

Article

Consistent Identification of Wiener Systems with Quantized Data Under Replay Attacks

Jingrong Liu ^{1,2}, Jingjie Gao ^{1,2} and Qingxiang Zhang ^{1,2,*}¹ School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China² Key Laboratory of Knowledge Automation for Industrial Processes, Ministry of Education, Beijing 100083, China

* Correspondence: qxzhang1001@xs.ustb.edu.cn

Abstract

This paper studies parameter estimation of Wiener systems with quantized inputs and multiple-valued observations in the presence of replay attacks. The proposed approach effectively addresses the destruction of the data temporal structure caused by replay attacks as well as the information loss induced by multiple-valued observations, thereby achieving consistent parameter estimation of Wiener systems under constrained observation conditions. First, the influence of replay attacks on the consistency of parameter estimation is theoretically analyzed. It is shown that the attack-free estimation algorithm converges to a biased limit under replay attacks, and an explicit expression of this limit is derived. To address this issue, two classes of consistent estimation algorithms are proposed. When the attack strategy is known, a compensation matrix is constructed to correct the compromised data, thereby eliminating the bias introduced by replay attacks and achieving consistent parameter estimation. When the attack strategy is unknown, a data preprocessing mechanism with structured markers is designed to estimate the attack strategy, based on which a compensated parameter estimation method is further developed. Moreover, the statistical properties of the proposed algorithms are analyzed, and the asymptotic normality of the estimators is established together with analytical expressions for the covariance matrices. On this basis, an optimization problem aimed at minimizing the estimation variance is formulated, yielding the optimal design parameters for the consistent estimation algorithms. Simulation results demonstrate that the proposed methods can effectively suppress the influence of replay attacks on system identification and achieve accurate and consistent parameter estimation for Wiener systems under multiple-valued observation.

Keywords: quantized system; multiple-valued observation; Wiener systems; replay attacks; consistent parameter estimation



Academic Editor: Francesco Bergadano

Received: 6 July 2026

Revised: 11 August 2026

Accepted: 12 August 2026

Published: 14 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

System identification is an important research area in control theory and signal processing, whose primary objective is to establish mathematical models of dynamic systems from input–output data [1–4]. In recent years, with the development of data-driven methodologies, system identification has gradually extended from traditional linear systems to complex nonlinear systems, while facing new challenges such as missing data, noise interference, and cyber attacks [5–8]. In particular, in networked environments, data integrity and security issues have become critical factors affecting identification performance [9,10].

Therefore, the development of robust and efficient system identification methods is of significant theoretical and practical importance.

Among various nonlinear models, Wiener systems have attracted extensive attention due to linear dynamics and static nonlinearity structure, which can effectively characterize many practical industrial processes, such as chemical reactions, communication systems, and biological systems [11–14]. In recent years, various approaches have been proposed for the parameter identification of Wiener systems. Zhang et al. [15] developed an identification method combining correlation analysis and intelligent optimization strategies for Wiener systems with input delays, where heuristic optimization algorithms were introduced to improve parameter estimation accuracy. Moreover, Wang et al. [16] incorporated deep learning architectures into Wiener system modeling and established a neural-network-based identification framework, thereby effectively characterizing complex nonlinear relationships and enhancing model representation capability. From a theoretical perspective, Bai and Reyland [17] systematically investigated the consistency and asymptotic properties of nonlinear system identification based on statistical learning theory and recursive estimation methods, providing an important theoretical foundation for parameter estimation of Wiener systems. Although considerable progress has been achieved, the intermediate variables in Wiener systems are generally unobservable, leading to strong nonconvexity and ill-posedness in parameter estimation. Consequently, the identification of Wiener systems remains highly challenging under complex observation conditions or attack environments.

In practical systems, limitations of sensors and communication bandwidth often require continuous outputs to be quantized, resulting in quantized observation problems. Quantization inevitably causes information loss, making traditional identification methods based on continuous observations difficult to apply directly [18–20]. To address this issue, Guo et al. [21] investigated parameter estimation for linear systems under binary observations and proposed a time-varying-threshold-based identification method, where the consistency of the estimator was established. Subsequently, related studies were gradually extended to multi-threshold quantization scenarios. Compared with binary measurements, multiple valued observations can provide richer observation information and thus alleviate information loss to some extent. However, multiple valued observations simultaneously introduces more complicated nonlinear observation mechanisms, leading to stronger nonconvexity and increased analytical difficulty in the identification problem. To tackle this issue, recent studies proposed parameter identification methods based on quasi-Newton algorithms and established corresponding asymptotic theoretical analysis frameworks [22]. Meanwhile, from a statistical perspective, quantizers, as a class of nonlinear sensors, alter the probability distribution characteristics of observation data and thereby directly affect the statistical performance of parameter estimation. For example, Forgione et al. [23] showed that quantization changes the covariance structure of estimation errors and influences the convergence rate.

Furthermore, with the development of networked control systems, the influence of cyber attacks on system identification has received increasing attention [24]. Among various attack types, replay attacks constitute a typical cyber attack strategy, where historical data are maliciously reused to replace current data, thereby destroying the temporal structure and statistical properties of the data and degrading the performance of identification algorithms. Zaman et al. [25] defined the replay attack mechanism and analyzed its impact on control systems.

Table 1 summarizes the most recent works on system identification under attacks, which can be categorized along three dimensions: the system model, the observation type, and the defense mechanism.

Table 1. Comparison of recent works on system identification under attacks.

Reference	System	Input Excitation	Observation	Attack	Defense
[5]	FIR	Periodic input	Quantized I/O	Replay	Stochastic defense
[8]	FIR	Periodic input	Multi-level	Tampering	Joint estimation
[26]	FIR	Periodic input	Binary	Replay	Compensation
[27]	FIR	Periodic input	Binary	Seq. replay	Compensation
This paper	Wiener	Quantized input	Multi-threshold	Replay	Structured Marking Compensation

As summarized in Table 1, existing studies on identification under cyber attacks mainly focus on linear FIR systems [5,8,26,27]. Compared with these works, this paper considers nonlinear Wiener systems, where the nonlinear function is located after the linear dynamics, making the identifiability analysis more challenging due to the unobservable intermediate state. Regarding observations, existing methods mainly consider binary-valued sensors [26,27] or single-threshold quantization [5,8], while this paper adopts multi-threshold quantization to obtain richer information and develops a corresponding weighted fusion strategy. Regarding defense mechanisms, existing works either assume known attack strategies [8,26,27] or rely on attack detection-based methods [5]. In contrast, this paper proposes a unified compensation-based framework that handles both known and unknown replay attack strategies. For the unknown case, a structured-marking mechanism is introduced to jointly estimate attack and system parameters. It can be seen from Table 1 that existing works do not simultaneously address nonlinear Wiener systems, multi-threshold quantized observations, and unknown replay attacks, which motivates the proposed identification framework.

Motivated by the complexity of system identification problems under the combined effects of nonlinear structures, quantized observations, and cyber attacks, this paper focuses on Wiener systems with quantized inputs and multiple-valued observations, and investigates the consistent parameter identification problem under replay attack environments. Since replay attacks destroy the temporal structure of observation data and alter their statistical properties, conventional estimation algorithms developed under attack-free assumptions generally fail in such scenarios. To address this issue, this paper systematically studies the problem from the perspectives of attack modeling and estimation compensation. The main contributions of this paper are summarized as follows:

- An estimation algorithm without replay attacks is developed for parameter identification of Wiener systems with multi-valued observations. The impact of multi-threshold quantization on the observation information structure is investigated, and its effects on parameter identifiability and estimation accuracy are systematically characterized.
- For nonlinear Wiener systems under the coexistence of quantized observations and replay attacks, explicit asymptotic bias expressions are derived for the attack-free estimator (Theorem 2), and consistent parameter estimation methods based on a compensation matrix are proposed for both known and unknown attack strategies (Theorems 3–4). The asymptotic normality of the estimators is rigorously established (Theorem 5).
- By exploiting the asymptotic covariance expressions, an optimal design parameter minimizing the estimation variance is obtained (Theorem 6), which provides a principled approach to trading off the marking overhead against the estimation accuracy.

The remainder of this paper is organized as follows. Section 2 formulates the identification problem. Section 3 presents the parameter estimation method and consistency analysis in the attack-free case. Section 4 develops consistent estimation algorithms for both known and unknown attack strategies and provides the corresponding theoretical proofs. Section 5 analyzes the asymptotic statistical properties of the proposed algorithms and

investigates the optimal consistent estimation strategy. Section 6 verifies the effectiveness of the proposed methods through numerical simulations. Finally, Section 7 concludes the paper and discusses future research directions.

2. Statement of the Problem

The following discrete-time Wiener system with single input and single output is studied:

$$\begin{cases} x_k = u_k\theta_1 + u_{k-1}\theta_2 + \cdots + u_{k-n_1+1}\theta_{n_1} \\ y_k = H(x_k, \eta) + d_k \end{cases} \quad (1)$$

where $\theta = [\theta_1, \theta_2, \dots, \theta_{n_1}]^T \in \mathbb{R}^{n_1}$ and $\eta = [\eta_1, \eta_2, \dots, \eta_{n_2}]^T \in \mathbb{R}^{n_2}$ denote the unknown parameter vectors associated with the linear and nonlinear subsystems, respectively. The nonlinear mapping $H(\cdot, \eta)$ is a static parameterized function defined on $\mathcal{D}_H \subseteq \mathbb{R}$. For $k \in \{1, 2, \dots, N\}$, u_k , x_k , and y_k represent the system input, internal variable, and system output, respectively, while d_k denotes additive measurement noise and N is the data length. The input u_k is quantized, taking values from a finite set $\{\mu_1, \dots, \mu_r\}$. Let μ_k denote the original input signal before quantization. After the quantization process, define the regression vector as $\Pi_k = [u_k, u_{k-1}, \dots, u_{k-n_1+1}]^T$, which yields $x_k = \Pi_k^T \theta$. Here, $\cdot^T$ denotes the transpose operation for vectors or matrices.

The output y_k passes through a multi-threshold quantization sensor with thresholds designed as $-\infty < C_0 < C_1 < \cdots < C_v < \infty$. After quantization, the multi-valued observation v_k^0 is obtained, which is defined by

$$v_k^0 = \sum_{i=1}^v i I_{\{C_{i-1} < y_k \leq C_i\}} \quad (2)$$

where

$$I_{\{C_{i-1} < y_k \leq C_i\}} = \begin{cases} 1, & C_{i-1} < y_k \leq C_i \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

The data received by the estimation center are denoted by v_k , which can be represented by the vector $\bar{v}_k = [v_k|_1, \dots, v_k|_v]^T$, where $v_k|i = I_{\{v_k \leq i\}}$, $i = 0, \dots, v$.

Figure 1 shows that the output y_k passes through a multi-threshold quantizer to produce v_k^0 , which is transmitted over a communication network subject to replay attacks. The received signal v_k is used by the estimation center to identify the unknown parameters θ and η . The noise d_k is additive. Let $\epsilon \in (0, 1]$ denote the probability that an attack occurs, and let τ be a fixed integer-valued delay introduced by the attacker. Define the binary variable $\kappa_k \in \{0, 1\}$ to characterize the attack status at time instant k , where $\kappa_k = 1$ corresponds to the presence of a replay attack and $\kappa_k = 0$ means that the transmission is attack-free. Under this setting, the signal received by the estimation center can be expressed as

$$v_k = (1 - \kappa_k)v_k^0 + \kappa_k v_{k-\tau}^0 \quad (4)$$

The replay attack described by (4) can be characterized by a tuple (ϵ, τ) , which represents the attack strategy.

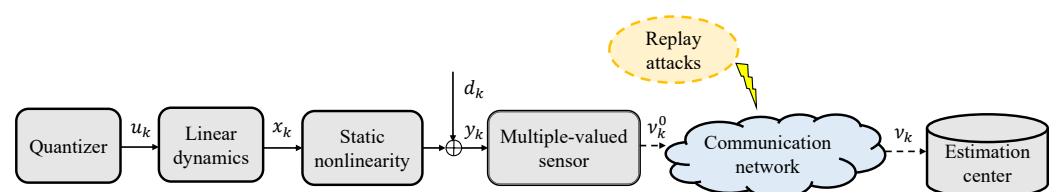


Figure 1. System structure.

To establish the consistency of the proposed identification algorithm under replay attacks, several assumptions are required regarding the attack process and measurement noise. These assumptions are introduced to characterize the uncertainty caused by cyber attacks and stochastic disturbances while ensuring that the collected data contain sufficient information for parameter recovery. First, since replay attacks may corrupt the transmitted observations and introduce statistical bias into the identification process, a constraint on the attack occurrence probability is necessary to guarantee that the impact of attacks remains limited. Specifically, the attack probability is assumed to be bounded away from one, which ensures that a sufficient amount of attack-free information is preserved for consistent estimation. Second, the statistical property of measurement noise directly affects the probability distribution of the quantized observations and the construction of the proposed estimator.

Assumption 1. *The attack process is energy-limited, and the attack occurrence probability satisfies $\Pr(\kappa_k = 1) = \epsilon \leq \bar{\epsilon} < 1$ for all $k \in \{1, \dots, N\}$, where $\bar{\epsilon} \in (0, 1)$ is a known upper bound on the attack probability. This assumption ensures that a non-vanishing fraction of the transmitted data remains attack-free at every time instant, which is essential for achieving consistent parameter estimation.*

Assumption 2. *The noise sequence d_k is assumed to be Gaussian and independently identically distributed (i.i.d.) with mean zero and known variance $\sigma^2 > 0$. The Gaussian assumption is standard in the identification literature and enables the construction of the estimator through the probit function $\Psi^{-1}(\cdot)$. Furthermore, $\Psi(\cdot)$ and $\phi(\cdot)$ denote the cumulative distribution function and probability density function associated with d_k , respectively.*

Given the quantized input sequence $\{u_k\}_{k=1}^N$ and the multi-valued observations $\{v_k\}_{k=1}^N$ received under potential replay attacks, the objective is to design consistent estimators θ_N and $\hat{\eta}_N$ such that $\hat{\theta}_N \rightarrow \theta$ and $\hat{\eta}_N \rightarrow \eta$ almost surely as $N \rightarrow \infty$.

3. Preliminaries

3.1. Parameter Estimation Algorithm in the Attack-Free Case

Before proceeding to the detailed mathematical derivation, we outline the logic of the attack-free estimation algorithm to help the reader navigate the notation. The algorithm consists of three steps. Step 1: For each input pattern U_i and each quantization threshold C_j , the empirical frequency of $\{v_k^0 \leq j\}$ among all data points with input pattern U_i is used to estimate $\omega_i = H(U_i\theta, \eta)$ via the inverse noise CDF $\Psi^{-1}(\cdot)$, yielding $\hat{\omega}_{N,i}^{(j)}$ (Equations (9) and (10)). Step 2: The estimates from all $v + 1$ thresholds are combined through a weighted average with weights λ_j to obtain $\hat{\omega}_N$ (Equation (11)). Step 3: The system parameters θ and η are recovered from $\hat{\omega}_N$ by solving the nonlinear equations via the mapping $q(\cdot)$ (Equations (12)–(14)). This three-step structure—frequency estimation, threshold fusion, and inverse mapping—carries over to the attack case with the key modification that the empirical frequencies must first be corrected by the compensation matrix (Section 4).

Assume that $\{u_k\}$ is an arbitrary input sequence taking values from the quantization set $\mathcal{U} = \{\mu_1, \dots, \mu_r\}$. The input sequence $\{u_k\}$ gives rise to the regression vectors Π_k^T , which admits $\mathcal{L} = r^{n_1}$ possible patterns, denoted by $U = \{U_1, \dots, U_{\mathcal{L}}\}$. Specifically,

$$\begin{cases} U_1 = [\mu_1, \mu_1, \dots, \mu_1, \mu_1]_{1 \times n_1} \\ U_2 = [\mu_1, \mu_1, \dots, \mu_1, \mu_2]_{1 \times n_1} \\ \vdots \\ U_{\mathcal{L}} = [\mu_r, \mu_r, \dots, \mu_r, \mu_r]_{1 \times n_1} \end{cases}$$

For a given input sequence $\{u_k\}$ and its corresponding regression sequence $\{\Pi_1^T, \dots, \Pi_N^T\}$, define $N_j = \sum_{k=1}^N \mathbf{I}_{\{\Pi_k^T = U_j\}}$, $j \in \mathbb{L} = \{1, \dots, \mathcal{L}\}$, which represents the number of occurrences of the pattern U_j .

Assumption 3. For a given data length N , define $\beta_j = \lim_{N \rightarrow \infty} \frac{\sum_{k=1}^N \mathbf{I}_{\{\Pi_k^T = U_j\}}}{N}$, $j = 1, \dots, \mathcal{L}$. Whenever the above limit exists and β_j is nonzero, the input pattern Π_j^T is regarded as persistently exciting. For notational simplicity, let $\mathbb{L}_0 = \{1, \dots, \mathcal{L}_0\}$ denote the index set associated with all persistently exciting input patterns. $\mathcal{L}_0 \geq n = n_1 + n_2$, thereby guaranteeing the identifiability of the system.

Consider the nonlinear function of the Wiener system in the following form:

$$\begin{cases} H(x_1, \eta) = \omega_1 \\ \vdots \\ H(x_{\mathcal{L}_0}, \eta) = \omega_{\mathcal{L}_0} \end{cases} \quad (5)$$

where $\omega = [\omega_1, \dots, \omega_{\mathcal{L}_0}]^T$.

Define the full-column-rank matrix $\bar{\Omega} = [\Pi_1, \dots, \Pi_{\mathcal{L}_0}]^T$, which guarantees the identifiability of the system. Choose n_1 vectors from Π_j^T , $j = 1, \dots, \mathcal{L}_0$, to construct a basis set. For convenience, let the selected vectors be $\Pi_1^T, \dots, \Pi_{n_1}^T$, which constitute the input matrix

$$\Omega = [\Pi_1, \dots, \Pi_{n_1}]^T \quad (6)$$

Each pattern Π_j^T , $j = 1, \dots, \mathcal{L}_0$, can be linearly represented by Π_i^T , $i = 1, \dots, n_1$, namely, $\Pi_j^T = \sum_{i=1}^{n_1} \gamma_{j,i} \Pi_i^T$, $j = 1, \dots, \mathcal{L}_0$, where $\gamma_{j,i}$ denotes the corresponding coefficients. Consequently, one has

$$x_j = \sum_{i=1}^{n_1} \gamma_{j,i} x_i, \quad j = 1, \dots, \mathcal{L}_0 \quad (7)$$

According to (5) and (7), it follows that

$$\mathbb{H}(x_1, \dots, x_{n_1}, \eta) = \omega \quad (8)$$

Equation (8) contains n unknown variables and n equations, and therefore admits a unique solution. In this case, system (1) is identifiable. Accordingly, the following assumption is introduced.

Assumption 4. Suppose that there exists a compact set $\mathcal{A} \subseteq \mathbb{R}^{\mathcal{L}_0}$ including the true output ω . For every $\alpha = [\alpha_1, \dots, \alpha_{\mathcal{L}_0}]^T \in \mathcal{A}$, the nonlinear equation $\mathbb{H}(x_1, \dots, x_{n_1}, \eta) = \alpha$ has a unique solution $q(\alpha) = [x_1, \dots, x_{n_1}, \eta]^T \in \mathbb{R}^n$. In addition, the mapping $q(\alpha)$ is continuous and bounded on $\mathcal{A}$.

In the absence of replay attacks, the bounded continuity of $q(\cdot)$ allows probabilities to be approximated through empirical frequencies. Then, by using two inverse mappings, the parameter estimates of the Wiener system with quantized inputs and multi-valued observations can be derived. The corresponding estimation procedure is given as follows.

Theorem 1. Consider system (1) and the multi-valued observations (2)–(3). Under Assumptions 1–4, as $N \rightarrow \infty$, the estimates $\hat{\theta}_N$ and $\hat{\eta}_N$ obtained from Algorithms (9)–(14) converge consistently to the true values θ and η , respectively.

- Estimation of ω

$$\hat{\omega}_N^{(j)} = [\hat{\omega}_{N,1}^{(j)}, \dots, \hat{\omega}_{N,\mathcal{L}_0}^{(j)}]^T \quad (9)$$

$$\hat{\omega}_{N,i}^{(j)} = C_j - \Psi^{-1} \left(\frac{1}{N_i} \sum_{k=1}^N \nu_k^0 |_j \mathbf{I}_{\{\Pi_k^T = U_i\}} \right) \quad (10)$$

where $i = 1, \dots, \mathcal{L}_0$ and $j = 0, \dots, v$.

- Weighted estimation of $\hat{\omega}_N$

$$\hat{\omega}_N = \sum_{j=0}^v \lambda_j \hat{\omega}_N^{(j)} \quad (11)$$

where λ_j denotes the weight associated with threshold C_j , satisfying $\lambda_0, \dots, \lambda_v \geq 0$ and $\lambda_0 + \dots + \lambda_v = 1$.

- Estimation of θ and η

After obtaining $\hat{\omega}_N$, the estimates of θ and η are given by

$$\hat{\theta}_N = \Omega^{-1} \hat{W}_N \quad (12)$$

and

$$\hat{\eta}_N = [\varrho_{n_1+1}(\hat{\omega}_N), \dots, \varrho_n(\hat{\omega}_N)]^T \quad (13)$$

where

$$\hat{W}_N = [\varrho_1(\hat{\omega}_N), \dots, \varrho_{n_1}(\hat{\omega}_N)]^T \quad (14)$$

Proof. In the absence of replay attacks, the probability that the estimation center receives data satisfying $\nu_k \leq j$ is given by

$$\begin{aligned} \mathbb{E}[\nu_k] &= \mathbb{E}[\nu_k^0] \\ &= \Pr\{\nu_k^0 \leq j\} \\ &= \Pr\{y_k \leq C_j\} \\ &= \Pr\{d_k \leq C_j - \omega_k\} \\ &= \Psi(C_j - \omega_k) \end{aligned} \quad (15)$$

Since $\Pi_k^T \in \{U_1, \dots, U_{\mathcal{L}_0}\}$, one has

$$\mathbb{E}[\nu_k^0] \in \{\mathbb{E}[\nu_k^0 |_j] \mathbf{I}_{\{\Pi_k^T = U_1\}}, \dots, \mathbb{E}[\nu_k^0 |_j] \mathbf{I}_{\{\Pi_k^T = U_{\mathcal{L}_0}\}}\} \quad (16)$$

According to the law of large numbers, as $N \rightarrow \infty$,

$$\frac{1}{N_i} \sum_{k=1}^N \nu_k^0 |_j \mathbf{I}_{\{\Pi_k^T = U_i\}} \rightarrow \Psi(C_j - \omega_i) \quad (17)$$

Substituting (17) into (10) yields

$$\hat{\omega}_{N,i}^{(j)} \rightarrow C_j - \Psi^{-1}(\Psi(C_j - \omega_i)) = \omega_i \quad (18)$$

Since each component converges to the true value of the nonlinear function, it follows directly that

$$\hat{\omega}_N \rightarrow \omega \quad (19)$$

Consequently,

$$\widehat{W}_N = \begin{bmatrix} q_1(\widehat{\omega}_N) \\ \vdots \\ q_{n_1}(\widehat{\omega}_N) \end{bmatrix} \rightarrow \begin{bmatrix} q_1(\omega) \\ \vdots \\ q_{n_1}(\omega) \end{bmatrix} = W \quad (20)$$

$$\widehat{\eta}_N = \begin{bmatrix} q_{n_1+1}(\widehat{\omega}_N) \\ \vdots \\ q_n(\widehat{\omega}_N) \end{bmatrix} \rightarrow \begin{bmatrix} q_{n_1+1}(\omega) \\ \vdots \\ q_n(\omega) \end{bmatrix} = \eta \quad (21)$$

According to (20),

$$\widehat{\theta}_N = \Omega^{-1} \widehat{W}_N \rightarrow \Omega^{-1} W = \Omega^{-1} \Omega \theta = \theta, \quad N \rightarrow \infty \quad (22)$$

The proof is completed. $\square$

3.2. Impact of Replay Attacks on the Attack-Free Estimation Algorithm

The estimation algorithm designed for the attack-free case may exhibit different convergence properties when replay attacks occur. Therefore, a theoretical analysis is conducted to reveal how replay attacks affect the convergence behavior of the algorithm. The analysis mainly focuses on the following two issues: whether replay attacks invalidate the original convergence results, and whether the algorithm still converges when consistency is destroyed, together with the theoretical expression of its limiting value. This subsection provides a detailed analysis of these problems.

Due to the temporal dependence introduced by replay attacks, the conditional probability structure among regression patterns plays a critical role in the performance analysis of the estimation algorithm. To this end, define an $\mathcal{L}_0 \times \mathcal{L}_0$ probability matrix as

$$\mathbb{R}(\tau) = \begin{bmatrix} \rho_{1,1}(\tau) & \rho_{1,2}(\tau) & \cdots & \rho_{1,\mathcal{L}_0}(\tau) \\ \rho_{2,1}(\tau) & \rho_{2,2}(\tau) & \cdots & \rho_{2,\mathcal{L}_0}(\tau) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{\mathcal{L}_0,1}(\tau) & \rho_{\mathcal{L}_0,2}(\tau) & \cdots & \rho_{\mathcal{L}_0,\mathcal{L}_0}(\tau) \end{bmatrix} \quad (23)$$

where $\rho_{i,j}(\tau) = \Pr\{\Pi_{k-\tau}^T = U_i \mid \Pi_k^T = U_j\}$ denotes the conditional probability that the regression pattern at time instant $k - \tau$ is U_i given that the regression pattern at time instant k is U_j . Each column of $\mathbb{R}(\tau)$ forms a probability distribution, namely, $\sum_{i=1}^{\mathcal{L}_0} \rho_{i,j}(\tau) = 1$. Furthermore, when $\tau = 0$, one has $\rho_{i,j}(0) = 1$ for $i = j$ and $\rho_{i,j}(0) = 0$ for $i \neq j$. Consequently, $\mathbb{R}(0) = \mathbb{I}_{\mathcal{L}_0}$, where $\mathbb{I}$ denotes the identity matrix.

Remark 1. In practice, $\mathbb{R}(\tau)$ can be estimated from the input sequence via empirical transition counts:

$$\widehat{\rho}_{i,j}(\tau) = \frac{\sum_{k=\tau+1}^N \mathbf{I}_{\{\Pi_{k-\tau}^T = U_i\}} \mathbf{I}_{\{\Pi_k^T = U_j\}}}{\sum_{k=\tau+1}^N \mathbf{I}_{\{\Pi_k^T = U_j\}}}, \quad i, j \in \{1, \dots, \mathcal{L}_0\}. \quad (24)$$

The above estimation can be implemented by counting the transitions between input patterns during a single scan of the data, requiring $O(N\mathcal{L}_0)$ operations. For the system considered in this paper, $\mathcal{L}_0 = r^{n_1} = 4$ with $r = 2$ and $n_1 = 2$, and thus $\mathbb{R}(\tau)$ can be directly constructed with negligible computational cost.

For high-dimensional systems, the main scalability issue arises from the exponential increase of the number of possible input patterns, i.e., $\mathcal{L}_0 = r^{n_1}$. For example, when $r = 2$ and $n_1 = 20$, $\mathcal{L}_0 = 2^{20} = 1,048,576$, and the full transition matrix contains more than 10^{12} elements. However,

in practical applications, only a subset of input patterns is usually activated by a finite-length input sequence. Let $\mathcal{S}$ denote the set of observed patterns and $\mathcal{L}_s = |\mathcal{S}|$ be its cardinality. Typically, $\mathcal{L}_s \ll \mathcal{L}_0$, and only the sparse transition matrix $\hat{\mathbb{R}}_s(\tau) \in \mathbb{R}^{\mathcal{L}_s \times \mathcal{L}_s}$ needs to be constructed. Consequently, the storage requirement is reduced from $O(\mathcal{L}_0^2)$ to $O(\mathcal{L}_s^2)$. This sparse implementation preserves the dominant transition information while significantly improving the scalability of the proposed compensation framework.

The attack signals contaminate the observation data and thereby introduce non-negligible estimation errors. To characterize such performance degradation, it is necessary to analyze the asymptotic properties of the estimator under attack environments.

Theorem 2. Consider the Wiener system (1) with multi-valued observations (2) and (3) under the (ϵ, τ) replay attack described by (4). Under Assumption 2, as the data length $N \rightarrow \infty$, the estimators $\hat{W}_N$, $\hat{\eta}_N$, and $\hat{\theta}_N$ converge almost surely to the following theoretical values, respectively:

$$\check{\omega} = [\varrho_1(\check{\omega}), \dots, \varrho_{n_1}(\check{\omega})]^T \quad (25)$$

$$\check{\eta} = [\varrho_{n_1+1}(\check{\omega}), \dots, \varrho_n(\check{\omega})]^T \quad (26)$$

$$\check{\theta} = \Omega^{-1}\check{\omega} \quad (27)$$

where $\check{\omega}$ is defined in (31).

Proof. In the presence of replay attacks, according to (1)–(4) and Assumption 2,

$$\begin{aligned} & \Pr\{v_k \leq a\} \\ &= \Pr\{v_k \leq a, \kappa_k = 0\} + \Pr\{v_k \leq a, \kappa_k = 1\} \\ &= \Pr\{\kappa_k = 0\} \Pr\{v_k^0 \leq a\} \\ &+ \Pr\{\kappa_k = 1\} \Pr\{v_{k-\tau}^0 \leq a\} \\ &= (1 - \epsilon) \Psi(C_a - \omega_k) + \epsilon \Psi(C_a - \omega_{k-\tau}) \\ &= (1 - \epsilon) \Psi(C_a - \omega_k) + \epsilon \sum_{m=1}^{\mathcal{L}_0} \rho_{m,k}(\tau) \Psi(C_a - \omega_m) \end{aligned} \quad (28)$$

According to the law of large numbers, as $N \rightarrow \infty$, one further has

$$\frac{1}{N_i} \sum_{k=1}^N v_k |_a \mathbf{I}_{\{\Pi_k^T = U_i\}} \rightarrow (1 - \epsilon) \Psi(C_a - \omega_i) + \epsilon \sum_{m=1}^{\mathcal{L}_0} \rho_{m,i}(\tau) \Psi(C_a - \omega_m) = \varepsilon_i^{(a)} \quad (29)$$

Substituting (29) into (9) and (10) yields

$$\hat{\omega}_N^{(a)} \rightarrow \check{\omega}^{(a)} = \begin{bmatrix} C_a - \Psi^{-1}(\varepsilon_1^{(a)}) \\ \vdots \\ C_a - \Psi^{-1}(\varepsilon_{\mathcal{L}_0}^{(a)}) \end{bmatrix} \quad (30)$$

Substituting (30) into (11) of the original estimation algorithm gives

$$\hat{\omega}_N \rightarrow \check{\omega} = \sum_{a=0}^v \lambda_a \check{\omega}^{(a)} \quad (31)$$

Finally, substituting (31) into (12)–(14) leads to (25)–(27). $\square$

4. Design of Consistent Estimation Algorithms

The above results show that replay attacks invalidate the convergence of the original estimation algorithm. To ensure accurate parameter estimation, consistent estimation methods are further developed for two scenarios: known attack strategies and unknown attack strategies.

4.1. Design of Consistent Estimation Algorithms: Known Attack Strategy

Define the compensation matrix $\aleph(\epsilon, \tau)$ as

$$\aleph(\epsilon, \tau) = \begin{bmatrix} 1 - \epsilon + \epsilon\rho_{1,1}(\tau) & \epsilon\rho_{2,1}(\tau) & \cdots & \epsilon\rho_{\mathcal{L}_0,1}(\tau) \\ \epsilon\rho_{1,2}(\tau) & 1 - \epsilon + \epsilon\rho_{2,2}(\tau) & \cdots & \epsilon\rho_{\mathcal{L}_0,2}(\tau) \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon\rho_{1,\mathcal{L}_0}(\tau) & \epsilon\rho_{2,\mathcal{L}_0}(\tau) & \cdots & 1 - \epsilon + \epsilon\rho_{\mathcal{L}_0,\mathcal{L}_0}(\tau) \end{bmatrix} \quad (32)$$

When the compensation matrix $\aleph(\epsilon, \tau)$ is invertible, denote its inverse matrix by $\tilde{\aleph}(\epsilon, \tau)$, which takes the following form:

$$\tilde{\aleph}(\epsilon, \tau) = \begin{bmatrix} q_{11}(\epsilon, \tau) & q_{12}(\epsilon, \tau) & \cdots & q_{1\mathcal{L}_0}(\epsilon, \tau) \\ q_{21}(\epsilon, \tau) & q_{22}(\epsilon, \tau) & \cdots & q_{2\mathcal{L}_0}(\epsilon, \tau) \\ \vdots & \vdots & \ddots & \vdots \\ q_{\mathcal{L}_0 1}(\epsilon, \tau) & q_{\mathcal{L}_0 2}(\epsilon, \tau) & \cdots & q_{\mathcal{L}_0 \mathcal{L}_0}(\epsilon, \tau) \end{bmatrix} \quad (33)$$

Remark 2. The compensation matrix $\aleph(\epsilon, \tau) = (1 - \epsilon)\mathbb{I}_{\mathcal{L}_0} + \epsilon\mathbb{R}(\tau)$ has a circulant structure due to the periodicity of the input sequence. Let $\tau_0 = \text{mod}(\tau - 1, \mathcal{L}_0) + 1$. According to the theory of circulant matrices, $\aleph(\epsilon, \tau)$ is invertible if and only if

$$1 - \epsilon + \epsilon e^{\frac{j2\pi\delta}{\mathcal{L}_0}(\tau_0 - 1)} \neq 0, \quad \delta = 0, \dots, \mathcal{L}_0 - 1. \quad (34)$$

The above condition indicates that singularity may occur only for the special case $\epsilon = 1/2$ together with specific replay delays. Under the simulation settings adopted in this paper (e.g., $\epsilon = 0.3$, $\mathcal{L}_0 = 4$), this singular case is avoided, and hence the compensated estimator is well-defined.

For $i = 1, \dots, \mathcal{L}_0$, the compensated estimator of $\omega_i^{(a)}$ is given by

$$\hat{\omega}_{N,i}^{(a)} = C_a - \Psi^{-1} \left(\sum_{j=1}^{\mathcal{L}_0} \left[q_{ij}(\epsilon, \tau) \frac{1}{N_j} \sum_{k=1}^N v_k |_a \mathbf{I}_{\{\Pi_k^T = U_j\}} \right] \right) \quad (35)$$

Combining (35) with (9) and (11), the weighted estimate $\hat{\omega}_{N,i} = \sum_{a=0}^v \lambda_j \hat{\omega}_{N,i}^{(a)}$ can be obtained. Substituting $\hat{\omega}_{N,i}$ into (12)–(14) yields the compensated estimates of the system parameters $\hat{\theta}_N$ and $\hat{\eta}_N$.

Theorem 3. As $N \rightarrow \infty$, under the known attack strategy (ϵ, τ) , the compensated estimate obtained from (9), (11) and (35) satisfies $\hat{\omega}_N \rightarrow \omega$. Furthermore, substituting it into (12)–(14) yields $\hat{\theta}_N \rightarrow \theta$ and $\hat{\eta}_N \rightarrow \eta$.

Proof. According to (29), the following relation can be obtained:

$$\begin{bmatrix} \frac{1}{N_1} \sum_{k=1}^N v_k |_a \mathbf{I}_{\{\Pi_k^T = U_1\}} \\ \vdots \\ \frac{1}{N_{\mathcal{L}_0}} \sum_{k=1}^N v_k |_a \mathbf{I}_{\{\Pi_k^T = U_{\mathcal{L}_0}\}} \end{bmatrix} \rightarrow \aleph(\epsilon, \tau) \begin{bmatrix} \Psi(C_a - \omega_1) \\ \vdots \\ \Psi(C_a - \omega_{\mathcal{L}_0}) \end{bmatrix} \quad (36)$$

Rearranging the above equation gives

$$\tilde{\mathbf{N}}(\epsilon, \tau) \begin{bmatrix} \frac{1}{N_1} \sum_{k=1}^N v_k |_{a \in \{\Pi_k^T = U_1\}} \\ \vdots \\ \frac{1}{N_{L_0}} \sum_{k=1}^N v_k |_{a \in \{\Pi_k^T = U_{L_0}\}} \end{bmatrix} \rightarrow \begin{bmatrix} \Psi(C_a - \omega_1) \\ \vdots \\ \Psi(C_a - \omega_{L_0}) \end{bmatrix} \quad (37)$$

Substituting (37) into (35) leads to $\hat{\omega}_{N,i}^{(a)} \rightarrow \omega_i$. Similarly, each component converges to the true value of the nonlinear function, and hence $\hat{\omega} \rightarrow \omega$. According to (20)–(22), one obtains $\hat{\theta}_N \rightarrow \theta$ and $\hat{\eta}_N \rightarrow \eta$. $\square$

4.2. Design of Consistent Estimation Algorithms: Unknown Attack Strategy

However, in practical scenarios, the estimation center generally cannot directly obtain the attack strategy. Therefore, a data preprocessing module is introduced to first estimate the attack strategy $(\hat{\epsilon}_N, \hat{\tau}_N)$, after which the identification algorithm for known attack strategies is employed to further estimate the system parameters.

As shown in Figure 2, the multi-valued observation data v_k^0 are not directly transmitted through the communication network. Instead, they are first processed by a signal preprocessing module to generate $\hat{v}_k^0$. The processed signal $\hat{v}_k^0$ may subsequently undergo replay attacks during transmission to the estimation center. In this case, the replay attack is launched on the preprocessed data $\hat{v}_k^0$, and the data received by the estimation center satisfy.

$$v_k = (1 - \kappa_k) \hat{v}_k^0 + \kappa_k \hat{v}_{k-\tau}^0 \quad (38)$$

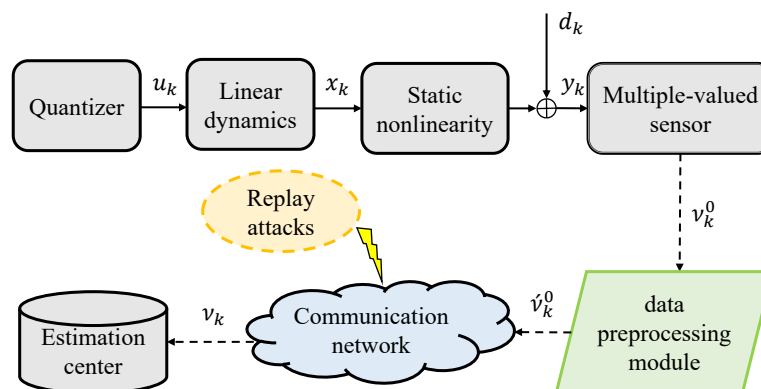


Figure 2. System structure with a data preprocessing module.

Next, the construction procedure of the preprocessed data is presented in three steps. Let $\bar{t} = \bar{\tau} + 2$. In the first step, select an index sequence $\{\zeta_i\}_{i=1}^{g_N}$ satisfying $\zeta_{i+1} - \zeta_i > 3\bar{t}$, and set $\zeta_{g_N+1} = N$. This condition guarantees that different marked intervals do not overlap under replay attacks. In the second step, construct the marked intervals. Define each marked interval as $\mathbb{M}_i = \{\zeta_i, \dots, \zeta_i + 3\bar{t} - 1\}$. The overall marked set is defined as $\mathbb{Z}_N^{(0)} = \bigcup_{i=1}^{g_N} \mathbb{M}_i$. In the third step, embed structured markers. Specifically, the following structure is constructed within each interval to obtain the preprocessed sequence $\hat{v}_k^0$:

$$\hat{v}_k^0 = \begin{cases} v_k^0, & k \notin \mathbb{Z}_N^{(0)}, \\ r, & k = \zeta_i + r, \\ r, & k = \zeta_i + \bar{t} + r \end{cases} \quad (39)$$

where $i = 1, 2, \dots, g_N$ and $r = 0, 1, \dots, \bar{t} - 1$.

As illustrated in Figure 3, the preprocessed data are constructed with total length $N_0 = 3g_N\bar{t}$. First, define $\tilde{N}_i = \sum_{i=1}^{N-N_0} I_{\{\Pi_k^T = U_i, k \notin Z_N^{(0)}\}}$. Next, define the function $\Im(\bar{t})(\{\chi_0, \chi_1, \dots, \chi_{\bar{t}-1}\})$, which returns the position index of the first occurrence of $\bar{t} - 1$ in the sequence $\{\chi_0, \chi_1, \dots, \chi_{\bar{t}-1}\}$. If the sequence does not contain $\bar{t} - 1$, the function returns 0, where $\chi \in \{0, 1, \dots, \bar{t} - 1\}$. The consistent estimation procedure of the system parameters is then carried out in two steps.

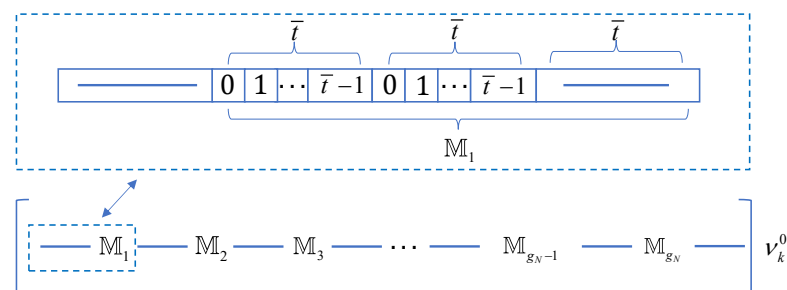


Figure 3. Illustration of data preprocessing.

In the first step, estimate the attack strategy $(\hat{\epsilon}_N, \hat{\tau}_N)$:

$$\hat{\epsilon}_N = \frac{1}{2g_N} \sum_{i=1}^{2g_N} I_{\{v_k \neq \bar{t}-1, v_k \in \mathbb{V}\}} \quad (40)$$

$$\hat{\tau}_N = \Im(\bar{t})(V_{\zeta_i, \bar{t}}) \quad (41)$$

where $\mathbb{V} = \{v_{\zeta_i + \bar{t} - 1}\} \cup \{v_{\zeta_i + 2\bar{t} - 1}\}$, and $V_{\zeta_i, \bar{t}} = \{v_{\zeta_i + \bar{t}}, \dots, v_{\zeta_i + 2\bar{t} - 1}\}$.

In the second step, estimate the system parameters $\hat{\theta}_N$ and $\hat{\eta}_N$:

$$\hat{\theta}_N = \Omega^{-1} [q_1(\hat{\delta}_N), \dots, q_{n_1}(\hat{\delta}_N)]^T \quad (42)$$

$$\hat{\eta}_N = [q_{n_1+1}(\hat{\delta}_N), \dots, q_n(\hat{\delta}_N)]^T \quad (43)$$

$$\hat{\delta}_N^{(a)} = [\hat{\delta}_{N,1}^{(a)}, \dots, \hat{\delta}_{N,\mathcal{L}_0}^{(a)}]^T \quad (44)$$

$$\hat{\delta}_N = \sum_{a=0}^v \lambda_a \hat{\delta}_N^{(a)} \quad (45)$$

$$\hat{\delta}_{N,i}^{(a)} = C_a - \Psi^{-1} \left(\sum_{j=1}^{\mathcal{L}_0} [q_{ij}(\hat{\epsilon}_N, \hat{\tau}_N) \omega_{N,j}^{(a)}] \right) \quad (46)$$

$$\omega_{N,j}^{(a)} = \frac{1}{\tilde{N}_j} \sum_{k=1}^{N-N_0} v_k |_a \mathbf{I}_{\{\Pi_k^T = U_j, k \notin Z_N^{(0)}\}} \quad (47)$$

where $i = 1, \dots, \mathcal{L}_0$ and $j = 1, \dots, \mathcal{L}_0$.

Theorem 4. Consider the Wiener system (1) with multiple-valued observations (2) and (3) under the (ϵ, τ) replay attack described by (4). Under Assumptions 2, 3, and 4, as $N \rightarrow \infty$, one has $N_0 \rightarrow \infty$. Then, $\hat{\epsilon}_N \rightarrow \epsilon$ and $\hat{\tau}_N \rightarrow \tau$; moreover, $\hat{\theta}_N \rightarrow \theta$ and $\hat{\eta}_N \rightarrow \eta$. Here, the identification algorithm for $\hat{\epsilon}_N$ and $\hat{\tau}_N$ is given by (40)–(41), while the identification algorithm for $\hat{\theta}_N$ and $\hat{\eta}_N$ is given by (42)–(47).

Proof. According to (28), when $k \notin Z_N^{(0)}$, one has $\Pr\{v_k \leq a\} = \epsilon^{(a)}$. When $k \in Z_N^{(0)}$, if $v_{\zeta_i + \bar{t} - 1} \neq \bar{t} - 1$ or $v_{\zeta_i + 2\bar{t} - 1} \neq \bar{t} - 1$ is observed, then the occurrence of a replay attack is indicated. Moreover, the observed $\hat{\tau}_N$ represents the delay introduced by the replay attack.

Since the marked samples are independently collected and their number satisfies $N_0 \rightarrow \infty$ as $N \rightarrow \infty$, the empirical frequencies of the marked observations converge to their corresponding probability distributions according to the law of large numbers. Therefore, the attack occurrence probability and the attack delay can be consistently identified from the marked samples.

Based on the above analysis, according to the law of large numbers, one obtains

$$\hat{\epsilon}_N \rightarrow \epsilon, \hat{\tau}_N \rightarrow \tau, \omega_{N,i}^{(a)} \rightarrow \epsilon_i^{(a)}, \text{w.p.1, } N \rightarrow \infty \quad (48)$$

The convergence of $\hat{\epsilon}_N$ and $\hat{\tau}_N$ indicates that the statistical characteristics of the unknown replay attack are asymptotically obtained. Consequently, the influence of replayed observations can be compensated by using the estimated attack parameters. Combining this compensation relationship with the probability estimation method in (46), one obtains

$$\hat{\delta}_{N,i}^{(a)} \rightarrow \omega_i \quad (49)$$

where ω_i denotes the original observation probability without replay attacks. Therefore, (49) implies that the proposed method can asymptotically recover the attack-free probability information from the corrupted observations.

Finally, according to the inverse relationship between the observation probabilities and the system parameters established in (9)–(13), the consistency of the recovered probabilities directly guarantees the consistency of the estimated parameters.

The proof is completed. $\square$

5. Performance Analysis of the Consistent Estimation Algorithm

5.1. Asymptotic Normality Analysis

After establishing the consistent estimation algorithm, it is necessary to further investigate its statistical properties, particularly the asymptotic normality of the estimators, which provides the theoretical foundation for the subsequent optimality analysis.

Define $\mathcal{T} = \begin{bmatrix} \Omega^{-1} & 0 \\ 0 & \mathbf{I}_{n_1} \end{bmatrix}$, $\hat{A}_N = \begin{bmatrix} \hat{\theta}_N \\ \hat{\eta}_N \end{bmatrix}$, and $A = \begin{bmatrix} \theta \\ \eta \end{bmatrix}$, where $\mathbf{I}_{n_1}$ represents the identity matrix of dimension n_1 .

Theorem 5. Let $\zeta = \frac{N_0}{N}$ and define $\tilde{r}_j = \lim_{N \rightarrow \infty} \tilde{r}_{j,N} = \lim_{N \rightarrow \infty} \frac{\tilde{N}_j}{N - N_0}$. According to Theorem 4, the estimator $\hat{A}_N$ constructed by Algorithms (42)–(47) is asymptotically normally distributed, and its limiting distribution is given by

$$\sqrt{N}(\hat{A}_N - A) \xrightarrow{d} \Sigma(\mathfrak{R}) \quad (50)$$

where

$$\Sigma(\mathfrak{R}) = (1 + \frac{\mathfrak{R}}{3\bar{f}})\Gamma_1 + (1 + \frac{3\bar{f}}{\mathfrak{R}})\Gamma_2 \quad (51)$$

$$\Gamma_1 = 3\bar{f}\varphi_1\epsilon\epsilon^T\varphi_1^T(\epsilon - \epsilon^2) \quad (52)$$

$$\Gamma_2 = \varphi_2[\text{diag}(\epsilon) - \text{diag}(\epsilon)^2]\varphi_2^T \quad (53)$$

$$\varphi_1 = \mathcal{T}\mathcal{Y}(\omega)\mathcal{S}\frac{\partial\tilde{\mathfrak{N}}(\epsilon, \tau)}{\partial\epsilon} \quad (54)$$

$$\varphi_2 = \mathcal{T}\mathcal{Y}(\omega)\mathcal{S}\text{diag}(\frac{1}{\tilde{r}_{j,N}})\tilde{\mathfrak{N}}(\epsilon, \tau) \quad (55)$$

$$\varepsilon = \sum_{a=1}^v \lambda_a [\varepsilon_1^{(a)}, \dots, \varepsilon_{\mathcal{L}_0}^{(a)}]^T \quad (56)$$

Proof. According to (6) and (43), (45), and (46), one obtains

$$\begin{aligned} & \sqrt{N} \begin{bmatrix} \hat{\theta}_N - \theta \\ \hat{\eta}_N - \eta \end{bmatrix} \\ &= \sqrt{N} (\hat{A}_N - A) \\ &= \sqrt{N} \mathcal{T}(\varrho(\hat{\delta}_N) - \varrho(\omega)) \end{aligned} \quad (57)$$

According to the mean value theorem, there exists a random vector $\bar{a}$, whose elements lie between $\hat{\delta}_N$ and ω . Define $\mathcal{Y}(\bar{a}) = \frac{\partial \varrho(z)}{\partial z} \big|_{z=\bar{a}} = [\frac{\partial \varrho_1(\bar{a})}{\partial \bar{a}}, \dots, \frac{\partial \varrho_n(\bar{a})}{\partial \bar{a}}]$. Then,

$$\varrho(\hat{\delta}_N) - \varrho(\omega) = \mathbb{Y}(\bar{a})(\hat{\delta}_N - \omega) \quad (58)$$

Combining (44) and (45) yields

$$\hat{\delta}_N - \omega = \sum_{a=1}^v \lambda_a (\hat{\delta}_N^{(a)} - \omega) \quad (59)$$

Extracting the i -th component of (59), for $i = 1, \dots, \mathcal{L}_0$, gives

$$\begin{aligned} & \hat{\delta}_{N,i}^{(a)} - \omega_i \\ &= C_a - \Psi^{-1} \left(\sum_{j=1}^{\mathcal{L}_0} [q_{ij}(\hat{\varepsilon}_N, \hat{\tau}_N) \omega_{N,j}^{(a)}] \right) - \omega_i \\ &= \Psi^{-1}(\Psi(C_a - \omega_i)) - \Psi^{-1} \left(\sum_{j=1}^{\mathcal{L}_0} [q_{ij}(\hat{\varepsilon}_N, \hat{\tau}_N) \omega_{N,j}^{(a)}] \right) \end{aligned} \quad (60)$$

Applying the mean value theorem again to (60), there exists a vector $\bar{b}$ whose elements lie between $[\Psi(C_a - \omega_1), \dots, \Psi(C_a - \omega_{\mathcal{L}_0})]^T$ and $[\sum_{j=1}^{\mathcal{L}_0} [q_{1j}(\hat{\varepsilon}_N, \hat{\tau}_N) \omega_{N,j}^{(a)}], \dots, \sum_{j=1}^{\mathcal{L}_0} [q_{\mathcal{L}_0 j}(\hat{\varepsilon}_N, \hat{\tau}_N) \omega_{N,j}^{(a)}]]^T$. Hence, (60) can be further written as

$$\begin{aligned} & \Psi^{-1}(\Psi(C_a - \omega_i)) - \Psi^{-1} \left(\sum_{j=1}^{\mathcal{L}_0} [q_{ij}(\hat{\varepsilon}_N, \hat{\tau}_N) \omega_{N,j}^{(a)}] \right) \\ &= \frac{1}{\psi(\bar{b}_i)} \left[\Psi(C_a - \omega_i) - \sum_{j=1}^{\mathcal{L}_0} [q_{ij}(\hat{\varepsilon}_N, \hat{\tau}_N) \omega_{N,j}^{(a)}] \right] \\ &= \frac{1}{\psi(\bar{b}_i)} \left[\sum_{j=1}^{\mathcal{L}_0} q_{ij}(\varepsilon, \tau) \varepsilon_j^{(a)} - \sum_{j=1}^{\mathcal{L}_0} [q_{ij}(\hat{\varepsilon}_N, \hat{\tau}_N) \omega_{N,j}^{(a)}] \right] \end{aligned} \quad (61)$$

Define $B_{N,i,1}^{(a)} = \sum_{j=1}^{\mathcal{L}_0} q_{ij}(\varepsilon, \tau) \varepsilon_j^{(a)} - \sum_{j=1}^{\mathcal{L}_0} q_{ij}(\hat{\varepsilon}_N, \hat{\tau}_N) \varepsilon_j^{(a)}$ and $B_{N,i,2}^{(a)} = \sum_{j=1}^{\mathcal{L}_0} q_{ij}(\hat{\varepsilon}_N, \hat{\tau}_N) \varepsilon_j^{(a)} - \sum_{j=1}^{\mathcal{L}_0} [q_{ij}(\hat{\varepsilon}_N, \hat{\tau}_N) \omega_{N,j}^{(a)}]$. Then, (61) can be expressed as the sum of $\frac{1}{\psi(\bar{b}_i)} B_{N,i,1}^{(a)}$ and $\frac{1}{\psi(\bar{b}_i)} B_{N,i,2}^{(a)}$. The two terms are analyzed separately below.

First, there exists a scalar $\bar{\epsilon}$ lying between $\hat{\epsilon}_N$ and ϵ . By the mean value theorem,

$$\begin{aligned} & \sqrt{N}B_{N,i,1}^{(a)} \\ &= \sqrt{N} \sum_{j=1}^{\mathcal{L}_0} \epsilon_j^{(a)} (q_{ij}(\epsilon, \tau) - q_{ij}(\hat{\epsilon}_N, \hat{\tau}_N)) \\ &= \sqrt{N} \sum_{j=1}^{\mathcal{L}_0} \epsilon_j^{(a)} [q_{ij}(\hat{\epsilon}_N, \tau) - q_{ij}(\hat{\epsilon}_N, \hat{\tau}_N)] - \frac{\sqrt{3\bar{f}N}}{\sqrt{N_0}} \sum_{j=1}^{\mathcal{L}_0} \epsilon_j^{(a)} \frac{\partial q_{ij}(x, \tau)}{\partial x} \Big|_{x=\bar{\epsilon}} \sqrt{\frac{N_0}{3\bar{f}}} (\hat{\epsilon}_N - \epsilon) \end{aligned} \quad (62)$$

The term $B_{N,i,2}^{(a)}$ can be decomposed as

$$\begin{aligned} & \sqrt{N}B_{N,i,2}^{(a)} \\ &= \sum_{j=1}^{\mathcal{L}_0} q_{ij}(\hat{\epsilon}_N, \hat{\tau}_N) \frac{\sqrt{N}}{\sqrt{\tilde{N}_i}} \sqrt{\tilde{N}_i} [\epsilon_j^{(a)} - \omega_{N,j}^{(a)}] \end{aligned} \quad (63)$$

Define $\mathcal{S} = \text{diag}(\frac{1}{\psi(\bar{b}_1)}, \dots, \frac{1}{\psi(\bar{b}_{\mathcal{L}_0})})$. Combining (57)–(63) yields

$$\sqrt{N}(\hat{A}_N - A) = D_{N,1} + D_{N,2} + D_{N,3} \quad (64)$$

where

$$D_{N,1} = -\mathcal{T}\mathcal{Y}(\bar{a})\mathcal{S} \frac{\sqrt{3\bar{f}N}}{\sqrt{N_0}} \sum_{a=1}^v \sum_{j=1}^{\mathcal{L}_0} \lambda_a \epsilon_j^{(a)} \begin{bmatrix} \frac{\partial q_{1j}(x, \tau)}{\partial x} \Big|_{x=\bar{\epsilon}} \sqrt{\frac{N_0}{3\bar{f}}} (\hat{\epsilon}_N - \epsilon) \\ \dots \\ \frac{\partial q_{\mathcal{L}_0 j}(x, \tau)}{\partial x} \Big|_{x=\bar{\epsilon}} \sqrt{\frac{N_0}{3\bar{f}}} (\hat{\epsilon}_N - \epsilon) \end{bmatrix} \quad (65)$$

$$D_{N,2} = \mathcal{T}\mathcal{Y}(\bar{a})\mathcal{S} \sum_{a=1}^v \lambda_a \begin{bmatrix} \sqrt{\frac{1}{\bar{r}_{1,N}} \cdot \frac{N}{N-N_0}} \sum_{j=1}^{\mathcal{L}_0} q_{1j}(\hat{\epsilon}_N, \hat{\tau}_N) \sqrt{\tilde{N}_1} [\epsilon_j^{(a)} - \omega_{N,j}^{(a)}] \\ \dots \\ \sqrt{\frac{1}{\bar{r}_{\mathcal{L}_0,N}} \cdot \frac{N}{N-N_0}} \sum_{j=1}^{\mathcal{L}_0} q_{\mathcal{L}_0 j}(\hat{\epsilon}_N, \hat{\tau}_N) \sqrt{\tilde{N}_{\mathcal{L}_0}} [\epsilon_j^{(a)} - \omega_{N,j}^{(a)}] \end{bmatrix} \quad (66)$$

$$D_{N,3} = \mathcal{T}\mathcal{Y}(\bar{a})\mathcal{S} \sum_{a=1}^v \sum_{j=1}^{\mathcal{L}_0} \lambda_a \epsilon_j^{(a)} \begin{bmatrix} \sqrt{N} [q_{1j}(\hat{\epsilon}_N, \tau) - q_{1j}(\hat{\epsilon}_N, \hat{\tau}_N)] \\ \dots \\ \sqrt{N} [q_{1j}(\hat{\epsilon}_N, \tau) - q_{\mathcal{L}_0 j}(\hat{\epsilon}_N, \hat{\tau}_N)] \end{bmatrix} \quad (67)$$

Since $\mathbb{E}(v_k - \mathbb{E}v_k)^2 = \mathbb{E}^2 v_k - (\mathbb{E}v_k)^2 = \mathbb{E}v_k - (\mathbb{E}v_k)^2, k \notin \mathbb{Z}_N^{(0)}$, according to (28), one has

$$\mathbb{E}(v_k - \mathbb{E}v_k)^2 = \epsilon_j - \epsilon_j^2, k \notin \mathbb{Z}_N^{(0)} \quad (68)$$

From (39) and (48), it follows that

$$\mathbb{E}(I_{\{v_k \neq \bar{t}-1, v_k \in \mathbb{V}\}}) - \mathbb{E}^2(I_{\{v_k \neq \bar{t}-1, v_k \in \mathbb{V}\}}) = \epsilon - \epsilon^2 \quad (69)$$

By the central limit theorem,

$$\sqrt{\frac{N_0}{3\bar{f}}} (\hat{\epsilon}_N - \epsilon) \xrightarrow{d} \mathcal{N}(0, \epsilon - \epsilon^2), N \rightarrow \infty \quad (70)$$

$$\sqrt{\tilde{N}_i} (\epsilon_j^{(a)} - \omega_{N,j}^{(a)}) \xrightarrow{d} \mathcal{N}(0, \epsilon_j - \epsilon_j^2), N \rightarrow \infty \quad (71)$$

Let ℓ_i denote the length of the normal data segment before the i -th data processing block, namely, the interval between two adjacent coding blocks. Assume that the random sequence $\{\ell_i\}$ is independent and identically distributed and follows a Poisson distribution with parameter $\mathfrak{R}$. Then, $\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^{g_N} \ell_i}{g_N} = \mathfrak{R}$. According to the data preprocessing scheme in Section 4.2, one has

$$\frac{N}{N_0} \rightarrow \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^{g_N} \ell_i + 3\bar{f}g_N}{3\bar{f}g_N} = 1 + \frac{\mathfrak{R}}{3\bar{f}} \quad (72)$$

$$\frac{N}{N - N_0} \rightarrow \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^{g_N} \ell_i + 3\bar{f}g_N}{\sum_{i=1}^{g_N} \ell_i} = 1 + \frac{3\bar{f}}{\mathfrak{R}} \quad (73)$$

According to (70)–(73),

$$D_{N,1} \xrightarrow{d} \mathcal{N}\left(0, \left(1 + \frac{\mathfrak{R}}{3\bar{f}}\right)\Gamma_1\right) \quad (74)$$

$$D_{N,2} \xrightarrow{d} \mathcal{N}\left(0, \left(1 + \frac{3\bar{f}}{\mathfrak{R}}\right)\Gamma_2\right) \quad (75)$$

$$D_{N,3} \rightarrow \mathbf{0}, \quad N \rightarrow \infty \quad (76)$$

Therefore,

$$\sqrt{N}(\hat{A}_N - A) = D_{N,1} + D_{N,2} + D_{N,3} \xrightarrow{d} \mathcal{N}\left(0, \left(1 + \frac{\mathfrak{R}}{3\bar{f}}\right)\Gamma_1 + \left(1 + \frac{3\bar{f}}{\mathfrak{R}}\right)\Gamma_2\right) \quad (77)$$

The proof is completed. $\square$

Remark 3. The asymptotic covariance matrix Σ reflects the influence of replay attacks on the identification accuracy. Specifically, the covariance term without attack represents the inherent statistical uncertainty of the binary-valued observations, while the additional attack-related term characterizes the uncertainty introduced by replayed data. A larger attack probability or replay duration increases the correlation of received data and reduces the effective identification information, leading to an increase of Σ and a degradation of estimation accuracy. Therefore, Σ provides a quantitative measure of the impact of replay attacks on parameter identification performance.

5.2. Optimality Analysis

The marking mechanism introduces N_0 marked data points, which are used exclusively for attack parameter estimation and are unavailable for system parameter identification. The fraction of data used for marking (overhead) is N_0/N . As $N \rightarrow \infty$, according to (72), this overhead fraction converges to

$$\frac{N_0}{N} \rightarrow \frac{3\bar{f}}{3\bar{f} + \mathfrak{R}}. \quad (78)$$

Thus, the design parameter $\mathfrak{R}$ directly controls the tradeoff between data preservation and attack estimation accuracy. A small $\mathfrak{R}$ corresponds to frequent marking, which provides more samples for accurate attack estimation but reduces the data available for system identification. Conversely, a large $\mathfrak{R}$ preserves more data for identification but provides fewer samples for attack estimation.

According to Theorem 5, the Poisson distribution parameter $\mathfrak{R}$ affects the asymptotic normality of the parameter estimator through the factors $(1 + \mathfrak{R}/3\bar{f})$ and $(1 + 3\bar{f}/\mathfrak{R})$ in the covariance expression. Therefore, an optimal consistent estimation algorithm can be designed to minimize the variance of the estimation error, namely, the discrepancy between the estimated values and the true parameters. Specifically, the term involving $(1 + \mathfrak{R}/3\bar{f})$

originates from the uncertainty in estimating the attack probability ϵ , and decreases as marking becomes more frequent ($\Re$ decreases). The term involving $(1 + 3\bar{f}/\Re)$ originates from the finite-sample uncertainty in the system parameter estimation and decreases as more data are preserved for identification ($\Re$ increases). The optimal $\Re^*$ balances these two competing sources of uncertainty. Consequently, this subsection aims to determine the optimal $\Re$ that minimizes the trace of the asymptotic covariance matrix.

Let $\text{trace}(\cdot)$ denote the trace of the target covariance matrix. Then, the optimality analysis can be formulated as the following parameter optimization problem:

$$\min_{\Re} \left\{ \left(1 + \frac{\Re}{3\bar{f}}\right) \text{trace}(\Gamma_1) + \left(1 + \frac{3\bar{f}}{\Re}\right) \text{trace}(\Gamma_2) \right\} \quad (79)$$

Denote the solution of the above optimization problem as the optimal design strategy $\Re^*$ for the consistent estimation algorithm.

Theorem 6. *The optimal design strategy $\Re^*$ for the consistent estimation algorithm given by (79) is*

$$\Re^* = 3\bar{f} \sqrt{\frac{\Lambda_2}{\Lambda_1}} \quad (80)$$

where $\Lambda_1 = \text{trace}(\Gamma_1) \geq 0$ and $\Lambda_2 = \text{trace}(\Gamma_2) \geq 0$.

Proof. According to the arithmetic–geometric mean inequality,

$$\text{trace}(\Sigma(\Re)) = \left(1 + \frac{\Re}{3\bar{f}}\right) \Lambda_1 + \left(1 + \frac{3\bar{f}}{\Re}\right) \Lambda_2 \geq \Lambda_1 + \Lambda_2 + 2\sqrt{\Lambda_1 \Lambda_2} \quad (81)$$

The equality holds if and only if $\frac{\Re}{3\bar{f}} \Lambda_1 = \frac{3\bar{f}}{\Re} \Lambda_2$. Since $\Re \geq 0$, solving the above equation yields $\Re^* = 3\bar{f} \sqrt{\frac{\Lambda_2}{\Lambda_1}}$. The proof is completed. $\square$

6. Numerical Simulations

To verify the effectiveness of the proposed consistent identification algorithm, this paper considers a class of nonlinear quantized systems and conducts numerical experiments in the presence of replay attacks.

The system output is defined as $y_k = 2^{u_k^T \theta} + \eta + d_k$, where the unknown parameter is $\theta = [2, 4]^T$, the bias term is $\eta = 1$, and the noise satisfies $d_k \sim \mathcal{N}(0, \sigma^2)$ with variance $\sigma = 3$. The input signal is finitely quantized as $u_k \in \{-0.1, 0.2\}$, thus yielding $L_0 = 4$ possible regression patterns. The system output is further processed through a multi-threshold quantizer, where the threshold set is $C = \{1, 2, 3\}$, $v = 3$.

To verify the accuracy of the attack-free estimation algorithm (Equations (12)–(14)), the attack-free scenario is first considered. As shown in Figure 4 (Left: Estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ converge to true values $\theta_1 = 2$ and $\theta_2 = 4$. Right: Estimate $\hat{\eta}$ converges to true value $\eta = 1$. Parameters: $N = 6000$, $\sigma = 3$, $\epsilon = 0$, $C = \{1, 2, 3\}$, $\lambda = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$), the proposed attack-free estimation algorithm can accurately estimate θ and η under quantized input and multi-threshold measurement constraints, thereby validating the effectiveness of the attack-free estimation algorithm.

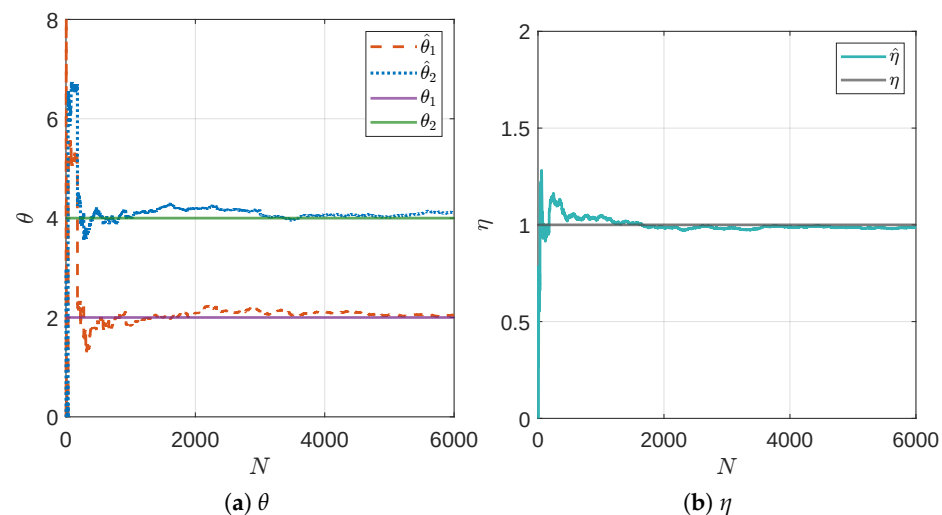


Figure 4. Effectiveness of the attack-free estimation algorithm (Equations (12)–(14)).

Next, the presence of replay attacks is considered, where the replay attack strategy is specified as $\epsilon = 0.3$ and $\tau = 1$. To verify whether the theoretical convergence values of θ and η given by (25)–(27) are correct when the original attack-free estimation algorithm is applied under replay attacks, the results shown in Figure 5 are provided (Left: Estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ converge to the biased limits $\check{\theta}_1$ and $\check{\theta}_2$, not to true values $\theta_1 = 2$ and $\theta_2 = 4$. Right: Estimate $\hat{\eta}$ converges to biased limit $\check{\eta}$, not to true $\eta = 1$. Parameters: $N = 6000$, $\sigma = 3$, $\epsilon = 0.3$, $\tau = 1$, $C = \{1, 2, 3\}$, $\lambda = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$). The simulation curves validate the correctness of the theoretical convergence values proposed in (25)–(27).

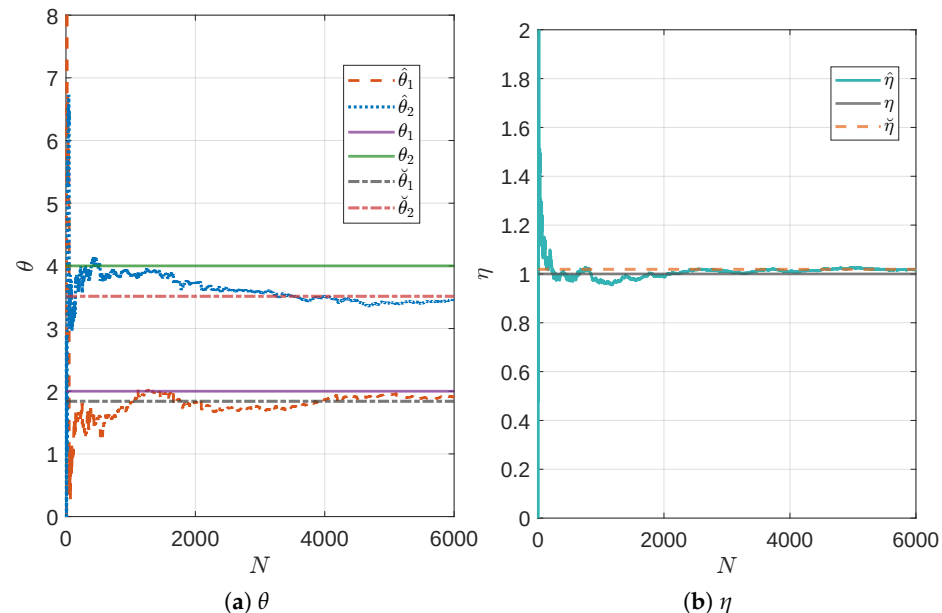


Figure 5. Effectiveness of the theoretical convergence results (Equations (25)–(27)).

To compensate for the parameter estimation bias caused by replay attacks, a consistent estimation algorithm with known attack strategy is first proposed. Assume that the estimator has prior knowledge of the attack strategy $(\epsilon, \tau) = (0.3, 1)$. As shown in Figure 6 (Left: Compensated estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ converge to true values $\theta_1 = 2$ and $\theta_2 = 4$. Right: Estimate $\hat{\eta}$ converges to $\eta = 1$. Parameters: $N = 30,000$, $\sigma = 3$, $\epsilon = 0.3$, $\tau = 1$, $C = \{1, 2, 3\}$, $\lambda = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$), the simulations demonstrate that the proposed consistent

estimation algorithm with known attack strategy can effectively eliminate the influence of replay attacks and achieve accurate parameter estimation.

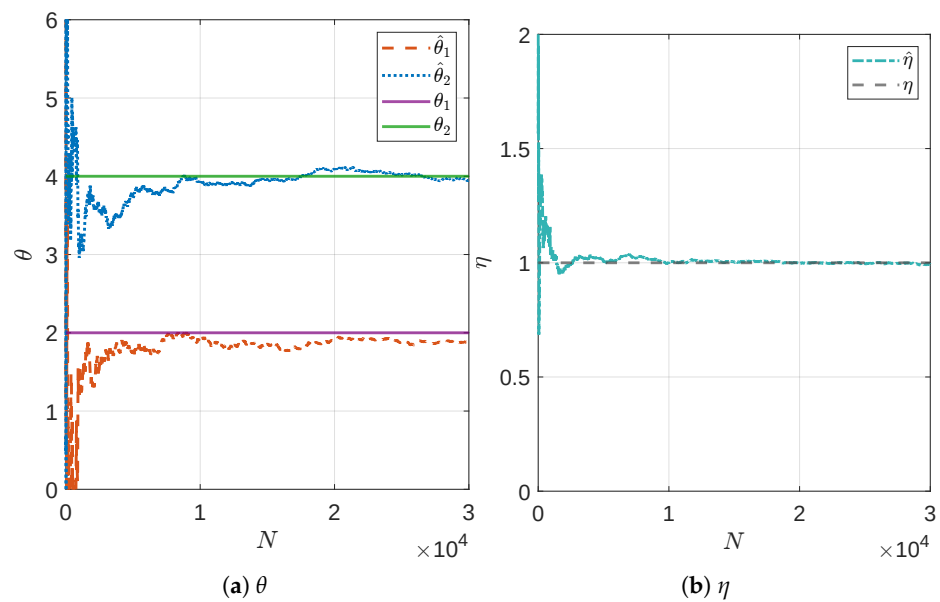


Figure 6. Consistent estimation performance with known attack strategy.

Next, the case of unknown attack strategy is considered to verify the effectiveness of the consistent identification algorithms (Equations (42)–(47)). Let $\bar{\tau} = 3$ and define $\bar{t} = \bar{\tau} + 2 = 5$. The preprocessed data sequence is generated recursively according to (39), where a specially structured detection sequence is constructed to effectively identify replay attacks. The parameter $\mathfrak{R}$ is chosen as $\mathfrak{R} = 20$, and the true attack strategy is set as $(0.3, 3)$. The results are shown in Figure 7 (Left: Estimates of attack strategy converge: $\hat{\epsilon}_N$ to $\epsilon = 0.3$, $\hat{\tau}_N$ to $\tau = 3$. Right: Estimates $\hat{\theta}_1$, $\hat{\theta}_2$, and $\hat{\eta}$ converge to true values $\theta_1 = 2$, $\theta_2 = 4$, and $\eta = 1$). Parameters: $N = 30,000$, $\sigma = 3$, $C = \{1, 2, 3\}$, $\lambda = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$. The simulations demonstrate that the consistent identification algorithms (Equations (42)–(47)) can accurately estimate the attack strategy and subsequently estimate the system parameters.

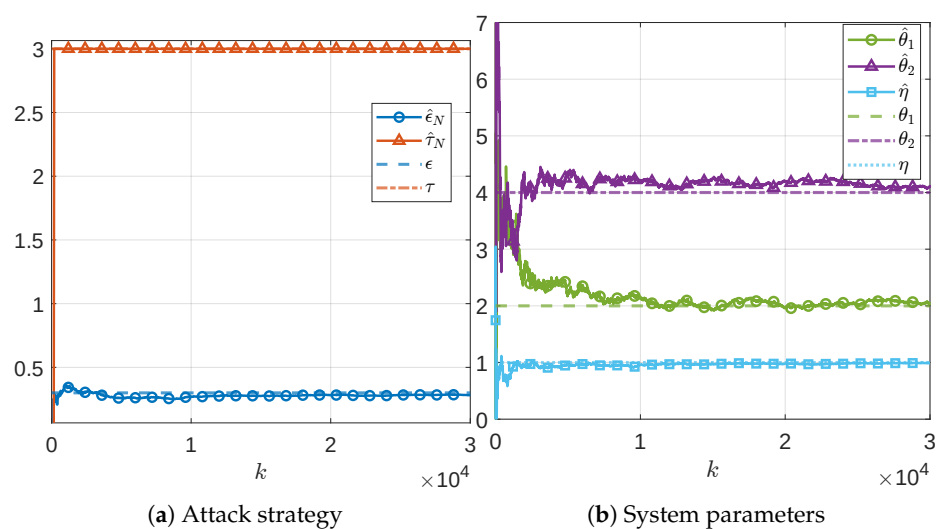


Figure 7. Consistent estimation performance with unknown attack strategy.

To verify Theorem 5, the asymptotic normality of θ_1 , θ_2 , and η is investigated through simulations, as shown in Figure 8. (Histograms show empirical distributions over

MC = 12,000 runs; solid curves show theoretical $\mathcal{N}(0, \Sigma)$ densities. Left: $\sqrt{N}(\hat{\theta}_{N,1} - \theta_1)$. Middle: $\sqrt{N}(\hat{\theta}_{N,2} - \theta_2)$. Right: $\sqrt{N}(\hat{\eta}_N - \eta)$. Parameters: $N = 30,000, \sigma = 3, \epsilon = 0.3, \tau = 3, \Re = 20, C = \{1, 2, 3\}, \lambda = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$.

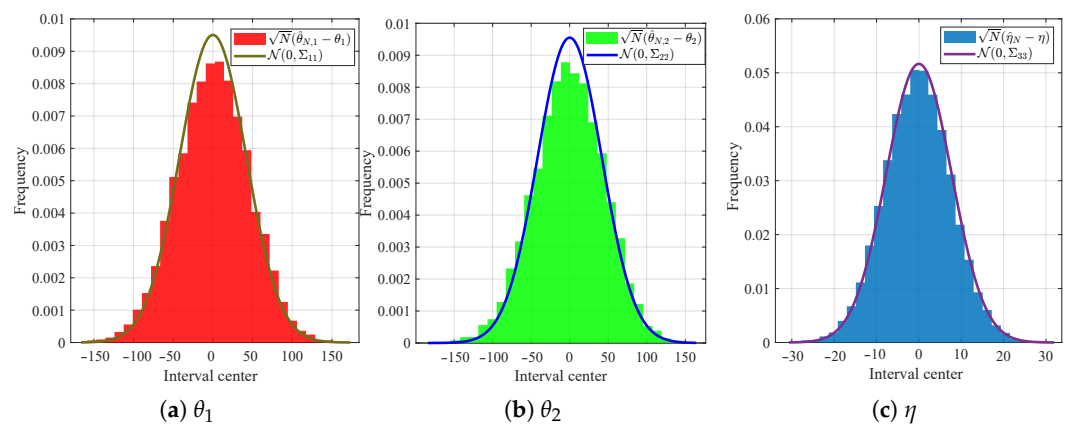


Figure 8. Asymptotic normality under the proposed consistent estimation algorithm with unknown attack strategy (42)–(47).

According to Theorem 6, the optimal defense parameter is $\Re^* = 86.0234$. Meanwhile, two non-optimal parameters are selected as $\Re_1 = 20$ and $\Re_2 = 130$. Parameter estimation is then carried out using the consistent identification algorithms (Equations (42)–(47)), and the results are shown in Figure 9 (The optimal $\Re^* = 86.0234$ achieves the lowest asymptotic trace. Non-optimal choices $\Re = 20$ and $\Re = 130$ produce larger asymptotic values. $N = 40,000, MC = 100$). The results indicate that the optimal consistent estimation strategy minimizes the trace of the covariance matrix and thereby achieves the minimum estimation error.

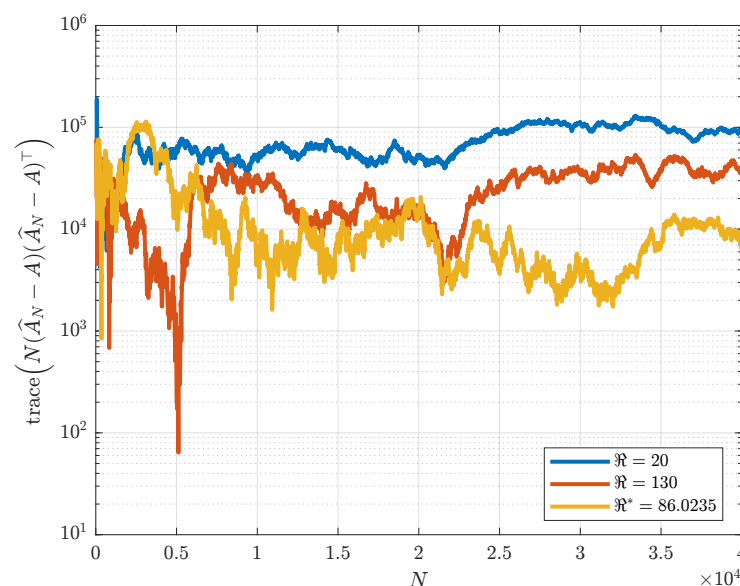


Figure 9. Trace of the covariance matrix $\text{trace}(N(\hat{A}_N - A)(\hat{A}_N - A)^T)$ under different values of $\Re$.

To further investigate the influence of the labeling overhead on the identification performance, an additional simulation is conducted. The evolution of the normalized estimation error under different labeled-data proportions is shown in Figure 10.

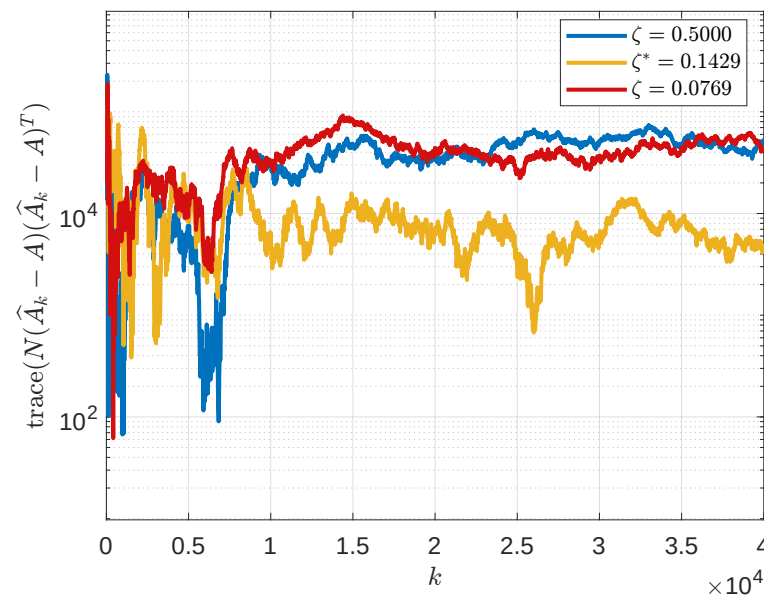


Figure 10. Effect of the labeled-data proportion ζ on the parameter identification performance.

To further validate the effectiveness of the proposed algorithm with unknown attack strategy, a comparative analysis is conducted against three representative methods. The first comparison method (denoted as M_2) estimates the attack parameters ϵ and τ purely from the statistical match rate of the received data without structured markers, and then applies the same compensation-based estimation procedure as in the consistent identification algorithms (Equations (42)–(47)). The second comparison method (denoted as M_3) adopts the dynamic watermarking framework [28]: a known pseudo-random watermark is injected into the system output, attacked segments are identified via sliding-window correlation detection and removed, and finally the attack-free estimation algorithm (12)–(14) is applied to the remaining clean data. The third comparison method (denoted as M_4) employs the classical chi-squared (χ^2) residual-based anomaly detector, which is one of the most widely recognized bad-data detection frameworks in the cyber-physical security literature [29,30]. In M_4 , the χ^2 statistic constructed from the observation residuals under the attack-free model is compared against a threshold to detect anomalies. The detected attacked segments are removed, and the attack-free estimation algorithm is applied to the retained clean data. Unlike M_1 , M_4 does not embed structured markers or estimate the attack strategy; instead, it relies on the statistical deviation between the observed and predicted measurement distributions to identify attack occurrences.

The results in Figures 11 and 12 demonstrate that:

1. The proposed method (M_1) achieves the smallest bias for all parameters, indicating that the structured marking mechanism enables more accurate estimation of the attack strategy and effective compensation.
2. M_2 shows larger estimation bias, especially for θ_2 and η , since the attack parameters estimated from empirical match rates are less reliable without deterministic markers, resulting in an inaccurate compensation matrix.
3. Both M_3 and M_4 (detection-and-removal methods) obtain larger RMSE values than M_1 due to the data removal process, which reduces the available samples and increases estimation variance. In contrast, the proposed compensation-based method retains the received data and achieves better estimation accuracy.
4. Comparing M_3 and M_4 , the χ^2 detector (M_4) achieves similar RMSE to the watermarking method (M_3), since both discard attacked segments. However, the χ^2 detector is a conventional methods from the cyber-physical security literature and does not

require signal modification (watermark injection), which preserves the original system dynamics and avoids potential interference with normal system operation. Nevertheless, its accuracy is still significantly lower than that of the proposed M_1 , confirming the superiority of the structured-marking-based compensation paradigm over the detection-and-removal paradigm.

5. M_1 and M_2 have comparable computational costs, whereas M_3 and M_4 require additional computation due to the sliding-window detection process. Specifically, M_4 needs to periodically update the covariance estimate of the residual buffer, leading to a higher per-sample computational overhead.

To assess the practical feasibility of the proposed algorithm, the computation time is evaluated against the data length N and the number of quantization levels v . Figure 13 shows the measured computation time as a function of N (with $v = 3$ fixed) and v (with $N = 30,000$ fixed). With $\mathcal{L}_0 = 4$ in the current simulation setting, the per-step cost is modest, and the total computation time remains well within practical limits even for $N = 10^5$.

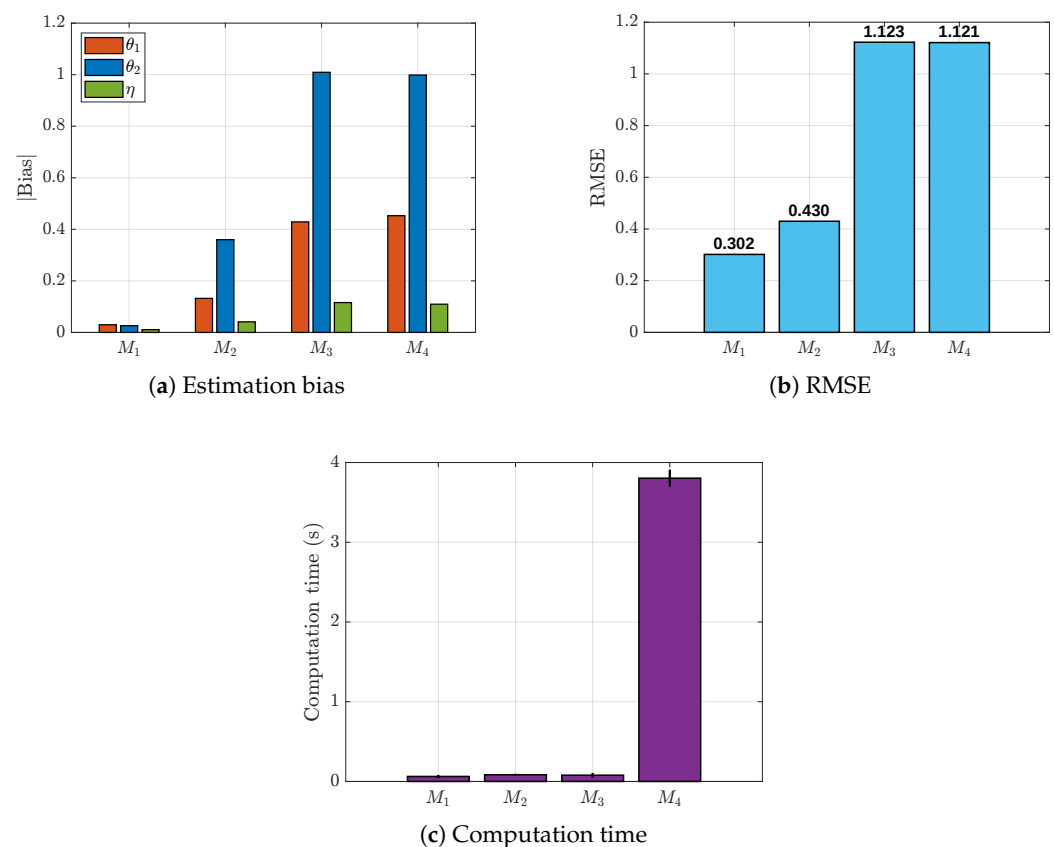


Figure 11. Comparison of the proposed method (M_1) with three existing approaches for unknown attack strategy (a) Absolute estimation bias of θ_1 , θ_2 , and η . (b) Overall RMSE of the parameter vector $[\theta^T, \eta]^T$. (c) Mean computation time with $\pm$ one standard deviation error bars. M_1 : proposed method (structured marking + attack strategy estimation + compensated estimation, Algorithms (40)–(47)); M_2 : empirical attack strategy estimation + compensation (no structured markers); M_3 : dynamic watermarking detection + removal + attack-free estimation [28]; M_4 : χ^2 residual-based detection + removal + attack-free estimation [29,30]. ($\epsilon = 0.3$, $\tau = 3$, $N = 40,000$, 30 Monte Carlo runs).

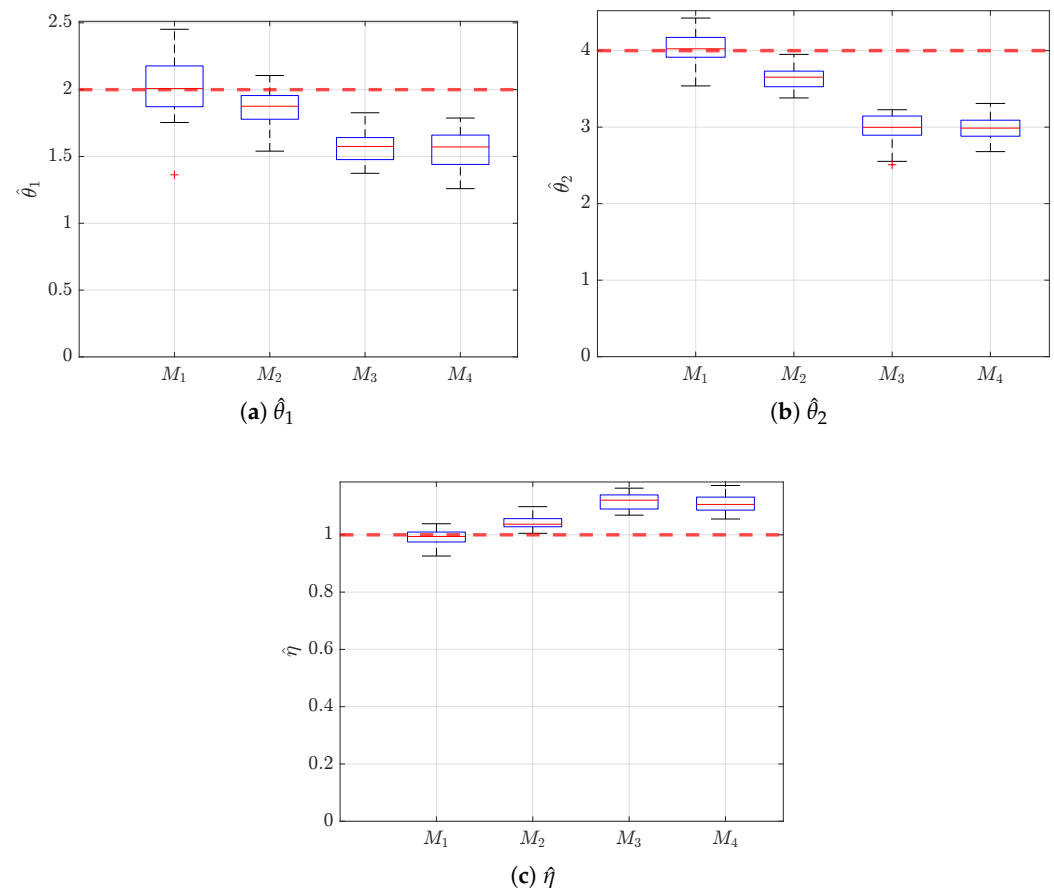


Figure 12. Boxplots of the parameter estimates obtained by the four methods over 30 Monte Carlo runs. Red dashed lines indicate the true parameter values ($\theta_1 = 2$, $\theta_2 = 4$, $\eta = 1$). M_1 : proposed method (structured marking + attack strategy estimation + compensated estimation, Algorithms (40)–(47)); M_2 : empirical attack strategy estimation + compensation (no structured markers); M_3 : dynamic watermarking detection + removal + attack-free estimation [28]; M_4 : χ^2 residual-based detection + removal + attack-free estimation [29,30].

The blue boxes represent the interquartile range, covering the 25th–75th percentiles of the estimates, while the red lines inside the boxes indicate the median estimates. The red dashed line denotes the true parameter value.

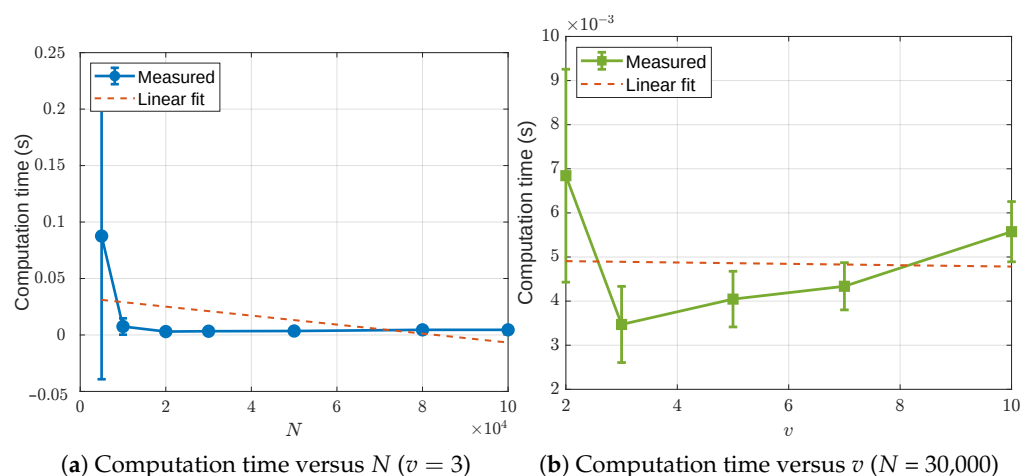


Figure 13. Computational cost of the proposed estimation algorithm. (a) Scaling with data length N under a fixed number of quantization levels $v = 3$. (b) Scaling with the number of quantization levels v under a fixed data length $N = 30,000$. The dashed lines indicate linear fits.

To assess the robustness of the compensated estimation algorithm to varying attack parameters, the estimation error $\|\hat{A}_N - A\|$ is evaluated under different attack probabilities and attack delays using the known attack strategy setting. Figure 14 shows that the algorithm converges consistently for all tested values. The convergence rate is slightly slower for larger ϵ because a larger fraction of data is corrupted, while the delay τ has a negligible effect on the asymptotic accuracy as long as it is correctly known.

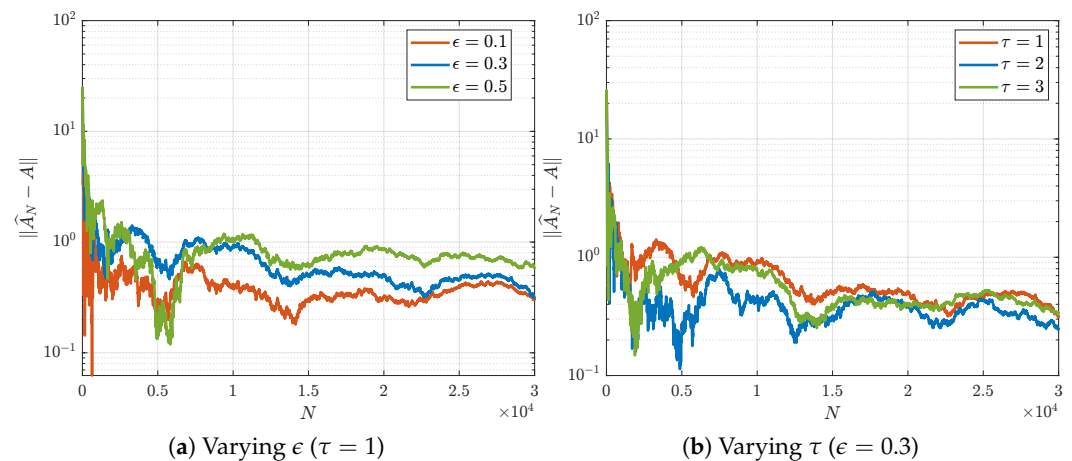


Figure 14. Sensitivity of the compensated estimation algorithm (known attack strategy) to varying attack parameters. (a) Estimation error $\|\hat{A}_N - A\|$ under $\epsilon = 0.1, 0.3, 0.5$ with $\tau = 1$. (b) Estimation error under $\tau = 1, 2, 3$ with $\epsilon = 0.3$. ($N = 30,000$, $\sigma = 3$).

7. Conclusions

This paper investigates the parameter identification problem for Wiener systems with quantized inputs and multi-threshold measurements under replay attacks. The main results are summarized as follows. First, the effect of replay attacks on the attack-free identification algorithm is investigated, and it is shown that the original algorithm loses consistency under replay attacks, and the corresponding theoretical biased expressions of these limit values are derived. Second, to address the performance degradation caused by replay attacks, two types of consistent estimation algorithms are proposed. When the attack strategy is known, unbiased consistent estimation is achieved by introducing a compensation matrix to correct the observed data. When the attack strategy is unknown, a structured-marking-based data preprocessing method is developed to effectively estimate the attack parameters, and then achieve consistent identification of the system parameters. Third, the asymptotic statistical properties of the proposed estimation algorithms are established. It is proved that the estimators satisfy asymptotic normality, and the covariance structure is explicitly derived. Based on this result, an optimal design method minimizing the trace of the covariance matrix is proposed, which provides a theoretical guideline for parameter selection in the proposed algorithms.

The proposed methods have potential applications in networked control systems, industrial process monitoring, and cyber-physical system security, especially in complex environments involving communication attacks and quantized observation constraints. It should be noted that the proposed results are obtained under the assumption of the considered noise model and replay attack structure. Specifically, the influence of non-stationary noise and more sophisticated replay attacks with time-varying attack probabilities or delays has not been fully investigated. Addressing these challenging scenarios will be an important direction for future research. Future research will focus on extending the proposed framework to more general noise conditions and complex replay attack scenarios, as well as investigating secure control strategies under quantization constraints.

Author Contributions: J.L.: Conceptualization, Methodology, Software, Writing—Original Draft; J.G.: Formal analysis, Writing—Original Draft; Q.Z.: Validation, Investigation, Writing—Review & Editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Persis, C.D.; Tesi, P. Formulas for data-driven control: Stabilization, optimality, and robustness. *IEEE Trans. Autom. Control* **2019**, *65*, 909–924. [\[CrossRef\]](#)
2. Liu, G.; Wang, J. Dendrite net: A white-box module for classification, regression, and system identification. *IEEE Trans. Cybern.* **2021**, *52*, 13774–13787. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Yu, P.; Hu, Y.; Wang, Y.; Jia, R.; Guo, J. Optimal consensus control strategy for multi-agent systems under cyber attacks via a stackelberg game approach. *IEEE Trans. Autom. Sci. Eng.* **2025**, *22*, 18875–18888. [\[CrossRef\]](#)
4. Noormohammadi-Asl, A.; Esrafilian, O.; Arzati, M.A.; Taghirad, H.D. System identification and h_∞-based control of quadrotor attitude. *Mech. Syst. Signal Process.* **2020**, *135*, 106358. [\[CrossRef\]](#)
5. Zhang, Q.; Zhao, Y.; Xue, W.; Guo, J. Parameter identification for FIR systems with simultaneously quantized inputs and outputs against replay attacks: A stochastic defense scheme. *IEEE Trans. Autom. Control* **2026**. [\[CrossRef\]](#)
6. Sharma, P.; Ajjarapu, V.; Vaidya, U. Data-driven identification of nonlinear power system dynamics using output-only measurements. *IEEE Trans. Power Syst.* **2021**, *37*, 3458–3468. [\[CrossRef\]](#)
7. Ji, Y.; Huang, Y.; Zeng, J.; Ren, L.; Chen, Y. A physical–data-driven combined strategy for load identification of tire type rail transit vehicle. *Reliab. Eng. Syst. Saf.* **2025**, *253*, 110493. [\[CrossRef\]](#)
8. Liu, W.; Wang, Y.; Guo, J. Consistent identification for FIR systems with multilevel quantized observations subjected to data tampering attacks: A joint estimation approach. *IEEE Trans. Ind. Electron.* **2025**, *73*, 7676–7686. [\[CrossRef\]](#)
9. Adil, M.; Farouk, A.; Abulkasim, H.; Ali, A.; Song, H.; Jin, Z. Ng-icps: Next generation industrial-cps, security threats in the era of artificial intelligence, and open challenges with future research directions. *IEEE Internet Things J.* **2024**, *12*, 1343–1367. [\[CrossRef\]](#)
10. Golabi, A.; Erradi, A.; Tantawy, A. Towards automated hazard analysis for CPS security with application to cstr system. *J. Process Control* **2022**, *115*, 100–111. [\[CrossRef\]](#)
11. Hou, J.; Wang, H.; Su, H.; Chen, F.; Liu, J. A bias-correction modeling method of Hammerstein–Wiener systems with polynomial nonlinearities using noisy measurements. *Mech. Syst. Signal Process.* **2024**, *213*, 111329. [\[CrossRef\]](#)
12. Dong, S.; Yu, L.; Zhang, W.-A.; Chen, B. Robust hierarchical identification of Wiener systems in the presence of dynamic disturbances. *J. Frankl. Inst.* **2020**, *357*, 3809–3834. [\[CrossRef\]](#)
13. Kang, Y.; Zhai, D.-H.; Liu, G.-P.; Zhao, Y.-B. On input-to-state stability of switched stochastic nonlinear systems under extended asynchronous switching. *IEEE Trans. Cybern.* **2015**, *46*, 1092–1105. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Giuseppe, Q.; Lacarbonara, W.; Masri, S.F. A review on computational intelligence for identification of nonlinear dynamical systems. *Nonlinear Dyn.* **2020**, *99*, 1709–1761. [\[CrossRef\]](#)
15. Zhang, Y.; Jia, L.; Li, F. Identification of wiener system with time delay using correlation analysis and swarm intelligence methods. *Trans. Inst. Meas. Control* **2026**, *48*, 1526–1536. [\[CrossRef\]](#)
16. Li, F.; Yang, Y.; Xia, Y. Identification for nonlinear systems modelled by deep long short-term memory networks based wiener model. *Mech. Syst. Signal Process.* **2024**, *220*, 111631. [\[CrossRef\]](#)
17. Bai, E.-W.; Reyland, J., Jr. Towards identification of Wiener systems with the least amount of a priori information: Iir cases. *Automatica* **2009**, *45*, 956–964. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Li, X.; Wang, T.; Guo, J.; Zhao, Y. Corrigendum to “Identification of ARMA models with binary-valued observations” [*Automatica* 149(2023) 110832]. *Automatica* **2025**, *175*, 112177. [\[CrossRef\]](#)
19. Zhang, G.; Yi, W.; Varshney, P.K.; Kong, L. Direct target localization with quantized measurements in noncoherent distributed mimo radar systems. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5103618. [\[CrossRef\]](#)
20. Tao, Y.; Tao, H.; Zhuang, Z.; Stojanovic, V.; Paszke, W. Quantized iterative learning control of communication-constrained systems with encoding and decoding mechanism. *Trans. Inst. Meas. Control* **2024**, *46*, 1943–1954. [\[CrossRef\]](#)

21. Guo, J.; Zhang, J.-F.; Zhao, Y. Adaptive tracking of a class of first-order systems with binary-valued observations and fixed thresholds. *J. Syst. Sci. Complex.* **2012**, *25*, 1041–1051. [[CrossRef](#)]
22. Wang, Y.; Zhao, Y.; Zhang, J.-F. Asymptotically efficient quasi-newton type identification with quantized observations under bounded persistent excitations. *Automatica* **2024**, *166*, 111722. [[CrossRef](#)]
23. Forgione, M.; Muni, A.; Piga, D.; Gallieri, M. On the adaptation of recurrent neural networks for system identification. *Automatica* **2023**, *155*, 111092. [[CrossRef](#)]
24. Jia, R.; Wang, T.; Xue, W.; Guo, J.; Zhao, Y. Multitime scale consensus algorithm of multiagent systems with binary-valued data under tampering attacks. *IEEE Trans. Ind. Inform.* **2025**, *21*, 9377–9388. [[CrossRef](#)]
25. Zaman, A.; Safarinejadian, B.; Birk, W. Security analysis and fault detection against stealthy replay attacks. *Int. J. Control* **2022**, *95*, 1562–1575. [[CrossRef](#)]
26. Guo, J.; Zhang, Q.; Zhao, Y. Identification of FIR systems with binary-valued observations under replay attacks. *Automatica* **2025**, *172*, 112001. [[CrossRef](#)]
27. Jiang, W.; Li, F.; Zhang, Q.; Jing, F.; Guo, J. Asymptotically consistent identification of FIR systems with binary-valued observations subject to sequence-type replay attacks. *Syst. Control Lett.* **2025**, *206*, 106286. [[CrossRef](#)]
28. Satchidanandan, B.; Kumar, P.R. Dynamic watermarking: Active defense of networked cyber–physical systems. *Proc. IEEE* **2016**, *105*, 219–240. [[CrossRef](#)]
29. Ye, D.; Zhang, T. Summation detector for false data-injection attack in cyber-physical systems. *IEEE Trans. Cybern.* **2019**, *50*, 2338–2345. [[CrossRef](#)] [[PubMed](#)]
30. Doostmohammadian, M.; Charalambous, T.; Shafie-khah, M.; Meskin, N.; Khan, U.A. Simultaneous distributed estimation and attack detection/isolation in social networks: Structural observability, kronecker-product network, and chi-square detector. In Proceedings of the 2021 IEEE International Conference on Autonomous Systems (ICAS), Montreal, QC, Canada, 11–13 August 2021; pp. 1–5.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.