reliable model updates in federated environments [18]. Nowroozi et al. exposed vulnerabilities in federated learning through data poisoning attacks, highlighting their detrimental impact on network security and the need for robust defenses to protect system dynamics [19]. Furthermore, Nabavirazavi et al. explored the

impact of aggregation function randomization to counter model poisoning, emphasizing the importance of accurate parameter estimation despite malicious interventions [20]. These studies underscore the importance of accurate parameter estimation in adversarial environments. However, despite these advances in federated learning and privacy-preserving training, less attention has been paid to the fundamental task of system identification—an essential step for capturing system dynamics and ensuring the effectiveness of both attack detection and privacy protection strategies.

System identification is a key process in ensuring that attack detection and privacy protection measures are fully effective. Given the challenges of data tampering and privacy preservation, this paper focuses on the problem of data privacy protection from the perspective of system identification, specifically targeting the issue of data tampering under binary observation conditions [21, 22].

System identification in the presence of binary observations and malicious tampering presents several challenges, including insufficient observational information, the presence of Gaussian noise combined with tampering interference, and the degradation of accuracy due to privacy noise [1, 23]. First, binary observations significantly reduce the available data since they only give the relationship of size between the system output and a constant (called the threshold), increasing uncertainty in parameter estimation. The subtle changes in continuous signals may be ignored in binary observations, resulting in information loss and increasing the bias and variance of FIR system parameter estimation, thereby reducing system identification accuracy. Second, tampering by attackers can disrupt the data patterns critical for accurate identification. Finally, privacy encryption mechanisms, while protecting data sensitivity by introducing random noise, may negatively affect estimation accuracy. To address these challenges, this paper applies a dual encryption approach based on differential privacy to both the system input and binary output. This strategy not only prevents attackers from inferring or tampering with the data but also, through improved estimation algorithms and convergence analysis, achieves reliable system parameter identification, balancing the dual objectives of privacy protection and identification accuracy.

The paper proposes a dual differential privacy encryption strategy for system identification: First, Laplace noise is applied to the system input to prevent reverse-engineering of the system's structure or parameters. Then, binary observations are perturbed probabilistically to comply with differential privacy, protecting sensitive information at the output from being accessed or altered. An improved parameter estimation algorithm is introduced, with asymptotic analysis proving the strong consistency and convergence of the proposed method. Numerical simulations demonstrate that high identification accuracy can be maintained, even under various adversarial attack scenarios. This work offers a novel perspective and technical foundation for applying differential privacy in cyber-physical systems, addressing both data security and system identification needs.

This paper makes the following contributions:

- A differential privacy encryption algorithm for FIR system inputs is proposed, enhancing data privacy and security.
- A differential privacy encryption strategy under binary observation conditions is investigated, presenting a discrete 0-1 sequence differential privacy encryption method.
- Parameter estimation results under differential privacy encryption are validated, demonstrating the effectiveness of the proposed method in ensuring both data security and accurate system identification.

The rest of this paper is organized as follows: Section 2 introduces the FIR system model and

discusses the data tampering threats under binary observations, formally stating the research problem. Section 3 details the dual differential privacy encryption methods for both input and output, along with the corresponding random noise models. Section 4 presents the parameter identification algorithm based on encrypted observation signals, accompanied by convergence analysis and strong consistency proofs. Section 5 explores the relationship between encryption budget and identification accuracy, presenting an optimization approach using Lagrange multipliers to determine the optimal encryption strength. Section 6 provides numerical simulations and compares results. Section 7 summarizes the work and discusses future research directions.

2. Problem formulation

The object of this study is a single-input single-output finite impulse response (FIR) system, and its model is given below:

$$y_k = a_1 u_k + a_2 u_{k-1} + \cdots + a_n u_{k-n+1} + d_k, \tag{1}$$

which can be simplified to the following form:

$$y_k = \phi_k^T \theta + d_k, \quad k = 1, 2, \ldots, \tag{2}$$

where u_k is the system input, $\phi_k = [u_k, u_{k-1}, \ldots, u_{k-n+1}]^T$ is the regression vector composed of the input signals, and $\theta = [a_1, a_2, \ldots, a_n]^T$ is the system parameter to be estimated for the system. d_k is independent and identically distributed Gaussian noise, with a mean of 0 and variance σ^2.

The output y_k is measured by a binary sensor with a threshold C, and the resulting binary observation is

$$s_k^0 = I_{\{y_k \le C\}} = \begin{cases} 1, & y_k \le C, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The binary observation model $s_k^0 = I_{\{y_k \le C\}}$ reflects a realistic scenario where only thresholded information is available from sensors. This situation arises frequently in practical FIR systems deployed in industrial monitoring or smart grid applications, where energy-efficient or low-cost sensors output only one-bit comparisons. The observation signal s_k^0 is transmitted through the network to the estimation center, but during the transmission process, it may be subject to data tampering by malicious attackers.

The signal s_k'' received by the estimation center is different from the original signal s_k^0, and the relationship between them is given by:

$$\Pr(s_k'' = 0 | s_k^0 = 1) = p, \quad \Pr(s_k'' = 1 | s_k^0 = 0) = q. \tag{4}$$

We abbreviate the above data tampering attack strategy as (p, q). In this case, the estimation center estimates the system parameter θ based on the received signal s_k'' [24, 25].

Remark 2.1 *Specifically, p denotes the probability that an original output of 0 is maliciously flipped to 1, and q denotes the probability that an original output of 1 is flipped to 0. These parameters characterize a probabilistic tampering model in which binary-valued observations are subject to asymmetric flipping attacks.*

Remark 2.2 *Binary quantization affects both noise characteristics and system identification accuracy: the additive noise becomes nonlinearly embedded in the binary output, and each observation carries limited information.*

Differential privacy is a powerful privacy protection mechanism, and its core idea is to add random noise to the query results, limiting the impact of a single data point on the output, thus ensuring that attackers cannot infer individual information by observing the output. This process not only ensures that the system's input and output data are not directly accessed by attackers, but also effectively prevents attackers from reverse-engineering sensitive system information through input data.

The use of differential privacy in FIR also has its pros and cons. Using differential privacy-based encryption algorithms can limit the impact of data tampering attacks on the FIR system, but encryption itself involves adding noise, which can also lead to data tampering. Therefore, designing an encryption algorithm that meets both the accuracy requirements for parameter estimation and the need to limit data tampering attacks is the key focus of our research.

Figure 1 presents the end-to-end workflow of the proposed privacy-preserving FIR system identification framework under potential data tampering attacks. The process begins with the input signal u_k, which is passed through a Laplace mechanism $\mathrm{Lap}(\Delta u/\epsilon)$ to produce a perturbed input u'_k, ensuring differential privacy at the input level. This perturbed input, together with the system's internal dynamics and additive noise d_k, is processed by the FIR system to produce the output y_k. The output is then passed through a binary-valued sensor that generates a thresholded observation $s_k^0 = I\{y_k \leq C\}$. To preserve output-side privacy, s_k^0 is further randomized via a randomized response mechanism, resulting in $s_k^{0'}$. This privatized observation is then transmitted over a communication network that may be subject to data tampering. During transmission, an attacker may flip the value of $s_k^{0'}$ with certain probabilities, yielding the final corrupted observation s''_k, which is received by the estimation center. The estimation center then applies a robust identification algorithm to estimate the unknown system parameters based on the received data. This pipeline integrates input perturbation, output privatization, and robustness against data manipulation into a unified privacy-preserving system identification framework.

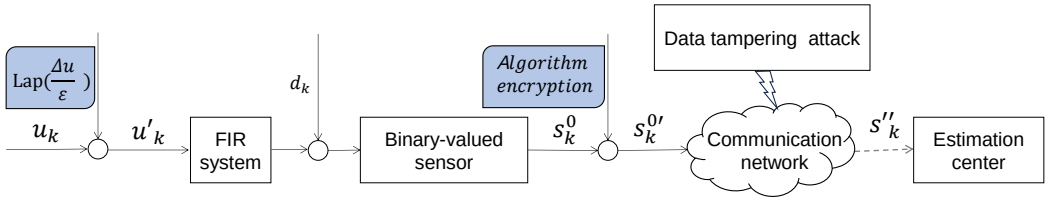

Figure 1. System configuration.

This paper aims to address the following two questions:

- How can the harm caused by data tampering be mitigated or prevented through differential privacy algorithms?
- How can we design a parameter estimation algorithm, and what are the parameter estimation algorithm and the optimal differential privacy algorithm?

3. Differential privacy encryption of system input and output

In this section, we consider an adversarial setting where the collected input-output data may be subject to tampering by untrusted components or compromised observers. Unlike traditional cryptographic methods such as MACs or digital signatures, which aim to detect and reject tampered data, our approach leverages differential privacy as a proactive defense. By adding calibrated randomness to both input and output channels, the encrypted signals inherently reduce the effectiveness of tampered data and enhance the robustness of system identification against malicious interference.

3.1. Differential privacy encryption of input

We first provide the standard definition of differential privacy. Let F be a randomized algorithm. For any two neighboring datasets x and x', their output results satisfy the following inequality:

$$\frac{\Pr(F(x) = S)}{\Pr(F(x') = S)} \leq e^{\varepsilon}. \tag{5}$$

Here, S is a possible output of the algorithm, and ε is the privacy budget, which measures the degree of difference between the output distributions of the algorithm [1, 23]. To ensure the privacy of the input data u_i in the FIR system, we encrypt the input signal using the Laplace mechanism of differential privacy. The Laplace mechanism was selected for input encryption due to its strong ε-differential privacy guarantee, fitting the bounded input signals of FIR systems. The Gaussian mechanism, among others, offering weaker (ε, δ)-differential privacy or may not be suitable for FIR application, was considered unsuitable.

By adding Laplace noise to u_i, we have:

$$u_i' = u_i + Lap\left(\frac{\Delta u_i}{\varepsilon_0}\right), \tag{6}$$

where $Lap\left(\frac{\Delta u_i}{\epsilon_0}\right)$ denotes a random variable drawn from the Laplace distribution with mean 0 and scale parameter $b = \frac{\Delta u_i}{\epsilon_0}$. The probability density function of the Laplace distribution is given by:

$$Lap(x \mid b) = \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right), \quad x \in \mathbb{R}, \tag{7}$$

where Δu_i is the sensitivity of the input u_i, representing the maximum change in the system's input, and the privacy budget ε_0 determines the magnitude of the noise. The sensitivity Δf is used to measure the maximum change of the query function f over neighboring datasets, defined as

$$\Delta f = \max_{x,x'} \|f(x) - f(x')\|, \tag{8}$$

where x and x' are neighboring datasets differing by only one row. From Eq (8), we can derive that

$$\Delta u_i = \max_{u_k, u_k'} \left|y_k - y_k'\right|, \tag{9}$$

where y_k and y'_k are the system outputs corresponding to the current input u_k and its neighboring input u'_k, respectively, as described by Eq (1). Substituting this in, we can get

$$\Delta u_i = \max_{u_k, u'_k} \left| \sum_{i=0}^{n-1} a_{i+1} \cdot (u_{k-i} - u'_{k-i}) \right|. \tag{10}$$

If the distance between the neighboring input sets u_k and u'_k is a fixed value o (or o is the chosen mean), the sensitivity can be expressed as

$$\Delta u_i = \left| \sum_{i=0}^{n-1} a_{i+1} \cdot o \right|. \tag{11}$$

Therefore, from Eqs (6) and (11), the encrypted result after adding noise to u_i is

$$u'_i = u_i + Lap\left(\frac{\left| \sum_{i=0}^{n-1} a_{i+1} \cdot o \right|}{\varepsilon_0} \right). \tag{12}$$

After encrypting the input u_i, the regression vector $\phi_k = [u_k, u_{k-1}, \ldots, u_{k-n+1}]^T$ formed by the input signals will also be affected by the Laplace noise. Let the noise vector be $L_k = [l_1, l_2, \ldots, l_n]^T$, where $l_i = Lap\left(\frac{\Delta u_i}{\varepsilon_0}\right)$. The encrypted regression vector is

$$\phi'_k = \phi_k + L_k. \tag{13}$$

Substituting the encrypted ϕ'_k into the system output expression Eq (2), we obtain the encrypted system output

$$y'_k = (\phi_k + L_k)^T \theta + d_k = \phi_k^T \theta + L_k^T \theta + d_k. \tag{14}$$

This implies that the statistical properties of the system output change under the influence of Laplace noise. By binarizing the system output y'_k, we obtain the observed signal s_k^0. Assume we have a binary sensor with threshold C, which converts y'_k into the binarized observation s_k^0:

$$s_k^0 = I_{\{y'_k \le C\}} = \begin{cases} 1, & y'_k \le C, \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

Based on Eqs (14) and (13), the center will estimate the system parameter θ based on the binarized signal s_k^0. By combining the tampered signal, the output s_k^0 can be expressed as

$$s_k^0 = I_{\{\phi_k^T \theta + L_k^T \theta + d_k \le C\}}. \tag{16}$$

By introducing Laplace noise, we ensure the privacy of the system input u_i. In the following sections, we will continue to discuss how to encrypt s_k^0 under the differential privacy mechanism and estimate the system parameter θ under data tampering attacks.

3.2. Differential privacy encryption of output

In the system, we define $\Pr(s_k^0 = 1) = \lambda_k$, which represents the probability that the signal s_k^0 equals 1 in each period n. To protect this probability information, we apply differential privacy encryption to s_k^0. Specifically, for each s_k^0, there is a probability γ of retaining the original value, and with probability $1 - \gamma$, we perform a negation operation. This operation can be formalized as

$$s_k^{0'} = \begin{cases} s_k^0, & \text{with probability } \gamma, \\ 1 - s_k^0, & \text{with probability } 1 - \gamma. \end{cases} \tag{17}$$

This encryption process introduces randomness and achieves the goal of differential privacy. This encryption method effectively ensures that, whether s_k^0 is 0 or 1, an external attacker would find it difficult to infer the original value of s_k^0 by observing $s_k^{0'}$, thus protecting data privacy. We have

$$\begin{aligned} \Pr(s_k^{0'} = 1) &= \lambda_k \cdot \gamma + (1 - \lambda_k) \cdot (1 - \gamma), \\ \Pr(s_k^{0'} = 0) &= (1 - \lambda_k) \cdot \gamma + \lambda_k \cdot (1 - \gamma). \end{aligned} \tag{18}$$

To analyze the encrypted s_k^0, we construct its likelihood function L to correct the perturbed encryption results. The likelihood function can be expressed as

$$L = [\lambda_k \cdot \gamma + (1 - \lambda_k)(1 - \gamma)]^m \cdot [(1 - \lambda_k) \cdot \gamma + \lambda_k(1 - \gamma)]^{n-m}, \tag{19}$$

where $m = \sum_{i=1}^n X_i$ represents the observed number of times $s_k^{0'} = 1$, and X_i is the value of the i-th encrypted signal $s_k^{0'}$. By maximizing this likelihood function, we can estimate the maximum likelihood estimate for λ_k after encryption. The maximization process leads to the estimation formula

$$\hat{\lambda}_k = \frac{\gamma - 1}{2\gamma - 1} + \frac{1}{2\gamma - 1} \cdot \frac{\sum_{i=1}^n X_i}{n}. \tag{20}$$

This estimate expresses the effect of the encrypted $s_k^{0'}$ on the true λ_k.

To further analyze the unbiasedness of the estimate, we can analyze the mathematical expectation and prove that the estimate is unbiased. The specific derivation is as follows:

$$\mathbb{E}(\hat{\lambda}_k) = \frac{1}{2\gamma - 1} \left[\gamma - 1 + \frac{1}{n} \sum_{i=1}^n X_i \right] = \lambda_k. \tag{21}$$

Based on the unbiased estimate, we can obtain the expected number N of signals equal to 1 in each period n as

$$N = \hat{\lambda}_k \times n = \frac{\gamma - 1}{2\gamma - 1} n + \frac{1}{2\gamma - 1} \cdot \sum_{i=1}^n X_i. \tag{22}$$

According to the definition of differential privacy, the relationship between the privacy parameter ε_0 and γ is

$$\gamma \leq e^{\varepsilon_1} \cdot (1 - \gamma). \tag{23}$$

By further derivation, we can obtain the calculation formula for the privacy budget ε_0 as

$$\varepsilon_1 = \ln\left(\frac{\gamma}{1 - \gamma}\right). \tag{24}$$

Through the above encryption mechanism [12], we ensure that the system satisfies differential privacy while maintaining high robustness against interference from attackers on the observation of s_k. During data transmission, the observed value s_k^0 may be subjected to attacks, and the tampering probability of the observed value s_k^0 is

$$\begin{cases} \Pr(s_k'' = 0|s_k^{0'} = 1) = p, \\ \Pr(s_k'' = 1|s_k^{0'} = 0) = q. \end{cases} \tag{25}$$

Assume that the system input $\{u_k\}$ is periodic with a period of n, i.e.,

$$u_{k+n} = u_k. \tag{26}$$

Let $\pi_1 = \phi_1'^T, \pi_2 = \phi_2'^T, \ldots, \pi_n = \phi_n'^T$, and the cyclic matrix formed by u_k is

$$\Phi = \left[\pi_1^T, \pi_2^T, \cdots, \pi_n^T\right]^T. \tag{27}$$

After the preservation process applied to both inputs and outputs of u_k and s_k^0, during data transmission, the observed value s_k^0 will be attacked, and both active and passive attacks on the encrypted data will be limited and disturbed by noise. In the next section, we will provide a parameter estimation algorithm.

4. Algorithm design and convergence of parameter identification

4.1. Algorithm design of parameter identification

Under external attacks, and in the sense of the Cramér-Rao lower bound, the optimal estimation algorithm for the encrypted parameter $\widehat{\theta}_N$ is given below:

$$\widehat{\theta}_N = \Phi^{-1}[\xi_{N,1}, \ldots, \xi_{N,n}]^T, \tag{28}$$

$$\xi_{N,i} = C - F^{-1}\left(\frac{1}{L_N}\sum_{l=1}^{L_N} s''_{(l-1)n+i}\right), \tag{29}$$

where Φ^{-1} is the inverse matrix in Eq (27), C is the threshold in Eq (3), F^{-1} is the inverse cumulative distribution function, and $s''_{(l-1)n+i}$ is the doubly encrypted s_k^0 in each period. For a data length N, it is divided into $L_N = \left\lfloor \frac{N}{n} \right\rfloor$, which is the integer part of N divided by the input period n, representing the number of data segments L_N [26].

4.2. Convergence of parameter identification

Definition 4.1 *An algorithm is said to have strong consistency if the estimated value $\hat{\theta}_N$ produced by the algorithm satisfies that as the sample size $N \to \infty$, the estimate converges to the true value of the parameter θ (or the encrypted true value $\bar{\theta}$) with probability 1, i.e., $\hat{\theta}_N \to \theta$ or $\hat{\theta}_N \to \bar{\theta}$ w.p.1 as $N \to \infty$.*

This conclusion follows the analysis in [24], which establishes strong consistency of estimators under binary-valued observation and bounded noise conditions.

Theorem 4.1 *For the parameter estimation under the attack strategy (p, q) in Eq (28) with encrypted conditions, the expanded form of the true value $\bar{\theta}$ is*

$$\begin{aligned}\hat{\theta}_N \to \bar{\theta} = \Phi^{-1}\Big[&C - F^{-1}\Big((2\gamma-1)(1-p-q)F_{Z_1}(C-\pi_1^T\theta) + (1-p)(1-\gamma) + q\gamma\Big), \ldots, \\ &C - F^{-1}\Big((2\gamma-1)(1-p-q)F_{Z_n}(C-\pi_n^T\theta) + (1-p)(1-\gamma) + q\gamma\Big)\Big]^T\end{aligned} \tag{30}$$

and the algorithm has strong consistency. Here, θ represents the parameter vector to be estimated in the system, and C, γ, Φ^{-1}, and π_k^T are given in Eqs (15), (17), and (27), respectively. F^{-1} is the inverse cumulative distribution function of noise d_k, and $F_{Z_k}(z)$ is the cumulative distribution function of the combined noise from Laplace noise $l_k^T \cdot \theta$ and Gaussian noise d_k.

Proof: Under data tampering attacks, in the period N, each period's s_k'' is independent and identically distributed. For $k = (l-1)n + i$, where $i = 1, \ldots, n$, the regression vector ϕ_k is periodic with period n, i.e., $\phi_{k+n} = \phi_k$. Define $\pi_i = \phi_i$, so $\phi_{(l-1)n+i} = \pi_i$. Since $\phi_{(l-1)n+i} = \pi_i$, and the noise terms L_k, d_k are i.i.d., the sequence $\{s''_{(l-1)n+i}\}_{l=1}^{L_N}$ is i.i.d. for each i.

By the strong law of large numbers, for each i:

$$\frac{1}{L_N}\sum_{l=1}^{L_N} s''_{(l-1)n+i} \to \mathbb{E}[s''_{(l-1)n+i}] \quad \text{w.p.1 as } N \to \infty. \tag{31}$$

Next, we compute $\mathbb{E}[s''_{(l-1)n+i}]$. Let $k = (l-1)n + i$. The encrypted input is $\phi'_k = \phi_k + L_k = \pi_i + L_k$, and the system output is:

$$y'_k = (\pi_i + L_k)^T\theta + d_k = \pi_i^T\theta + L_k^T\theta + d_k. \tag{32}$$

Define $Z_k = L_k^T\theta + d_k$. The binary signal is:

$$s_k^0 = I_{\{y'_k \le C\}} = I_{\{\pi_i^T\theta + Z_k \le C\}} = I_{\{Z_k \le C - \pi_i^T\theta\}}. \tag{33}$$

Thus:

$$\lambda_k = \Pr(s_k^0 = 1) = \Pr(Z_k \le C - \pi_i^T\theta) = F_{Z_i}(C - \pi_i^T\theta). \tag{34}$$

The output encryption gives $s_k^{0'}$:

$$\mathbb{E}[s_k^{0'}] = \gamma\lambda_k + (1-\gamma)(1-\lambda_k) = (2\gamma-1)\lambda_k + (1-\gamma). \tag{35}$$

After the attack strategy (p, q):

$$\mathbb{E}[s_k''] = (1-p)\mathbb{E}[s_k^{0'}] + q(1 - \mathbb{E}[s_k^{0'}]) = (1-p-q)\mathbb{E}[s_k^{0'}] + q. \tag{36}$$

Substitute $\mathbb{E}[s'_k]$:

$$\mathbb{E}[s_k''] = (1-p-q)[(2\gamma-1)F_{Z_i}(C - \pi_i^T\theta) + (1-\gamma)] + q. \tag{37}$$

Simplify:

$$\mathbb{E}[s_k''] = (2\gamma-1)(1-p-q)F_{Z_i}(C - \pi_i^T\theta) + (1-p)(1-\gamma) + q\gamma. \tag{38}$$

Define:

$$\eta_i = (2\gamma-1)(1-p-q)F_{Z_i}(C - \pi_i^T\theta) + (1-p)(1-\gamma) + q\gamma. \tag{39}$$

Thus:

$$\frac{1}{L_N}\sum_{l=1}^{L_N} s''_{(l-1)n+i} \to \eta_i \quad \text{w.p.1}\,. \tag{40}$$

From the definition of $\hat{\theta}_N$:

$$\xi_{N,i} = C - F^{-1}\left(\frac{1}{L_N}\sum_{l=1}^{L_N} s''_{(l-1)n+i}\right) \to C - F^{-1}(\eta_i) \quad \text{w.p.1}\,. \tag{41}$$

Therefore:

$$\bar{\theta}_N = \Phi^{-1}\left[\xi_{N,1},\ldots,\xi_{N,n}\right]^T \to \Phi^{-1}\left[C - F^{-1}(\eta_1),\ldots,C - F^{-1}(\eta_n)\right]^T. \tag{42}$$

Note that the total observation noise comprises the superposition of the original Gaussian noise d_k and the added Laplace noise due to differential privacy encryption. Since both noise sources are independent and have finite variance, their sum forms a new independent noise process with bounded second moments. This preserves the excitation condition needed for strong consistency. Here, the strong consistency follows from the strong law of large numbers. From Eqs (27)–(29) and (34), the expanded form of $\bar{\theta}$ is obtained, thus completing the proof.

5. The optimality of encryption

5.1. Modeling of the optimization problem for the optimal encryption strategy

Our goal is to study how to maximize encryption within the permissible range of estimation errors.

To clarify the impact of different encryptions on the results, we analyze them separately based on the independence of input and output noise.

For the input noise encryption, the noise variance σ_L^2 introduced by input encryption, which affects $\hat{\theta}$ through the system model, is as follows:

$$D_0 = \sigma_L^2 = 2\left(\frac{1}{\varepsilon_0}\right)^2 \sum_{i=1}^{n}(\theta_i \Delta u_i)^2. \tag{43}$$

For the output noise encryption, since s'_k is a binary random variable, its variance is

$$\mathrm{Var}(s'_k) = \Pr(s'_k = 1)(1 - \Pr(s'_k = 1)) = (1 - \gamma + \lambda_k(2\gamma - 1))(\gamma - \lambda_k(2\gamma - 1)). \tag{44}$$

From Eq (20), we can derive the variance of the maximum likelihood estimate $\hat{\lambda}_k$ as

$$\mathrm{Var}(\hat{\lambda}_k) = \left(\frac{1}{(2\gamma - 1)n}\right)^2 \sum_{k=1}^{n} \mathrm{Var}(s'_k). \tag{45}$$

Substituting the result from Eq (44), we can get

$$\mathrm{Var}(\hat{\lambda}_k) = \frac{\mathrm{Var}(s'_k)}{(2\gamma - 1)^2 n} = \frac{(1 - \gamma + \lambda_k(2\gamma - 1))(\gamma - \lambda_k(2\gamma - 1))}{(2\gamma - 1)^2 n}. \tag{46}$$

The parameter estimate $\hat{\theta}$ depends on λ_k, and D_1 is defined as

$$D_1 = \mathrm{Var}(\hat{\lambda}_k) = \frac{(1-\gamma+\lambda_k(2\gamma-1))(\gamma-\lambda_k(2\gamma-1))}{(2\gamma-1)^2 n}. \tag{47}$$

To solve the optimal encryption problem, we model the optimization of the optimal encryption strategy:

$$\begin{aligned} &\max_{\varepsilon_0,\varepsilon_1} (D_0 + D_1) \\ &s.t.\ \|\bar{\theta} - \theta\| \le \epsilon, \end{aligned} \tag{48}$$

where ε_0 and ε_1 are the differential privacy budgets for input and output, respectively, and D_0 and D_1 are the functions that measure the impact of input and output encryption noise on the parameter estimate $\bar{\theta}$, where $D_0 = 2\left(\frac{1}{\varepsilon_0}\right)^2 \sum_{i=1}^{n}(\theta_i \Delta u_i)^2$ and $D_1 = \frac{\mathrm{Var}(s'_k)}{(2\gamma-1)^2 n}$. $\|\bar{\theta} - \theta\|$ is the parameter estimation error, which must be less than or equal to the given threshold ϵ, and both $\bar{\theta}$ and θ must be greater than zero.

Theorem 5.1 *In the case of differential privacy with the preservation process applied to both inputs and outputs, σ_Z^2 is monotonically decreasing with respect to the privacy parameter of input ε_0, which indicates the smaller ε_0 is, the stronger the privacy protection provides, but the worse the estimation accuracy is.*

Proof: $L_k^T\theta$ is a linear combination of Laplace noise, and d_k is Gaussian noise, while Z_k is the sum of these two. Then, the variance of each l_i is

$$\mathrm{Var}(l_i) = 2\left(\frac{\Delta u_i}{\varepsilon_0}\right)^2. \tag{49}$$

Thus,

$$\mathrm{Var}(L_k^T\theta) = \sum_{i=1}^{n} \theta_i^2 \,\mathrm{Var}(l_i) = 2\left(\frac{1}{\varepsilon_0}\right)^2 \sum_{i=1}^{n}(\theta_i \Delta u_i)^2. \tag{50}$$

The total variance of Z_k is

$$\sigma_Z^2 = \mathrm{Var}(Z_k) = \mathrm{Var}(L_k^T\theta) + \mathrm{Var}(d_k) = \sigma_L^2 + \sigma_d^2, \tag{51}$$

where

$$\sigma_L^2 = 2\left(\frac{1}{\varepsilon_0}\right)^2 \sum_{i=1}^{n}(\theta_i \Delta u_i)^2. \tag{52}$$

To rigorously prove monotonicity, we can verify it by analyzing the derivative of σ_Z^2 with respect to ε_0. Taking the derivative of σ_Z^2 with respect to ε_0,

$$\frac{d\sigma_Z^2}{d\varepsilon_0} = \frac{d}{d\varepsilon_0}\left(2\left(\frac{1}{\varepsilon_0}\right)^2 \sum_{i=1}^{n}(\theta_i \Delta u_i)^2 + \sigma_d^2\right). \tag{53}$$

Since σ_d^2 is a constant and does not depend on ε_0, we get

$$\frac{d\sigma_Z^2}{d\varepsilon_0} = -4\frac{1}{\varepsilon_0^3}\sum_{i=1}^{n}(\theta_i\Delta u_i)^2. \tag{54}$$

Since $\varepsilon_0 > 0$ and $\sum_{i=1}^{n}(\theta_i\Delta u_i)^2 > 0$, we have $\frac{d\sigma_Z^2}{d\varepsilon_0} < 0$. This implies that as ε_0 increases, σ_Z^2 decreases monotonically. According to system estimation theory, as the noise variance decreases, the estimation accuracy improves. Therefore, as ε_0 increases, σ_Z^2 decreases, making the estimated result $\hat{\theta}$ closer to the true value θ. This concludes the proof.

Theorem 5.2 *In the case of differential privacy with the preservation process applied to both inputs and outputs, $\bar{\theta}$ is monotonically decreasing with respect to ε_1. The smaller ε_1 is, the stronger the privacy protection provides.*

Proof: From Eq (24), we have $\gamma = \frac{e^{\varepsilon_1}}{e^{\varepsilon_1}+1}$ and $2\gamma - 1 = \frac{e^{\varepsilon_1}-1}{e^{\varepsilon_1}+1}$.

Substituting into Eq (30), we can express $\bar{\theta}$ as

$$\begin{aligned}\bar{\theta} = \Phi^{-1}\Bigg[&C - F^{-1}\left(\frac{e^{\varepsilon_1}-1}{e^{\varepsilon_1}+1}\cdot F_{Z_1}(C-\pi_1^T\theta) + 1 - F_{Z_1}(C-\pi_1^T\theta) - \frac{e^{\varepsilon_1}}{e^{\varepsilon_1}+1}\right),\ldots,\\ &C - F^{-1}\left(\frac{e^{\varepsilon_1}-1}{e^{\varepsilon_1}+1}\cdot F_{Z_n}(C-\pi_n^T\theta) + 1 - F_{Z_n}(C-\pi_n^T\theta) - \frac{e^{\varepsilon_1}}{e^{\varepsilon_1}+1}\right)\Bigg]^T.\end{aligned} \tag{55}$$

By simplifying each term, we obtain

$$\frac{(e^{\varepsilon_1}-1)F_{Z_i}(C-\pi_i^T\theta) + (e^{\varepsilon_1}+1)(1-F_{Z_i}(C-\pi_i^T\theta)) - e^{\varepsilon_1}}{e^{\varepsilon_1}+1}. \tag{56}$$

Simplifying the numerator, it can be inferred that

$$\begin{aligned}&(e^{\varepsilon_1}-1)F_{Z_i}(C-\pi_i^T\theta) + (e^{\varepsilon_1}+1)(1-F_{Z_i}(C-\pi_i^T\theta)) - e^{\varepsilon_1}\\ &= e^{\varepsilon_1}F_{Z_i}(C-\pi_i^T\theta) - F_{Z_i}(C-\pi_i^T\theta) + e^{\varepsilon_1} + 1 - e^{\varepsilon_1}F_{Z_i}(C-\pi_i^T\theta) - F_{Z_i}(C-\pi_i^T\theta) - e^{\varepsilon_1}\\ &= -2F_{Z_i}(C-\pi_i^T\theta) + 1.\end{aligned} \tag{57}$$

Substituting each term with $\frac{-2F_{Z_i}(C-\pi_i^T\theta)+1}{e^{\varepsilon_1}+1}$, the expression for $\bar{\theta}$ becomes

$$\bar{\theta} = \Phi^{-1}\left[C - F^{-1}\left(\frac{-2F_{Z_1}(C-\pi_1^T\theta)+1}{e^{\varepsilon_1}+1}\right),\ldots,C - F^{-1}\left(\frac{-2F_{Z_n}(C-\pi_n^T\theta)+1}{e^{\varepsilon_1}+1}\right)\right]^T. \tag{58}$$

This concludes the proof.

5.2. The optimal solution for modeling

In parameter estimation, the estimation error is typically related to the variance of the observation noise. Differential privacy encryption for both input and output introduces noise that affects the bias and variance of parameter estimation.

The expected value of the parameter estimate $\hat{\theta}$ is $\bar{\theta}$, i.e.,

$$E[\hat{\theta}] = \bar{\theta}.$$

The deviation between $\bar{\theta}$ and the true parameter θ is caused by the encryption noise. Our goal is to solve for $\|\bar{\theta} - \theta\|$, and express it as a function of ε_0 and ε_1.

The mean square error (MSE) of the parameter estimate can be expressed as the square of the bias plus the variance:

$$\text{MSE}(\hat{\theta}) = \|E[\hat{\theta}] - \theta\|^2 + \text{Var}(\hat{\theta}). \tag{59}$$

Since we are concerned with the bias of the estimate $\|\bar{\theta} - \theta\|$, and the variance of the estimate also depends on D_0 and D_1, we need to express both the bias and variance as functions of ε_0 and ε_1.

To accurately express $\|\bar{\theta} - \theta\|$, we start from the root mean square error (RMSE):

$$\text{RMSE}(\hat{\theta}) = \sqrt{\text{trace}(\text{Var}(\hat{\theta}))}. \tag{60}$$

The total variance of the system's parameter estimate is

$$\text{Var}(\hat{\theta}) = \left(\sigma_d^2 + D_0 + D_1\right)(\Phi^T\Phi)^{-1}, \tag{61}$$

where σ_d^2 is the system's inherent noise variance, D_0 and D_1 are the variances introduced by input and output encryption noise, and Φ is the matrix from Eq (27).

Thus, the root mean square error of the estimate can be expressed as

$$\text{RMSE}(\hat{\theta}) = \sqrt{\text{trace}\left(\left(\sigma_d^2 + D_0 + D_1\right)(\Phi^T\Phi)^{-1}\right)}. \tag{62}$$

Since $\Phi^T\Phi$ is a known design matrix, we can represent the estimation error as

$$\|\bar{\theta} - \theta\| \leq k\sqrt{\sigma_d^2 + D_0 + D_1}, \tag{63}$$

where $k = \sqrt{\text{trace}((\Phi^T\Phi)^{-1})}$.

Based on the above derivation, the optimization problem can be reformulated as:

$$\begin{aligned} \max_{\varepsilon_0,\varepsilon_1} \quad & D_0 + D_1 \\ \text{s.t.} \quad & \|\bar{\theta} - \theta\| \leq \epsilon \\ & \sigma_d^2 + D_0 + D_1 \leq \epsilon_{\text{eff}}^2, \end{aligned}$$

where $\epsilon_{\text{eff}}^2 = \epsilon^2/k^2 - \sigma_d^2$.

Now, we construct the Lagrangian objective function

$$L(\varepsilon_0, \varepsilon_1, \lambda) = D_0 + D_1 - \lambda\left(D_0 + D_1 - \epsilon_{\text{eff}}^2\right). \tag{64}$$

For ε_0, we have

$$D_0 = 2\left(\frac{1}{\varepsilon_0}\right)^2 \sum_{i=1}^{n} (\theta_i \Delta u_i)^2. \tag{65}$$

Take partial derivatives with respect to ε_0, and we have

$$\frac{\partial D_0}{\partial \varepsilon_0} = -4\left(\frac{1}{\varepsilon_0^3}\right)\sum_{i=1}^{n} (\theta_i \Delta u_i)^2. \tag{66}$$

For ε_1, we have

$$D_1 = \frac{\mathrm{Var}(s_k')}{(2\gamma - 1)^2 n}, \tag{67}$$

where $\gamma = \frac{e^{\varepsilon_1}}{e^{\varepsilon_1}+1}$, and we can compute

$$\frac{\partial \gamma}{\partial \varepsilon_1} = \frac{e^{\varepsilon_1}}{(e^{\varepsilon_1}+1)^2}. \tag{68}$$

Now, further differentiate D_1. First, rewrite D_1 as

$$D_1 = \frac{\mathrm{Var}(s_k^{0'})}{(2\gamma - 1)^2 n}. \tag{69}$$

Differentiate $(2\gamma - 1)^2$ with respect to ε_1, and it can be obtained that

$$\frac{\partial}{\partial \varepsilon_1}[(2\gamma - 1)^2] = 4 \cdot (2\gamma - 1) \cdot \frac{e^{\varepsilon_1}}{(e^{\varepsilon_1}+1)^2}. \tag{70}$$

Now, applying the chain rule, we have

$$\frac{\partial D_1}{\partial \varepsilon_1} = -\frac{8\,\mathrm{Var}(s_k^{0'}) \cdot (2\gamma - 1)e^{\varepsilon_1}}{n(e^{\varepsilon_1}+1)^4}. \tag{71}$$

By the above analysis, the optimal values for ε_0 and ε_1 need to satisfy the following constraint:

$$D_0 + D_1 = \epsilon_{\text{eff}}^2. \tag{72}$$

Taking partial derivatives with respect to ε_0 and ε_1, we get

$$\frac{\partial L}{\partial \varepsilon_0} = \frac{\partial D_0}{\partial \varepsilon_0} - \lambda \frac{\partial}{\partial \varepsilon_0}(D_0 + D_1) = 0, \tag{73}$$

$$\frac{\partial L}{\partial \varepsilon_1} = \frac{\partial D_1}{\partial \varepsilon_1} - \lambda \frac{\partial}{\partial \varepsilon_1}(D_0 + D_1) = 0. \tag{74}$$

Substituting the previously computed derivatives, we have

$$-4\left(\frac{1}{\varepsilon_0^3}\right)\sum_{i=1}^{n}(\theta_i \Delta u_i)^2 - \lambda \cdot 0 = 0, \tag{75}$$

$$\frac{-8\,\mathrm{Var}(s_k^{0'})}{n} \cdot \frac{(2\gamma - 1)e^{\varepsilon_1}}{(e^{\varepsilon_1}+1)^4} - \lambda \cdot 0 = 0. \tag{76}$$

For ε_0, since $\frac{\partial}{\partial \varepsilon_0}(D_0 + D_1) = 0$, the derivative is completely determined by the objective function. To maximize the objective function under the constraint, we first check if the condition holds. Then, we can solve for ε_0 to satisfy the given constraint.

For ε_1, we also solve for λ to obtain the optimal solution.

The optimal solutions for ε_0 and ε_1 are

$$\varepsilon_0 = \sqrt{\frac{2}{1 + \epsilon_{\text{eff}}^2}}, \tag{77}$$

$$\varepsilon_1 = \frac{2}{3(1 + \epsilon_{\text{eff}}^2)}. \tag{78}$$

Remark 5.1 *These solutions are confirmed based on the value of λ, and the specific numerical solution can be obtained through numerical computation.*

6. Numerical simulation

Consider the following system:

$$\begin{cases} y_k = a_1 u_k + a_2 u_{k-1} + d_k, \\ s_k = I_{\{y_k \le C\}}, \quad k = 1, \ldots, N, \end{cases} \tag{79}$$

where $a_1 = 10$ and $a_2 = 5$ are the unknown parameters of the system, d_k is $\mathcal{N}(0, 100)$, and the threshold $C = 7$. The sample length for this simulation is set to $N = 5000$. The system input signals u_k are set as cyclic signals, with a value of $[3, 5]$ in one cycle. The threshold ϵ is set to 12. The sensitivity Δu is computed based on the maximum possible deviation between two neighboring input sequences, resulting in $o = |5 - 3| = 2$.

6.1. System input under different intensities of Laplacian noise

We encrypt the input signal u_k with Laplace noise under three different privacy budget parameters and simulate to verify its effects under different encryption strengths. According to the Laplace noise mechanism, the added noise follows the distribution:

$$u'_k = u_k + \mathrm{Lap}\left(\frac{\Delta u}{\varepsilon}\right), \tag{80}$$

where u'_k is the encrypted input signal, Δu represents the sensitivity of the input signal, and ε is the privacy budget parameter.

Figure 2 shows the u_k signal under three different privacy budget parameters. The red line represents the strong encryption effect with a privacy budget of 0.1, the green line represents the moderate encryption effect with a privacy budget of 0.5, and the blue line represents the weak encryption effect with a privacy budget of 1.0. From the figure, it can be seen that as the privacy budget increases, the encryption strength decreases, and the amplitude of the noise significantly reduces. This result verifies the relationship between the strength of the Laplace noise and the privacy budget parameter.

From Figure 2, it can be observed that when the privacy budget parameter is small, i.e., under strong encryption, the noise amplitude is the largest, and the impact on the original signal u_k is most significant. On the other hand, under weak encryption, the noise amplitude is the smallest, and the impact on the original signal is minimal. The results of Laplace noise encryption at different strengths indicate that the stronger the encryption, the larger the noise amplitude, the more significant the interference on the original signal u_k, and the better the privacy protection; the weaker the encryption, the smaller the noise amplitude, the less interference on the original signal, and the weaker the privacy protection.

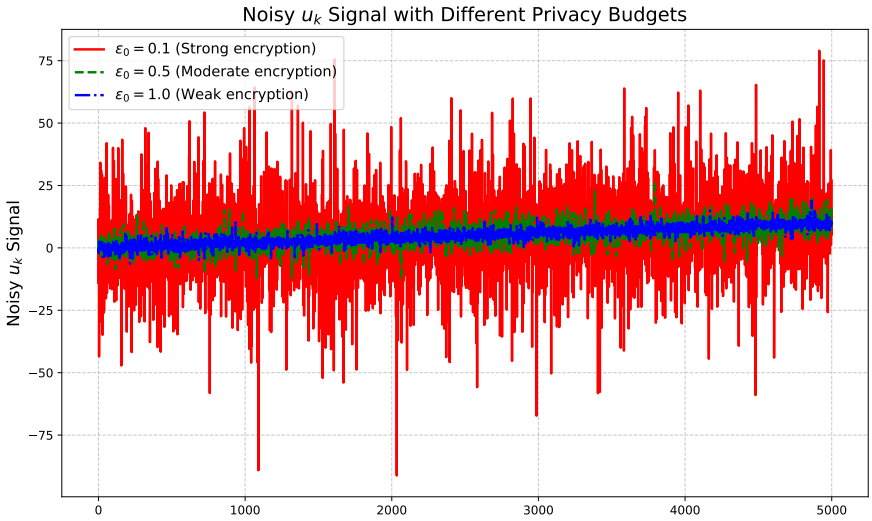

Figure 2. The u_k signal under different privacy budgets.

6.2. *System output under binarization and noise addition*

The system output's binarized observation signal s_k has been encrypted with different privacy budget parameters. The encryption process introduces Laplace noise with varying strengths on the binarized s_k to protect privacy. After encryption, each s_k retains its original value with probability γ and is flipped with probability $1 - \gamma$. This process is expressed as

$$s'_k = \begin{cases} s_k, & \text{with probability } \gamma, \\ 1 - s_k, & \text{with probability } 1 - \gamma. \end{cases} \tag{81}$$

Figure 3 shows the binarized encrypted signal s'_k generated under three different privacy budget parameters ($\varepsilon = 0.1$, $\varepsilon = 0.5$, $\varepsilon = 1.0$), with the red, green, and blue lines corresponding to strong, moderate, and weak encryption effects, respectively.

From Figure 3, it is evident that the encryption strength significantly affects the degree of randomization of s'_k. As the privacy budget parameter ε decreases (i.e., stronger encryption), the probability of flipping s'_k increases, leading to higher randomization of s'_k. Conversely, when the privacy budget increases (i.e., weaker encryption), s'_k closely approximates the original binarized signal s_k. This result verifies the relationship between encryption strength and signal randomization, namely, the stronger the encryption, the higher the randomization of s'_k; the weaker the encryption, the more s'_k retains the pattern of the original signal.

6.3. *Simulation validation of Theorem 4.1*

Equation (30) provides the full expression for parameter estimation $\widehat{\theta}$ under encryption, and we observe the accuracy of the parameter estimation in the simulation.

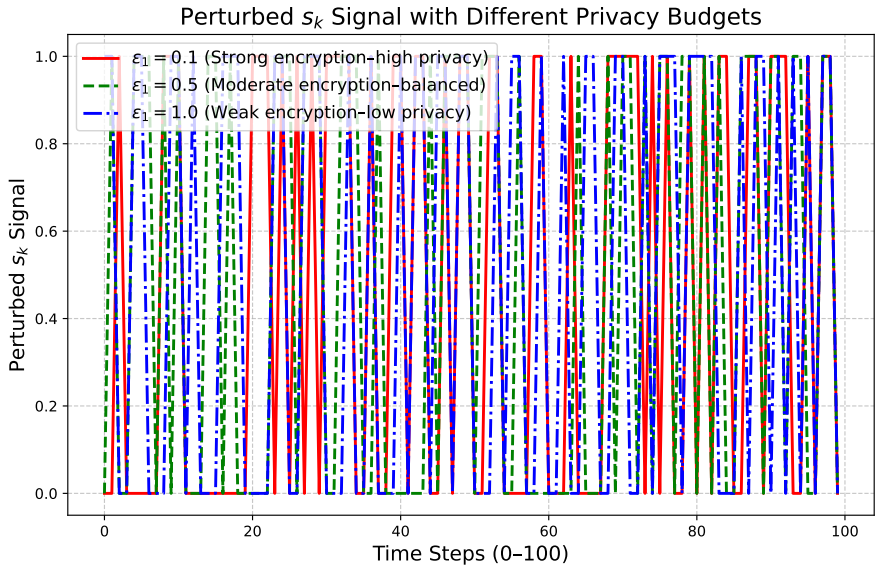

Figure 3. Simulation of s_k after noise is added following binarization.

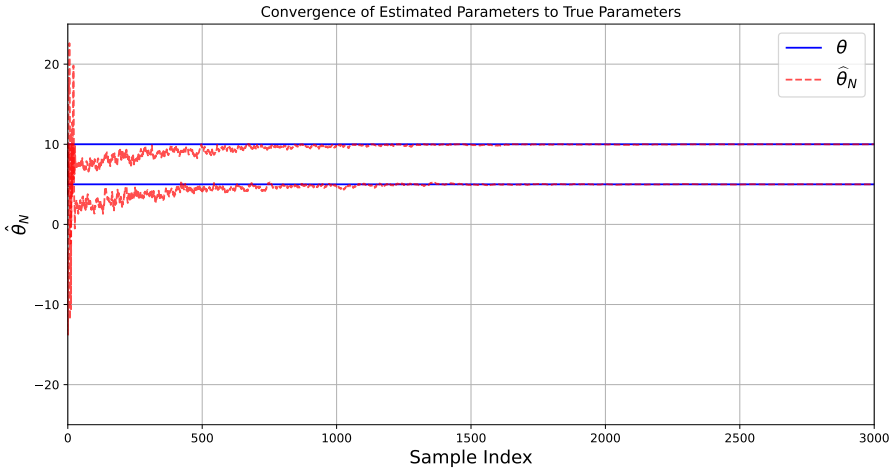

Figure 4. Simulation results: Parameter estimation error and deviation from the true value under different privacy budgets.

The selected privacy budget values are $\varepsilon_0 = 1.0$ and $\varepsilon_1 = 1.0$. The simulation results are shown in Figure 4, where, as the sample size increases, the parameter estimation approaches the encrypted value $\bar{\theta}$.

6.4. Error analysis under different noise intensities

We evaluate the impact of the privacy budget ε_0 encryption on the input u_k through simulations. Figure 5 shows the encryption effects under different privacy budgets. The error impact is represented

by $\left\|\widehat{\theta}-\theta\right\|$. The red line represents strong encryption with a privacy budget of 0.1, the green line represents moderate encryption with a privacy budget of 0.5, and the blue line represents weak encryption with a privacy budget of 1.0. From the error results, the simulation graph confirms the monotonicity Theorem 5.1. The smaller the privacy budget, the stronger the encryption, the larger the noise amplitude, and the greater the error between the estimated value and the true value. Conversely, when the privacy budget increases, the noise amplitude decreases, and the estimation error becomes smaller.

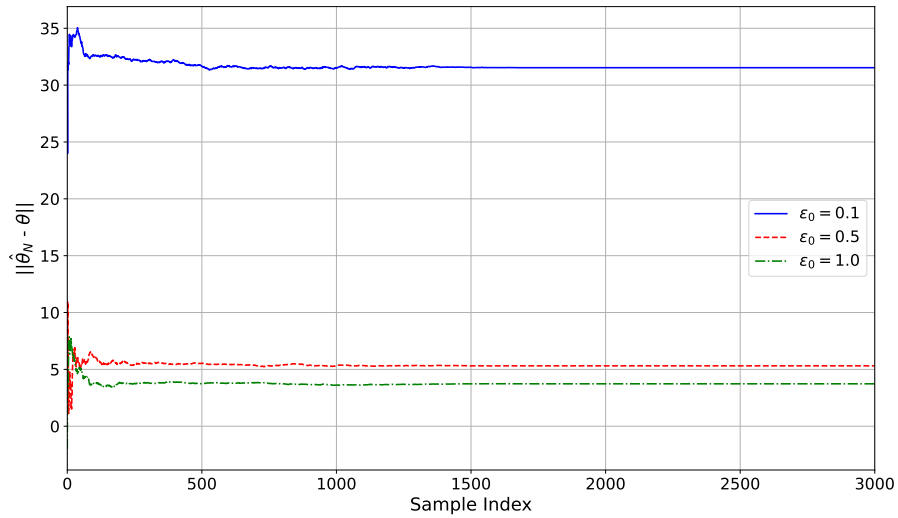

Figure 5. Encryption effects of the input signal u_k under different privacy budgets.

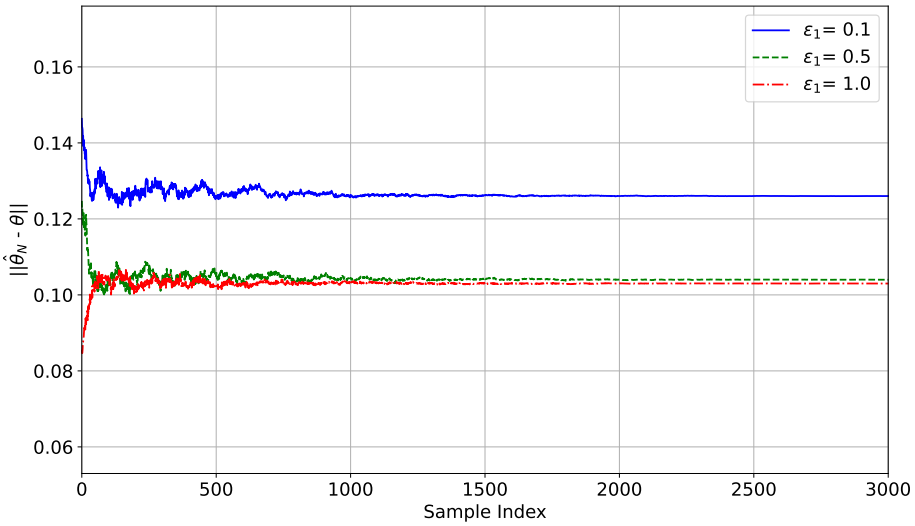

Figure 6. Convergence curve of parameter estimation error under different privacy budgets ε_1.

Next, we simulate the effect of adding differential privacy noise to the system output s_k and observe the impact of different privacy budget parameters ε_1 on the parameter estimation error. The simulation results are shown in Figure 6, which display the changes in parameter estimation error as the sample size varies for $\varepsilon_1 = 0.1$, $\varepsilon_1 = 0.5$, and $\varepsilon_1 = 1.0$. The result shown in Figure 6 is consistent with Theorem 5.2.

6.5. Evaluation of input magnitude influence under dual privacy protection

In this section, we further investigate how the magnitude of input signals affects the identification accuracy of FIR systems under dual differential privacy constraints. Specifically, both the input vector $\boldsymbol{u}_k = [u_k, u_{k-1}]^\top$ and the binary-valued output s'_k are simultaneously perturbed by independent Laplace noise satisfying ϵ-differential privacy, denoted as ϵ_0 and ϵ_1, respectively.

To evaluate the estimation behavior under different input scales, we simulate two configurations: one using low-magnitude inputs $(u_k, u_{k-1}) = (3, 5)$, and another using higher-magnitude inputs $(30, 50)$, while keeping all other conditions identical (privacy budgets $\epsilon_0 = \epsilon_1 = 0.5$, threshold $C = 70$, and Gaussian noise variance $\sigma^2 = 100$).

The simulation results (Figure 7) reveal a significant contrast. While the lower magnitude input case results in relatively fluctuating estimation error due to stronger sensitivity to both noise and binarization, the higher input magnitude quickly leads to a nearly constant estimation error curve. This phenomenon is attributed to the high values of $y_k = a_1 u_k + a_2 u_{k-1} + d_k$ being mostly above the threshold C, rendering the binary outputs nearly constant and less informative for learning.

This experiment highlights the trade-off between input energy and privacy-resilient identifiability in FIR systems. It emphasizes the importance of designing input signals that balance observability with privacy-preserving distortion in binary measurement contexts.

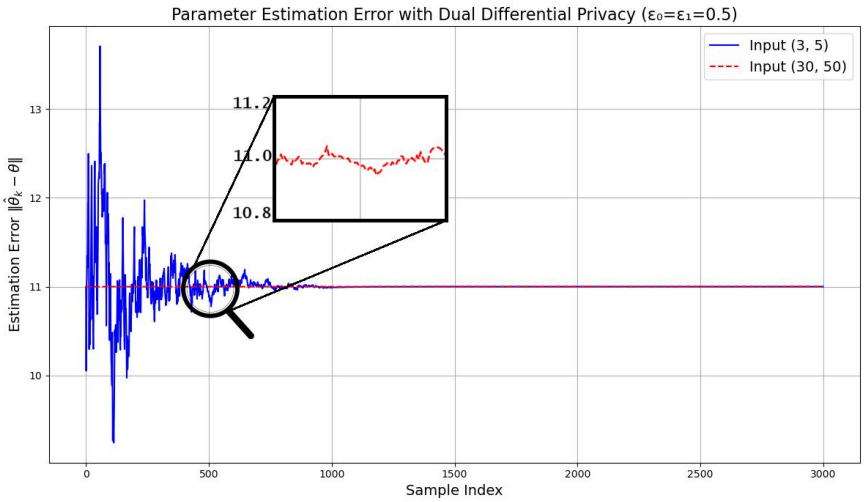

Figure 7. Estimation error comparison under dual differential privacy: $(3, 5)$ vs. $(30, 50)$ input.

6.6. *Comparison of noise distributions*

We apply two noise distributions to the binarized output s'_k:

1) **Gaussian noise**: Zero mean, variance $V = 1.0$, standard deviation $\sigma = \sqrt{V} = 1.0$.
2) **Laplacian noise**: Zero mean, variance $V = 1.0$, scale parameter $b = \sqrt{V/2} \approx 0.707$.

Both distributions maintain equal variance for a fair comparison. We simulate $N = 5000$ samples, adding Gaussian and Laplacian noise to s'_k and estimating parameters iteratively. The exponential moving average (EMA) of the estimation error is computed for both noise types.

Figure 8 presents the EMA estimation error for Gaussian and Laplacian noise over 5000 samples.

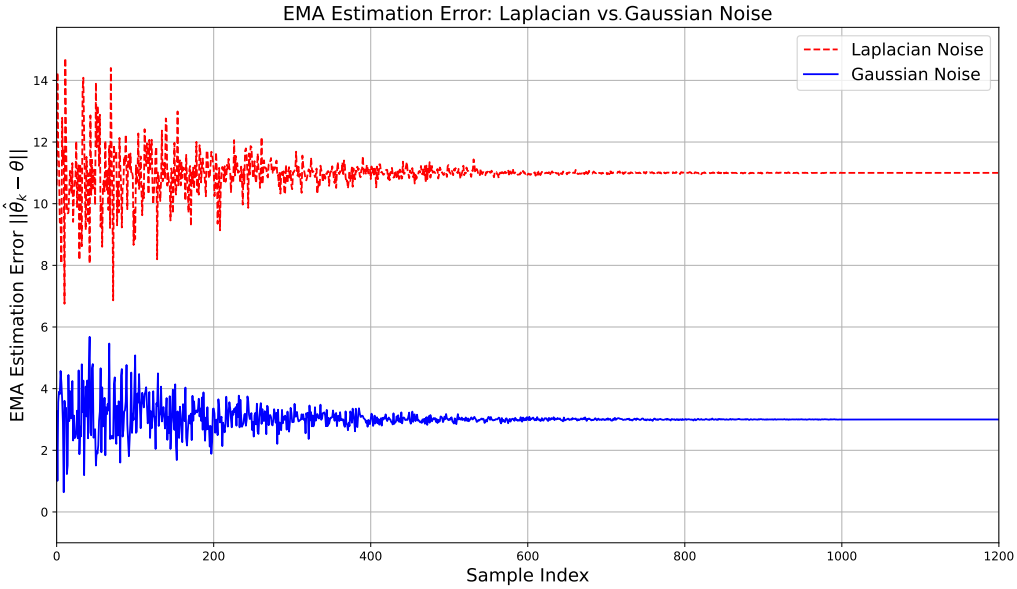

Figure 8. EMA estimation error for Gaussian (blue) and Laplacian (red) noise with equal variance ($V = 1.0$, $\epsilon = 0.5$).

The results, shown in Figure 8, demonstrate that Laplacian noise yields higher estimation errors due to its heavier-tailed distribution, indicating stronger privacy protection at the cost of accuracy, thus validating the privacy-accuracy trade-off of our approach.

6.7. *Impact of different privacy budgets on parameter estimation*

To investigate the impact of different privacy budget parameters ε_0 and ε_1 on the system parameter estimation, we adjust ε_0 and ε_1 to change the encryption strengths of the input and output, which affects both the estimation error and the degree of privacy protection in the system, thus helping to find the optimal solution.

Using the chain rule, we compute the partial derivatives of ε_0 and ε_1, i.e., $\frac{\partial D_0}{\partial \varepsilon_0} = -\frac{2000}{\varepsilon_0^3}$ and $\frac{\partial D_1}{\partial \varepsilon_1} = -\frac{8\,\mathrm{Var}(s'_k)(2\gamma-1)e^{\varepsilon_1}}{n(e^{\varepsilon_1}+1)^4}$. Through numerical methods, we solve for the optimal solution, and under the estimation error constraint $\epsilon = 12$, we calculate $\epsilon_{\text{eff}}^2 = \frac{\epsilon^2}{k^2} - \sigma_d^2$, yielding $\epsilon_{\text{eff}}^2 \approx 4777$. By setting the

range for input encryption strength ε_0, combined with the optimization target and constraints, the optimal solution is found to be $\varepsilon_0 \approx 4.3$ and $\varepsilon_1 \approx 0.07$, with corresponding values of $D_0 \approx 4598.9$ and $D_1 \approx 51.06$.

We selected three different privacy budget parameter sets for comparison, as follows:

- Optimal solution: $\varepsilon_0 = 4.3$, $\varepsilon_1 = 0.07$.
- Scheme 1: $\varepsilon_0 = 4.28$, $\varepsilon_1 = 0.2$.
- Scheme 2: $\varepsilon_0 = 5.0$, $\varepsilon_1 = 0.06$.

For each parameter set, we calculated the corresponding D_0, D_1, and $D_0 + D_1$, and compared the parameter estimation errors.

The calculation results for the three parameter sets are summarized in the table below:

Table 1. Values of D_0, D_1, and $D_0 + D_1$ under different privacy budgets.

Privacy budget parameters	D_0	D_1	$D_0 + D_1$
Optimal solution ($\varepsilon_0 = 4.3$, $\varepsilon_1 = 0.07$)	4598.9	51.06	4649.96
Scheme 1 ($\varepsilon_0 = 4.28$, $\varepsilon_1 = 0.2$)	4643.3	6.22	4649.52
Scheme 2 ($\varepsilon_0 = 5.0$, $\varepsilon_1 = 0.06$)	3400	69.47	3469.47

From Table 1, we can see that the optimal solution set yields the best overall encryption strength $D_0 + D_1$. This also corroborates that the optimal solution reaches the extremum, and to get even closer to the extremum, higher precision would be required. Moreover, this optimal solution does not focus on the trade-off between input and output, and if there are more specific requirements, additional constraints should be included in the solution process. If higher privacy protection strength is needed, the ϵ limit can be relaxed to balance the encryption strengths for both input and output.

7. Conclusion remarks

This paper investigates the application of differential privacy encryption to reduce the risk of data tampering in FIR system identification under binary observation conditions. Two different differential privacy algorithms are proposed to ensure data security and privacy. The experimental evaluation confirms that the proposed method not only effectively protects sensitive information, but also maintains the accuracy of parameter estimation. These findings validate the effectiveness of the proposed scheme in this paper. Future work may explore optimizing the trade-off between privacy protection and estimation accuracy, as well as extending the approach to more complex system models and real-world applications.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The authors declare no conflicts of interest in this paper.

References

1. C. Dwork, A. Roth, *The Algorithmic Foundations of Differential Privacy*, Now Publishers Inc., 2014. http://dx.doi.org/10.1561/0400000042
2. Y. Kim, R. Eum, S. Park, Stealthy sensor attack detection and real-time performance recovery for resilient CPS, *IEEE Trans. Ind. Inf.*, **17** (2021), 7412–7422. http://dx.doi.org/10.1109/TII.2021.3052182
3. J. Guo, Q. Zhang, Y. Zhao, Identification of FIR Systems with binary-valued observations under replay attacks, *Automatica*, **172** (2025), 112001. https://doi.org/10.1016/j.automatica.2024.112001
4. Y. Jiang, S. Wu, H. Yang, H. Luo, Z. Chen, S. Yin, et al., Secure data transmission and trustworthiness judgement approaches against cyber-physical attacks in an integrated data-driven framework, *IEEE Trans. Syst. Man Cybern.: Syst.*, **52** (2022), 7799–7809. https://doi.org/10.1109/TSMC.2022.3164024
5. J. Glavaš, I. Uroda, B. Mandić, Managing digital transformation in public administration, in *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, (2021), 1466–1469. https://doi.org/10.23919/MIPRO52101.2021.9596775
6. H. Wang, L. Wang, Y. Yang, M. Hu, Z. Jia, Z. Chen, et al., A secure and efficient public data auditing solution for the cloud, in *2024 16th International Conference on Communication Software and Networks (ICCSN)*, (2024), 28–32. https://doi.org/10.1109/ICCSN63464.2024.10793330
7. S. Saratkar, A. Chaudhari, T. Thute, R. Raut, G. Thakre, H. Kumar, Assessment of heart-attack prediction using fuzzy rule based system, in *2024 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, (2024), 1–6. https://doi.org/10.1109/ICCUBEA61740.2024.10774808
8. D. B. Rawat, J. J. P. C. Rodrigues, I. Stojmenovic, *Cyber-Physical Systems: From Theory to Practice*, CRC Press, 2015.
9. L. Ljung, *System Identification: Theory for the User*, 2^{nd} edition, Prentice Hall, 1999.
10. M. Pouliquen, E. Pigeon, O. Gehan, A. Goudjil, Identification using binary measurements for IIR systems, *IEEE Trans. Autom. Control*, **65** (2020), 786–793. https://doi.org/10.1109/TAC.2019.2921657
11. J. Guo, J. F. Zhang, Y. Zhao, Adaptive tracking of a class of first-order systems with binary-valued observations and fixed thresholds, *J. Syst. Sci. Complexity*, **25** (2012), 1041–1051. https://doi.org/10.1007/s11424-012-1257-0
12. T. Wang, X. Zhang, J. Feng, X. Yang, A comprehensive survey on local differential privacy toward data statistics and analysis, *Sensors*, **20** (2020), 7030. https://doi.org/10.3390/s20247030
13. D. Ding, Q. Han, Z. Wang, X. Ge, A survey on model-based distributed control and filtering for industrial cyber-physical systems, *IEEE Trans. Ind. Inf.*, **15** (2019), 2483–2499. https://doi.org/10.1109/TII.2019.2905295
14. G. P. Liu, Networked learning predictive control of nonlinear cyber physical systems, *J. Syst. Sci. Complexity*, **33** (2020), 1719–1732. https://doi.org/10.1007/s11424-020-0243-1

15. M. S. Mahmoud, M. M. Hamdan, U. A. Baroudi, Modeling and control of cyber-physical systems subject to cyberattacks: A survey of recent advances and challenges, *Neurocomputing*, **338** (2019), 101–115. https://doi.org/10.1016/j.neucom.2019.01.099

16. R. Taheri, M. Shojafar, F. Arabikhan, A. Gegov, Unveiling vulnerabilities in deep learning-based malware detection: Differential privacy driven adversarial attacks, *Comput. Secur.*, **146** (2024), 104035. https://doi.org/10.1016/j.cose.2024.104035

17. S. Nabavirazavi, R. Taheri, S. S. Iyengar, Enhancing federated learning robustness through randomization and mixture, *Future Gener. Comput. Syst.*, **158** (2024), 28–43. https://doi.org/10.1016/j.future.2024.04.009

18. R. Taheri, F. Arabikhan, A. Gegov, N. Akbari, Robust aggregation function in federated learning, in *International Conference on Information and Knowledge Systems*, (2023), 168–175. https://doi.org/10.1007/978-3-031-51664-1_12

19. E. Nowroozi, I. Haider, R. Taheri, M. Conti, Federated learning under attack: Exposing vulnerabilities through data poisoning attacks in computer networks, *IEEE Trans. Netw. Serv. Manage.*, **22** (2025), 822–831. https://doi.org/10.1109/TNSM.2025.3525554

20. S. Nabavirazavi, R. Taheri, M. Shojafar, S. S. Iyengar, Impact of aggregation function randomization against model poisoning in federated learning, in *22nd IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom 2023)*, (2024), 165–172.

21. R. Tian, J. Mei, Privacy preserving resilient constrained consensus for multi-agent systems via state decomposition, in *2024 43rd Chinese Control Conference (CCC)*, (2024), 5806–5811. https://doi.org/10.23919/CCC63176.2024.10662823

22. J. Cortés, G. E. Dullerud, S. Han, J. Le Ny, S. Mitra, G. J. Pappas, Differential privacy in control and network systems, in *2016 IEEE 55th Conference on Decision and Control (CDC)*, (2016), 4252–4272. https://doi.org/10.1109/CDC.2016.7798915

23. C. Dwork, F. McSherry, K. Nissim, A. Smith, Calibrating noise to sensitivity in private data analysis, in *Journal of Privacy and Confidentiality*, (2006), 17–51. https://doi.org/10.1007/11681878_14

24. J. Guo, X. Wang, W. Xue, Y. Zhao, System identification with binary-valued observations under data tampering attacks, *IEEE Trans. Autom. Control*, **66** (2020), 1041–1055. https://doi.org/10.1109/TAC.2020.3029325

25. A. Teixeira, I. Shames, H. Sandberg, K. H. Johansson, A secure control framework for resource-limited adversaries, *Automatica*, **51** (2015), 135–148. https://doi.org/10.1016/j.automatica.2014.10.067

26. J. Guo, R. Jia, R. Su, Y. Zhao, Identification of FIR systems with binary-valued observations against data tampering attacks, *Trans. Syst. Man Cybern.: Syst.*, **53** (2023), 1041–1055. https://doi.org/10.1109/TSMC.2023.3276352

AIMS Press