Automatica 179 (2025) 112387

Contents lists available at ScienceDirect

Automatica

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

False Noise Attack Detection for differentially-private distributed control of microgrids☆

Feng Ye [a,b,c], Xianghui Cao [a,b,*], Lin Cai [c], Mo-Yuen Chow [d]

[a] School of Automation, Southeast University, Nanjing 210096, China
[b] Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education, Nanjing 210096, China
[c] Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 3P6, Canada
[d] UM-SJTU Joint Institute, Shanghai Jiao Tong University, Shanghai 200240, China

ARTICLE INFO

Article history:
Received 7 November 2023
Received in revised form 9 September 2024
Accepted 7 April 2025
Available online 28 May 2025

Keywords:
Distributed energy management
Differential privacy
False noise attack
Attack detection
State estimation
Information entropy

ABSTRACT

Privacy preserving in distributed control is getting more attention, and differential privacy (DP) is the common tool to protect data privacy, in which additive noise is applied in the algorithm function. However, DP can be leveraged by false noise (FN) attacks because attack vectors can be disguised as artificial noise in DP. FN attacks are a concern as the stealth attacks are hard to detect. Moreover, DP in distributed control makes FN attack detection more difficult. Hence, detecting FN attacks in privacy-preserving distributed control is critical and challenging. In this paper, taking distributed energy management systems as the control object, we propose a novel peer-to-peer attack detection approach, named False Noise Attack Detection (FNAD). In FNAD, each device observes the power decisions of its neighbors based on the data from its two-hop neighbors, estimates the power decisions of its neighbors by a Kalman filter, and updates the detection index of each neighbor according to the residues of the Kalman filter at each iteration. The detection index is developed based on information entropy, without any prior knowledge of the FN attacks. If a device's detection index is out of well-defined thresholds, its neighbors can perform a majority vote to decide whether it is malicious. We theoretically prove the detection effect of FNAD against three representative attacks in the literature and analyze the advantages of FNAD compared with the traditional methods. The effectiveness of FNAD is demonstrated by extensive simulations.

1. Introduction

With the rapid developments in science and technology in power system and automatic control fields, distributed control of microgrids has emerged as a novel form of power grid control in recent years due to its low costs of computation and communication (Zhu & Martinez, 2012). Usually, microgrids consist of renewable energy resources (RERs) including wind turbines and photovoltaic (PV) systems, distributed energy storage devices (DESDs), loads, etc. (Almehizia, Al-Masri, & Ehsani, 2019; Chen, Cai, & Su, 2023; Muhtadi, Pandit, Nguyen, & Mitra, 2021). As for the normal operation of the microgrid, economic dispatch (ED) is a fundamental need, which aims at meeting the power balance demand and operating at the minimum cost (Rahbari-Asr, Zhang, & Chow, 2015; Yang, Tan, & Xu, 2013). Distributed control of ED has been developed, forming consensus-based distributed energy management systems (DEMSs) (Ananduta & Ocampo-Martinez, 2021; Qin, Wan, Yu, Li, & Li, 2019; Wang, Li, Zhang, & Wang, 2019; Yang et al., 2013). Meanwhile, privacy preservation is critical in microgrids, as users do not want to disclose their private information to distrustful neighbors. Furthermore, privacy disclosure offers key data for potential cyber attacks. Even if there is no direct private data transmission in DEMSs, adversaries can infer the private information by continuous eavesdropping (Ye, Cao, Chow, & Cai, 2024; Ye, Cheng, Cao, & Chow, 2021; Zhao, Chen, He, & Cheng, 2018).

To preserve the private information in DEMSs, several privacy-preserving methods have been proposed, such as differential privacy (DP) and homomorphic encryption. Meanwhile, An, Duan, Chow, and Duel-Hallen (2019) proposed a resilient bargaining game for DEMSs, in which privacy-preserving is considered as a critical need. Among them, DP has received extensive attention. Zhao, Chen, et al. (2018) proposed two DP-based algorithms, where zero-sum and decaying noises are added to the algorithm. Nevertheless, since the above-mentioned privacy-preserving algorithms are designed based on adding noise, attackers can inject

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Shun-ichi Azuma under the direction of Editor Thomas Parisini.

* Corresponding author at: School of Automation, Southeast University, Nanjing 210096, China.

E-mail addresses: fye97@outlook.com (F. Ye), xhcao@seu.edu.cn (X. Cao), cai@ece.uvic.ca (L. Cai), moyuen.chow@sjtu.edu.cn (M.-Y. Chow).

https://doi.org/10.1016/j.automatica.2025.112387
0005-1098/

Nomenclature

Parameters

$\boldsymbol{I}$	Identity matrix
$\boldsymbol{O}$	Zero matrix
Δt	Time length of each step
$\mathcal{H}(\cdot)$	Information entropy of a random sequence
ρ	Penalty factor of Lagrange function
$\underline{E}_i$, $\overline{E}_i$	Lower and upper bounds of energy storage
$\underline{J}_k^{th}$, $\overline{J}_k^{th}$	Lower and upper thresholds at iteration k
$\underline{P}_i$, $\overline{P}_i$	Lower and upper bounds of P_i
$p(t)$	Unit electric price at time t
w_{ij}	Adjacency weight for devices i and j

Sets and Indices

$\mathcal{E}$	Set of communication edges
$\mathcal{V}$	Set of all devices
$\mathcal{V}_L$	Set of loads
$\mathcal{V}_G$	Set of main grid interface (MGI)
$\mathcal{V}_R$	Set of renewable energy resources (RERs)
$\mathcal{V}_S$	Set of distributed energy storage devices (DESDs)

Variables

$\hat{\boldsymbol{x}}_{i,k}^-$, $\hat{\boldsymbol{x}}_{i,k}$	Prior and posterior estimate of $\boldsymbol{x}_{i,k}(t)$
$\boldsymbol{K}_{i,k}$	Kalman gain
$\boldsymbol{P}_{i,k}^-$, $\boldsymbol{P}_{i,k}$	Prior and posterior estimate error covariance of $\hat{\boldsymbol{x}}_{i,k}$
$\boldsymbol{x}_{i,k}$	Vector form of $P_{i,k}(t)$ at all time steps
$\boldsymbol{y}_{i,k}$	Vector form of $\tilde{P}_{i,k}(t)$ at all time steps
$\hat{\lambda}_{i,k}(t)$	Incremental cost estimation by device i at time step t
$\underline{\hat{P}}_{i,k}(t)$, $\overline{\hat{P}}_{i,k}(t)$	Estimated lower and upper bounds of device i's power
$\hat{\xi}_{i,k}(t)$	Power imbalance estimation by device i at time step t
$\lambda(t)$	Incremental cost at time step t
$\mu_{1,i}(t)$, $\mu_{2,i}(t)$	Lagrange multipliers
$\tilde{P}_{i,k}(t)$	Observation of $P_{i,k}(t)$ at iteration k
$\xi(t)$	Global power imbalance at time step t
$E_i(t)$	Energy stored in i at the beginning of time step t
$J_{i,k}$	Detector index of device i at iteration k
$P_{i,k}(t)$	Power decision of device i at iteration k
$P_i(t)$	Power of device i at time step t

false noise (FN) instead of normal noise into privacy-preserving algorithms to cause outage or undermine the economics of microgrids (Li, Yang, Xia, & Dai, 2022; Wei, Wan, & He, 2020; Yang & Zhai, 2022, 2024; Ye, Cao, Cheng, & Chow, 2023). In other words, adding-noise privacy-preserving algorithms reduce the resilience to cyber attacks, which limits their application in microgrids.

An FN attack can be equivalent to a false data injection (FDI) attack in terms of attack effect, one of the most critical cyber attacks, in which attackers inject designed false data into the target and disturb its normal operation (Chen et al., 2021; Liang, Zhao, Luo, Weller, & Dong, 2017; Ye et al., 2023). Usually, FDI attacks are hard to detect given their stealthy property (Higgins, Teng, & Parisini, 2021). There have been several major FDI attacks in the world. For instance, in the 2015 Ukraine power grid hack, hackers used the BlackEnergy 3 malware to launch the FDI attack against the information systems of the power grid (Liang, Weller, Zhao, Luo, & Dong, 2017). Compared with FDI attacks that only malicious devices actively inject false data, all devices in privacy-preserving DEMSs inject noise to the original data, which could be seen as false data injection from the perspective of neighbors. The only difference is that benign devices generate noise according to predefined rules while malicious devices generate FN to disturb the normal operation of DEMSs. Hence, in this paper, FN attacks are applied to describe these attacks. Moreover, adding-noise privacy-preserving algorithms make FN attacks harder to detect, as noise in the DP-based algorithms has similar characteristic to attack vectors. Distinguishing FN from normal noise in DP-based algorithms is a critical problem. Hence, it is critical to develop a method that can effectively detect FN attacks in privacy-preserving distributed control of microgrids.

Since FN attacks are stealthy, they are hard to detect using traditional common detection methods. Among the traditional methods, the chi-square detector and Kullback–Leibler Divergence (KLD) are the main tools. The chi-square detector transforms the residue of the Kalman filter into a new variable and checks if the variable follows a chi-squared distribution at the chosen significance level. However, the chi-square detector only considers attack vectors with large values and has the weakness of a high false alarm rate, which makes it ineffective at detecting well-designed attack vectors. The KLD detector compares the probability density functions (PDFs) of the compromised and normal residues, and KLD would be positive if the PDF of the compromised residue is different from that of the normal residue (Guo, Shi, Johansson, & Shi, 2018). However, KLD needs sufficient statistical knowledge of attack vectors. In other words, KLD cannot detect FN attacks with zero prior knowledge, making it unsuitable for real-time dynamic systems such as microgrids.

In this paper, we consider the internal dynamic process in DP-DEMSs and propose a novel peer-to-peer (P2P) FN attack detection method, named False Noise Attack Detection (FNAD). The proposed detection algorithm consists of three phases. In the first phase, each neighbor of a device individually observes the power decisions of the target device based on data from its two-hop neighbors. In the second phase, using the observation results, each neighbor estimates the power decision by the Kalman filter. In the last phase, each neighbor evaluates the detection index of the target device using the parameter estimation results at each iteration, and all neighbors vote to decide if a device is malicious when the detection index is above a well-defined threshold. The detection index is designed based on information entropy, while FNAD directly processes the discrete-time residue sequence without prior knowledge of the FN attack.

The main contributions of this paper are three-fold.

- Considering DP-DEMS, we propose a novel FN attack detection method called FNAD, in which the neighbors estimate the power decisions of the target device, and update the detection index based on the residue of estimation results, where the detection index is computed without any prior information about the attack.
- We analyze the detection performance of FNAD based on information entropy theory and theoretically prove the detectability against three kinds of representative FN attacks in the literature.

- We analyze the advantages of FNAD over traditional detection methods, noting that FNAD can effectively detect FN attacks without prior knowledge and preserve the privacy of each device.

The rest of this paper is structured as follows. Section 3 introduces the DEMS, the distributed control of ED, and FN attack models. Section 4 proposes the FN attack detection method. Section 5 analyzes the detection performance of FNAD. Section 6 illustrates the performance of FNAD by case studies and Section 7 concludes this paper.

Notations: In this paper, $\mathbb{R}$ denotes the set of real numbers, and $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. For a random variable x, $x \sim \mathbb{N}(\mu, \sigma^2)$ means x is normally distributed with mean μ and variance σ^2, and $\mathbb{E}[x]$ denotes the expectation of the random variable x. $\mathbb{P}\{\Phi\}$ stands for the probability of the event Φ happens. $\text{sgn}(\cdot)$ denotes the sign function.

2. Related works

FDI attacks and their detection have been a hot topic, with several attack strategies and detection methods proposed.

Regarding FDI attacks in grids, Lakshminarayana, Kammoun, Debbah, and Poor (2021) proposed enhanced FDI attack vectors based on random matrix theory, which can bypass detection under conditions of limited measurements. Yang, He, Wang, Qiu, and Ai (2022) proposed a blind FDI attack strategy against the state estimation of power grids, which does not require any information about the grid parameters and topology. To bypass bad data detection, Tian et al. (2022) proposed an algorithm that adds noise to state variables, ensuring stealthiness against both traditional bad data detectors and deep learning-based detectors.

Since FDI attacks threaten the operation of grids, researchers have developed various detection methods and resilient control methods to mitigate these FDI attacks (Cui, Qu, Gao, Xie, & Yu, 2020; Musleh, Chen, & Dong, 2020). One main detection method in smart grids is called moving target defense (MTD). To evaluate the effectiveness and the stealth of MTD, Liu, Zhao, Zhang, and Deng (2022) derived explicit approximations of measurement residuals and designed an explicit residual-based MTD to make a trade-off between effectiveness and stealth. Xu, Jaimoukha, and Teng (2023) considered noises in the system and proposed robust MTD to guarantee the detection rate in the worst case. Besides MTD, other detection methods exist. Su, Li, Gao, Huang, and Li (2021) developed a robust sliding mode observer that generates residual signals to reconstruct and detect FDI attacks. In order to detect the FDI attacks as well as preserve privacy, Li, Wei, Li, Dong, and Shahidehpour (2022) proposed an algorithm integrating Transformer, federated learning, and homomorphic encryption, in which Transformer utilizes its multi-head self-attention mechanism to detect FDI attacks, while homomorphic encrypted federated learning ensures the privacy preservation. However, the aforementioned MTD methods are designed for centralized control architectures and cannot be applied to DEMSs. To detect FN in privacy-preserving discrete-time average consensus, He, Cai, Cheng, Pan, and Shi (2019) proposed a P2P neighbor monitoring detection method. Whereas, the internal operation mechanism of DEMSs is much more complex than that of discrete-time average consensus, and the method in He et al. (2019) cannot be directly applied in DEMSs.

There has been limited research on FDI attacks in DEMS. Duan and Chow (2019a) proposed an attack to gain extra benefits, where a DESD designs malicious data based on the number of neighbors and extra power charging or discharging, and broadcast the false data to its neighbors. Zhao, He, Cheng, and Chen (2017) analyzed the impact of different FDI attack vectors on DEMSs and identified the conditions under which these vectors do not affect the normal operation of microgrids. In Ye et al. (2023), a collusive attack was proposed to decrease the economics of DEMS, where a malicious DESD and a malicious load collusively design the extra power consumption. The attack in Ye et al. (2023) is in a stealthy manner as it does not increase communication data or times, and it is hard to be detected by neighbors of the attackers. As for the detection methods of FDI attacks in DEMS, the works in Duan and Chow (2019b) and Cheng and Chow (2020) proposed reputation-based detection approaches, where neighbors cross-check the correctness of transmitted data, and calculate the reputation index. If the reputation index of a device is below the threshold, neighbors determine if the device is malicious by major voting. However, the works in Duan and Chow (2019b) and Cheng and Chow (2020) do not account for the internal dynamic process of devices when updating power decisions and privacy preserving mechanisms in DEMSs. These drawbacks limit detection performance.

3. Problem setup

3.1. System model

Consider a microgrid consisting of the main grid interfaces (MGIs), distributed energy storage devices (DESDs), renewable energy resources (RERs) containing wind turbines and PV, and loads. The sets of all devices, MGIs, DESDs, RERs, and loads are denoted by $\mathcal{V}$, $\mathcal{V}_G$, $\mathcal{V}_S$, $\mathcal{V}_R$, and $\mathcal{V}_L$, respectively, and $\mathcal{V} = \mathcal{V}_G \cup \mathcal{V}_S \cup \mathcal{V}_R \cup \mathcal{V}_L$ holds. Specifically, RERs directly operate with their maximum power points, which is not adjustable. Thus, we call loads and RERs uncontrollable devices. The power usage of other devices is adjustable. Similarly, we call DESDs and GMIs controllable devices, and the sets of controllable and uncontrollable devices are denoted by $\mathcal{V}_C = \mathcal{V}_G \cup \mathcal{V}_S$ and $\mathcal{V}_U = \mathcal{V}_R \cup \mathcal{V}_L$, respectively. In this microgrid, let e_{ij} denote the communication link of devices i and j, where $e_{ij} = 1$ means device i can receive messages from device j, $e_{ij} = 0$ otherwise. All devices in the microgrid are connected and formed a network topology $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{E}$ represents the set of e_{ij}. In this paper, we assume that $e_{ij} = e_{ji}$, i.e., all communications link between two devices are bidirectional. Then it is obvious that $\mathcal{G}$ is an undirected graph. For device i, neighbors of device i are the devices that satisfy $e_{ij} = e_{ji} = 1$. Let $\mathcal{N}_i$ represent the set of device i's neighbors, and $d_i = |\mathcal{N}_i|$ denote the degree of device i.

3.2. Distributed power control for ED

For a microgrid, ED is a fundamental need for its normal operation. Consider a finite time horizon $\mathcal{T} = \{1, 2, \ldots, t_{\max}\}$ where $|\mathcal{T}| \geq 2$, and let $P_i(t)$ denote the purchased power of each device at time $t \in \mathcal{T}$. For the whole microgrid, costs are the electricity bought by GMIs. For GMIs, they buy electricity from the main grid if the power generation from RERs and power discharging from DESDs cannot satisfy the needs of the loads, and the cost of buying electricity at each time step is as follows:

$$C_i(t) = p(t)P_i(t)\Delta t, \tag{1}$$

where $C_i(t)$ denotes the cost of buying electricity by device i, $p(t)$ is the unit electricity price, Δt denotes the time interval of each time step. Hence, the global economic optimization problem over the finite time horizon $\mathcal{T}$ can be formulated as follows (Ye et al., 2021):

$$\min_{\{P_i(t), i \in \mathcal{V}_C\}} \quad \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{V}_G} C_i(t), \tag{2}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{V}_L} P_i(t) = \sum_{i\in\mathcal{V}\backslash\mathcal{V}_L} P_i(t), \tag{3}$$

$$\underline{P}_i \le P_i(t) \le \overline{P}_i, \ \forall i \in \mathcal{V}\backslash\mathcal{V}_R \tag{4}$$

$$E_i(t+1) = E_i(t) - P_i(t)\Delta t, \ \forall i \in \mathcal{V}_S \tag{5}$$

$$\underline{E}_i \le E_i(t) \le \overline{E}_i, \ \forall i \in \mathcal{V}_S \tag{6}$$

where $\underline{P}_i$ and $\overline{P}_i$ denote device i's lower and upper bounds of power, $\underline{E}_i$ and $\overline{E}_i$ denote device i's lower and upper bounds of energy storage, $E_i(t)$ denotes the energy storage at the beginning of time step t. The above problem (2)–(6) is usually solved by constructing the augmented Lagrange function and gradient descent, where the Lagrange Multiplier for constraint (3), (6) is defined by $\lambda(t)$, $\mu_1(t)$, $\mu_2(t)$, respectively. Meanwhile, the penalty factor ρ is utilized to accelerate the convergence rate.

Given the communication and computation constraints, we apply the distributed ED (2)–(6) by consensus-based distributed control. At each iteration, each device estimates the global power imbalance $\xi(t)$ and the Lagrange Multiplier $\lambda(t)$ (which is also named incremental cost), where

$$\xi(t) = \sum_{i\in\mathcal{V}_L} P_i(t) - \sum_{i\in\mathcal{V}\backslash\mathcal{V}_L} P_i(t), \tag{7}$$

because $\xi(t)$ and $\lambda(t)$ cannot be directly obtained by each device. The estimations of $\xi(t)$ and $\lambda(t)$ of device i at iteration k are denoted by $\hat{\xi}i,k(t)$ and $\hat{\lambda}i,k(t)$, respectively. Meanwhile, each controllable device makes its local power decision $P_{i,k}(t)$ based on $\hat{\xi}i,k(t)$ and $\hat{\lambda}i,k(t)$. By updating $\hat{\xi}i,k(t)$, $\hat{\lambda}i,k(t)$, and $P_{i,k}(t)$, the distributed algorithm converges to the optimal solution, and each device obtains the optimal operation point.

On the other hand, DP is a wildly applied measurement for privacy preservation. The definition of DP is introduced as follows.

Definition 1 (*(ϵ, δ)-DP*)**.** Considering two different dataset $\mathcal{D}$ and $\mathcal{D}'$, in which only one element is different. With regard to a privacy-preserving algorithm $\mathcal{A}$ and output set $\mathcal{O} \subseteq \text{Ra}(\mathcal{A})$, where $\text{Ra}(\mathcal{A})$ is the domain of the output under mechanism $\mathcal{A}$, if and only if

$$\mathbb{P}\{\mathcal{A}(\mathcal{D}) \in \mathcal{O}\} \le e^{\epsilon}\mathbb{P}\left\{\mathcal{A}(\mathcal{D}') \in \mathcal{O}\right\} + \delta \tag{8}$$

holds, then the algorithm $f(\cdot)$ satisfies (ϵ, δ)-DP.

In practice, the form of the privacy-preserving algorithm $\mathcal{A}$ would be $\mathcal{A}(x) = x + \theta$, where x denotes the original data, and θ stands for artificial noises for privacy-preserving purposes. In DEMSs, the artificial noises are added to the transferred data to preserve privacy of power information. The detailed updating rules are as follows (Ye et al., 2021; Zhao, Chen, et al., 2018):

Power imbalance estimation: Each device coordinates with its neighbors and updates the power imbalance estimation by

$$\hat{\xi}^{+}_{i,k}(t) = \hat{\xi}_{i,k}(t) + \psi_{i,k}(t), \tag{9}$$

$$\hat{\xi}_{i,k+1}(t) = w_{ii}\hat{\xi}^{+}_{i,k}(t) + \sum_{j\in\mathcal{N}_i} w_{ij}\hat{\xi}^{+}_{j,k}(t) + \begin{cases} P_{i,k+1}(t) - P_{i,k}(t), & \text{if } i \in \mathcal{V}_L \\ P_{i,k}(t) - P_{i,k+1}(t), & \text{otherwise,} \end{cases} \tag{10}$$

where $\hat{\xi}_{i,0}(t) = P_{i,0}(t)$ for all loads and $\hat{\xi}_{i,0}(t) = -P_{i,0}(t)$ for other devices, w_{ij} is the adjacency weight (Ye et al., 2021), and $\psi_{i,k}(t)$ denotes the artificial noise. $\psi_{i,k}(t)$ is generated by

$$\psi_{i,k}(t) = \begin{cases} \theta_{i,0}(t), & k = 0 \\ \rho_1^k\theta_{i,k}(t) - \rho_1^{k-1}\theta_{i,k-1}(t), & k > 0, \end{cases} \tag{11}$$

where $\theta_{i,k}(t)$ is a normally distributed random variable $\boldsymbol{\theta}_{i,k} \sim \mathbb{N}(\mathbf{0}, \boldsymbol{R}_1)$, $\boldsymbol{\theta}_{i,k} \in \mathbb{R}^{|\mathcal{T}|}$ denotes the vector form of $\theta_{i,k}(t)$ at all time steps, $\mathbf{0} \in \mathbb{R}^{|\mathcal{T}|}$ is the vector with all elements 0, $\boldsymbol{R}_1$ is the covariance of $\boldsymbol{\theta}_{i,k}$, which is a diagonal matrix, and $0 < \rho_1 < 1$ is a decaying factor.

Incremental cost estimation: Each device coordinates with its neighbors and updates the incremental cost estimation by

$$\hat{\lambda}^{+}_{i,k}(t) = \hat{\lambda}_{i,k}(t) + \varphi_{i,k}(t), \tag{12}$$

$$\hat{\lambda}_{i,k+1}(t) = w_{ii}\hat{\lambda}^{+}_{i,k}(t) + \sum_{j\in\mathcal{N}_i} w_{ij}\hat{\lambda}^{+}_{j,k}(t) + \eta\hat{\xi}_{i,k}(t), \tag{13}$$

where η is a constant coefficient, $\hat{\lambda}_{i,0}(t) = 0$, and $\varphi_{i,k}(t)$ denotes the artificial noise. Similarly, $\varphi_{i,k}(t)$ is generated by

$$\varphi_{i,k}(t) = \begin{cases} \vartheta_{i,0}(t), & k = 0 \\ \rho_2^k\vartheta_{i,k}(t) - \rho_2^{k-1}\vartheta_{i,k-1}(t), & k > 0, \end{cases} \tag{14}$$

where $\vartheta_{i,k}(t)$ is also a normally distributed random variable $\boldsymbol{\vartheta}_{i,k} \sim \mathbb{N}(\mathbf{0}, \boldsymbol{R}_2)$, $\boldsymbol{R}_2$ is the covariance of $\boldsymbol{\vartheta}_{i,k} \in \mathbb{R}^{|\mathcal{T}|}$, which is also a diagonal matrix, and $0 < \rho_2 < 1$ is a decaying factor.

Power decision: Each controllable device updates the local power decision as follows.

- $\forall i \in \mathcal{V}_l$, $P_{i,k+1}(t)$ is updated by

$$P_{i,k+1}(t) = \left[P_{i,k}(t) - \eta\left(p(t)\Delta t - \hat{\lambda}_{i,k}(t) - \rho\hat{\xi}_{i,k}(t)\right)\right]^{\overline{P}_i}_{\underline{P}_i}, \tag{15}$$

  where for real numbers c, $\underline{c}$, and $\overline{c}$, $[c]^{\overline{c}}_{\underline{c}} \triangleq \min\{\overline{c}, \max\{\underline{c}, c\}\}$.
- $\forall i \in \mathcal{V}_S$, $P_{i,k+1}(t)$ is updated by equation in Box I with

$$\mu_{1,i,k+1}(t) = \left[\mu_{1,i,k}(t) + \eta\sum_{s=1}^{t} P_{i,k}(s)\Delta t + \eta\left(\underline{E}_i - E_i(1)\right)\right]^{\infty}_{0}, \tag{17}$$

$$\mu_{2,i,k+1}(t) = \left[\mu_{2,i,k}(t) - \eta\sum_{s=1}^{t} P_{i,k}(s)\Delta t - \eta\left(E_i(1) - \overline{E}_i\right)\right]^{\infty}_{0}, \tag{18}$$

  where $\mu_{1,i,k}(t)$ and $\mu_{2,i,k}(t)$ are the tentative results of $\mu_{1,i}(t)$ and $\mu_{2,i}(t)$ at iteration k, respectively.

Lemma 1 (*Zhao, Chen, et al., 2018*)**.** *The privacy-preserving DEMS converges to its optimal operation point:*

$$\lim_{k\to\infty} \Delta\hat{P}_{i,k}(t) = 0, \ \lim_{k\to\infty} \hat{\lambda}_{i,k}(t) = \lambda^*(t),$$

$$\lim_{k\to\infty} P_{i,k}(t) = P_i^*(t), \tag{19}$$

where $\lambda^*(t)$ and $P_i^*(t)$ are the optimal solutions to problem (2)–(6).

3.3. FN attack models

FN attacks in DEMS can be divided into three categories with respect to the attack targets: attack against the estimate of the incremental cost, the estimate of the power imbalance, and power decision.

In the estimate of power imbalance of P2P detection, attacks against the estimate of incremental cost are easy to be detected by (13). Hence, most attacks aim at the estimate of power imbalance and power decisions. The three typical attacks are summarized as follows:

Attack 1 For the power imbalance estimation process, FN is injected into $\hat{\xi}_{i,k}(t)$ at each iteration, i.e.,

$$\hat{\xi}^{a}_{i,k}(t) = \hat{\xi}_{i,k}(t) + v^{\xi}_{i,k}(t), \tag{20}$$

where $\hat{\xi}^{a}_{i,k}(t)$ represents the false estimate of power imbalance, $v^{\xi}_{i,k}(t)$ denotes the injected FN to $\hat{\xi}_{i,k}(t)$, which is zero-sum and can cause DEMSs deviate from the optimal operation points.

As for Attack 1, if $v^{\xi}_{i,k}(t)$ is not zero-sum, the sum of power imbalance estimation would not be zero-sum either. Since the

$$P_{i,k+1}(t)=\begin{bmatrix} P_{i,k}(t)+\eta\left(\hat{\lambda}_{i,k}(t)+\rho\hat{\xi}_{i,k}(t)\right)-\eta\Delta t\sum_{l=t}^{t_{\max}}\rho\left[\underline{E}_i-E_i(1)+\sum_{s=1}^{l}P_{i,k}(s)\Delta t\right]_0^{\infty} \\ +\eta\Delta t\sum_{l=t}^{t_{\max}}\left(\mu_{2,i,k}(l)-\mu_{1,i,k}(l)\right)+\eta\Delta t\sum_{l=t}^{t_{\max}}\rho\left[E_i(1)-\overline{E}_i-\sum_{s=1}^{l}P_{i,k}(s)\Delta t\right]_0^{\infty} \end{bmatrix}_{\underline{P}_i}^{\overline{P}_i} \tag{16}$$

Box I.

sum of power imbalance estimation is equal to the global power imbalance (Duan & Chow, 2019b), Attack 1 can directly destabilize DEMSs by causing frequency fluctuations in AC microgrids. If $v_{i,k}^{\xi}(t)$ is zero-sum, it may lead to suboptimal operation, i.e., undermining the economics of DEMSs. Usually, $v_{i,k}^{\xi}(t)$ is well-designed to avoid power imbalance in the microgrid and bypass the physical detection based on power imbalance because the existing physical detection methods are typically very effective to detect such attacks.

For the power decision process, there are two different ways to launch FN attacks:

Attack 2 The attackers design and inject the FN $v_{i,k}^{P}(t)$ to $P_{i,k}(t)$:

$$P_{i,k}^{a}(t)=P_{i,k}(t)+v_{i,k}^{P}(t), \tag{21}$$

where $P_{i,k}^{a}(t)$ denotes the false power decision, and $P_{i,k}^{a}(t)$ substitutes for $P_{i,k}(t)$ in the power imbalance estimation update (10). $v_{i,k}^{P}(t)$ can cause DEMSs deviate from normal operation points. Note that even if uncontrollable devices do not change their power decisions, they can still launch Attack 2.

Attack 3 The attackers tamper with the parameters including $p(t)$, $E_i(1)$, $\underline{E}_i$, $\overline{E}_i$.

As for Attacks 2 and 3, if the attack vectors of Attacks 2 and 3 are time-invariant, the DEMS may operate with the suboptimal point; while if the attack vectors are time-variant, the DEMS may not converge,i.e., causing outage of the microgrid.

3.4. Problem statement

Consider the privacy-preserving DEMS introduced above, where at least one malicious device launches an FDI attack and undermines the economics of the DEMS in a stealthy manner. Even though several P2P detection methods such as the algorithm in Duan and Chow (2019b) where devices cross-validate the correctness of neighbors' information via two-hop neighbors' data, some attacks can bypass the detection. Moreover, the FN has similar characteristics to artificial noise for privacy preservation, which increases the difficulty of distinguishing between benign and malicious noise. Hence, we need to develop a novel detection method to fill the gap.

4. The proposed detection method

In this section, we introduce the proposed FNAD algorithm in detail. The basic idea is that each device collects two-hop neighbors' information and estimates neighbors' real-time power decisions. Based on observed power decisions with noise, devices persistently estimate neighbors' power decisions. Hence, the FNAD algorithm can detect whether misbehavior happens by the estimation results. The entire FNAD process can be divided into three parts: P2P monitoring and power decision observation, power decision estimation, and attack detection. Power decision observation is based on P2P monitoring and data forwarding, and the observation results are used in the power decision estimation that is realized by Kalman filter. Then we design the detection index and detect potential FN attacks according to the residue of Kalman filter.

4.1. P2P monitoring and power decision observation

For the DEMS, since there is no centralized controller in the microgrid, attack detection must rely on P2P monitoring. Each device in the microgrid can only obtain power imbalance estimation and incremental cost estimation data from its neighbors. Therefore, $\hat{\xi}i,k(t)$ and $\hat{\lambda}i,k(t)$ are the only clues for P2P monitoring and attack detection. However, $\hat{\xi}i,k(t)$ and $\hat{\lambda}i,k(t)$ from one-hop neighbors are insufficient for devices to detect misbehavior. Hence, we make the following assumptions.

Assumption 1. Each device can obtain the transmitted data $\hat{\xi}_{i,k}(t)$ and $\hat{\lambda}_{i,k}(t)$ of two-hop neighbors by data forwarding.

For DEMSs, although data forwarding increases communication costs, it can provide sufficient information for P2P monitoring. If a device tampers with the actual data of two-hop neighbors, it is more likely to be suspected during the state and parameter estimation phase. The reasons are as follows: it is difficult for attackers to design FN that both launch FN attacks and bypass detection when forwarding information of two-hop neighbors. Additionally, there may be cases where three or more devices are common neighbors to each other. This compels the devices to follow the data forwarding rule, which contributes to the stable operation of DEMSs. Note that having only one neighbor should be avoided; if the single neighbor is malicious and tampers with the forwarding data, it is difficult for the remaining normal devices to cross-check the data integrity.

Assumption 2. Each device shares its internal parameters of its neighbors including $p(t)$, $E_i(1)$, $\underline{E}_i$, $\overline{E}_i$.

In Assumption 2, sharing internal parameters is feasible because DP can preserve the power decision information. For each device, even if its neighbors know the internal parameters, they cannot deduce the actual real-time operation states of the target device.

The above assumption forms the foundation for power decision observation. Under Assumption 1, if devices forward received messages from their neighbors at each iteration, then, according to He, Cai, and Guan (2018), the neighbors of device $i\in\mathcal{V}\backslash\mathcal{V}_L$ can observe device i's power decision at iteration $k>0$ by

$$\begin{aligned}\tilde{P}_{i,k}(t)=&\sum_{h=0}^{k-1}\sum_{j\in\mathcal{N}_i\cup\{i\}}w_{ij}\hat{\xi}_{j,h}^{+}(t)-\sum_{h=0}^{k}\hat{\xi}_{i,h}^{+}(t)\\ =&\tilde{P}_{i,k-1}(t)+\sum_{j\in\mathcal{N}_i\cup\{i\}}w_{ij}\hat{\xi}_{j,k-1}^{+}(t)-\hat{\xi}_{i,k}^{+}(t),\end{aligned} \tag{22}$$

where $\tilde{P}_{i,k}(t)$ is the observation of $P_{i,k}(t)$. Similarly, if the device $i\in\mathcal{V}_L$, its neighbors can observe its power decision at iteration $k>0$ by

$$\begin{aligned}\tilde{P}_{i,k}(t)=&\sum_{h=0}^{k}\hat{\xi}_{i,h}^{+}(t)-\sum_{h=0}^{k-1}\sum_{j\in\mathcal{N}_i\cup\{i\}}w_{ij}\hat{\xi}_{j,h}^{+}(t)\\ =&\tilde{P}_{i,k-1}(t)-\sum_{j\in\mathcal{N}_i\cup\{i\}}w_{ij}\hat{\xi}_{j,k-1}^{+}(t)+\hat{\xi}_{i,k}^{+}(t).\end{aligned} \tag{23}$$

Lemma 2. *For each target device, the power decision observation method (22)–(23) can accurately observe its actual convergence power decision if the target device obeys the updating rule of (10).*

Proof. See Appendix A.

4.2. State estimation for local power decisions

Based on the above power decision observation, each device estimates the states of neighbors in (15)–(18), containing the power decisions and the lower and upper bounds of power.

Here we assume that the devices do not reach any boundary conditions, i.e., $\underline{P}_i < P_{i,k}(t) < \overline{P}_i$ for each device, $\mu_{1,i,k}(t) = \mu_{2,i,k}(t) = 0$ for each DESD. According to (15)–(18), we have the updating rules of different devices' power decisions as follows:

$$\begin{aligned} P_{i,k+1}(t) &= P_{i,k}(t) \\ &+\eta \begin{cases} \hat{\lambda}_{i,k}(t)+\rho\hat{\xi}_{i,k}(t)-p(t)\Delta t, & \text{if } i \in \mathcal{V}_G \\ \hat{\lambda}_{i,k}(t)+\rho\hat{\xi}_{i,k}(t)+f_i(P_{i,k}(t))-g_i(P_{i,k}(t)), & \text{if } i \in \mathcal{V}_S \\ 0, & \text{if } i \in \mathcal{V}_U, \end{cases} \end{aligned} \tag{24}$$

$$f_i(P_{i,k}(t)) \triangleq \Delta t \sum_{l=t}^{t_{\max}} \rho \left[E_i(1) - \overline{E}_i - \sum_{s=1}^{l} P_{i,k}(s)\Delta t \right]_0^{\infty},$$

$$g_i(P_{i,k}(t)) \triangleq \Delta t \sum_{l=t}^{t_{\max}} \rho \left[\underline{E}_i - E_i(1) + \sum_{s=1}^{l} P_{i,k}(s)\Delta t \right]_0^{\infty}.$$

By rewriting and substituting (10) into (22), we can derive the observation of power decision by

$$\tilde{P}_{i,k}(t) = P_{i,k}(t) + \text{sgn}(\tau)\rho_1^k \theta_{i,k}(t), \tag{25}$$

where

$$\tau = \begin{cases} 1, & \text{if } i \in \mathcal{V}_L \\ -1, & \text{otherwise.} \end{cases} \tag{26}$$

We can rewrite (25) into the vector form as follows:

$$\boldsymbol{y}_{i,k} = \boldsymbol{x}_{i,k} + \text{sgn}(\tau)\rho_1^k \boldsymbol{\theta}_{i,k}, \tag{27}$$

where $\boldsymbol{x}_{i,k} \in \mathbb{R}^{|\mathcal{T}|}$ and $\boldsymbol{y}_{i,k} \in \mathbb{R}^{|\mathcal{T}|}$ denote the vector form of $P_{i,k}(t)$ and $\tilde{P}_{i,k}(t)$ at all time steps, respectively. Let $\boldsymbol{R}_{i,k} \in \mathbb{R}^{|\mathcal{T}|\times|\mathcal{T}|}$ denote the covariance of noise $\rho_1^k \boldsymbol{\theta}_{i,k}$, and $\boldsymbol{R}_{i,k} = \rho_1^{2k}\boldsymbol{R}_1$ holds.

Then we can utilize the Kalman filter to estimate the power decisions. For ease of describing the detection process, we rewrite all variables as vector or matrix forms: let $\hat{\boldsymbol{\xi}}_{i,k} \in \mathbb{R}^{|\mathcal{T}|}$, $\hat{\boldsymbol{\lambda}}_{i,k} \in \mathbb{R}^{|\mathcal{T}|}$, $\boldsymbol{f}_i(\cdot)$, $\boldsymbol{g}_i(\cdot)$, and $\boldsymbol{p} \in \mathbb{R}^{|\mathcal{T}|}$ denote the vectors of $\hat{\xi}_{i,k}(t)$, $\hat{\lambda}_{i,k}(t)$, $f_i(\cdot)$, $g_i(\cdot)$, and $p(t)$ at all time steps, respectively. Based on Assumption 2, Lemma 2, and (24)–(27), by using the Kalman filter, we have the states estimate updating rule:

$$\begin{aligned} \hat{\boldsymbol{x}}_{i,k}^- &= \hat{\boldsymbol{x}}_{i,k-1} \\ &+\eta \begin{cases} \hat{\boldsymbol{\lambda}}_{i,k-1}+\rho\hat{\boldsymbol{\xi}}_{i,k-1}-\boldsymbol{p}\Delta t, & \text{if } i \in \mathcal{V}_G \\ \hat{\boldsymbol{\lambda}}_{i,k-1}+\rho\hat{\boldsymbol{\xi}}_{i,k-1}+\boldsymbol{f}_i(\hat{\boldsymbol{x}}_{i,k-1})-\boldsymbol{g}_i(\hat{\boldsymbol{x}}_{i,k-1}), & \text{if } i \in \mathcal{V}_S \\ \boldsymbol{0}, & \text{if } i \in \mathcal{V}_U, \end{cases} \end{aligned} \tag{28}$$

$$\boldsymbol{P}_{i,k}^- = \boldsymbol{P}_{i,k-1}, \tag{29}$$

$$\boldsymbol{K}_{i,k} = \boldsymbol{P}_{i,k}^-(\boldsymbol{P}_{i,k}^- + \boldsymbol{R}_{i,k})^{-1}, \tag{30}$$

$$\hat{\boldsymbol{x}}_{i,k} = \hat{\boldsymbol{x}}_{i,k}^- + \boldsymbol{K}_{i,k}(\boldsymbol{y}_{i,k} - \hat{\boldsymbol{x}}_{i,k}^-), \tag{31}$$

$$\boldsymbol{P}_{i,k} = (\boldsymbol{I} - \boldsymbol{K}_{i,k})\boldsymbol{P}_{i,k}^-, \tag{32}$$

where $\hat{\boldsymbol{x}}_{i,k}^-$ and $\hat{\boldsymbol{x}}_{i,k}$ denote the prior and posterior estimate of $\boldsymbol{x}_{i,k}(t)$, respectively, $\boldsymbol{f}_i(\cdot)$ and $\boldsymbol{g}_i(\cdot)$ denote the vector form of $f_i(\cdot)$ and $g_i(\cdot)$, respectively, $\boldsymbol{I}$ denotes the identity matrix, $\boldsymbol{P}_{i,k}^- \in \mathbb{R}^{|\mathcal{T}|\times|\mathcal{T}|}$ and $\boldsymbol{P}_{i,k} \in \mathbb{R}^{|\mathcal{T}|\times|\mathcal{T}|}$ denote the prior and posterior estimate error covariance of $\hat{\boldsymbol{x}}_{i,k}$, respectively, and $\boldsymbol{K}_{i,k} \in \mathbb{R}^{|\mathcal{T}|\times|\mathcal{T}|}$ represents the Kalman gain. For the initial value of the Kalman filter, $\hat{\boldsymbol{x}}_{i,0}^- = \boldsymbol{0}$, $\boldsymbol{P}_{i,0}^- = \mathbb{E}[\boldsymbol{x}_{i,0}\boldsymbol{x}_{i,0}^T]$. Here we denote

$$\begin{aligned} k_{i,1} \triangleq \min_k \big\{ &\boldsymbol{f}_i(\hat{\boldsymbol{x}}_{i,k}) + \boldsymbol{g}_i(\hat{\boldsymbol{x}}_{i,k}) \neq \boldsymbol{0} | \\ &\boldsymbol{f}_i(\hat{\boldsymbol{x}}_{i,k-1}) + \boldsymbol{g}_i(\hat{\boldsymbol{x}}_{i,k-1}) = \boldsymbol{0} \big\}, \end{aligned} \tag{33}$$

which is positive, and means the iteration that the ith DESD reaches the energy storage limit (6).

Remark 1. The Kalman filter is suitable for the estimation of power decisions. The system model (24) does have process noise, which can be viewed as the special case of standard Kalman filter, i.e., the covariance of process noise is a zero matrix. The observation noise in (27) is $\rho_1^k \boldsymbol{\theta}_{i,k}$, which is zero-mean Gaussian distributed. Also, $\rho_1^k \boldsymbol{\theta}_{i,k}$ is not correlated at different iterations. Thus, the system model and observation model satisfy the requirements of the Kalman filter, and the estimation results are unbiased. While if the observation noise is non-Gaussian distributed, the Kalman filter cannot be the optimal state estimator. Hence, zero-mean Gaussian distribution of the observation noise is a necessary condition of valid use of the proposed algorithm.

Remark 2. Note that $k_{i,1}$ only exists when the scheduled energy storage $E_i(t)$ is outside the normal range, i.e., $\exists t \in \mathcal{T}$, $E_i(t) > \overline{E}_i$ or $E_i(t) < \underline{E}_i$. For each DESD, if $k_{i,1}$ does not exist, the proposed detection algorithm can normally work. While if $k_{i,1}$ exists, the estimation (28)–(32) can be only used at iterations $k \leq k_{i,1}$ for DESDs. This is because $\mu_{1,i,k}(t)$ and $\mu_{2,i,k}(t)$ are equal to 0 and (28)–(32) hold at iterations $k \leq k_{i,1}$; while for iterations $k > k_{i,1}$, $\mu_{1,i,k}(t)$ and $\mu_{2,i,k}(t)$ are greater than 0 and (28)–(32) no longer hold.

Lemma 3. *If the target device is normal, then the state and parameter estimation processes by Kalman filter converge, and*

$$\lim_{k\to\infty} \boldsymbol{K}_{i,k} = \boldsymbol{I}, \ \lim_{k\to\infty} \boldsymbol{P}_{i,k} = \boldsymbol{O}, \tag{34}$$

where $\boldsymbol{I}$ denotes the identity matrix, $\boldsymbol{O}$ denotes the zero matrix, and the dimensions of the two matrices are $|\boldsymbol{x}_i| \times |\boldsymbol{x}_i|$.

Proof. See Appendix B.

4.3. Attack detection method

In this subsection, we introduce the design of detection index and the FN attack detection process.

4.3.1. Design of detection index

Based on the Kalman filter, we can construct the *residual*:

$$\boldsymbol{z}_{i,k} = \boldsymbol{y}_{i,k} - \hat{\boldsymbol{x}}_{i,k}^-, \tag{35}$$

where $\boldsymbol{z}_{i,k}$ denotes the residual, and $\boldsymbol{z}_{i,k}$ follows the normal distribution, i.e., $\boldsymbol{z}_{i,k} \sim \mathbb{N}(0, \boldsymbol{Q}_{z_{i,k}})$, and

$$\boldsymbol{Q}_{z_{i,k}} \triangleq \boldsymbol{P}_{i,k}^- + \boldsymbol{R}_{i,k} \tag{36}$$

denotes the covariance of $\boldsymbol{z}_{i,k}$. Then we construct a residual-based variable

$$\nu_{i,k} = \boldsymbol{z}_{i,k}^T \boldsymbol{Q}_{z_{i,k}}^{-1} \boldsymbol{z}_{i,k}, \tag{37}$$

which is χ^2 distributed with $|\mathcal{T}|$ freedom degrees. Then we propose an information entropy-based method to detect the potential FN attacks, where the information entropy of a random variable and KLD of two random variables are calculated as follows:

Definition 2 (*Information Entropy*)**.** The entropy $\mathcal{H}(X)$ of a discrete random variable X is defined by

$$\mathcal{H}(X) = -\sum_{x\in X} p(x)\ln p(x), \tag{38}$$

where $p(x)$ denotes the probability of event x, ln denotes the logarithms to base natural constant e, and we define $0\ln 0 = 0$.

Definition 3 (*KLD*)**.** Assume there are two random sequences x_k and y_k with PDFs f_{x_k} and f_{y_k}, respectively. The KLD between the two random sequences is denoted as

$$D(x_k|y_k) \triangleq \int_{\{\zeta_k|f_{x_k}(\zeta_k)>0\}} f_{x_k}(\zeta_k)\ln\frac{f_{x_k}(\zeta_k)}{f_{y_k}(\zeta_k)}\,d\zeta_k. \tag{39}$$

It is easy to obtain that $D(x_k|y_k) = 0$ if and only if $f_{x_k} = f_{y_k}$. Moreover, usually $D(x_k|y_k) \neq D(y_k|x_k)$.

Let $\boldsymbol{z}^a_{i,k}$ and $\nu^a_{i,k}$ denote the compromised residue and corresponding χ^2 distributed variable in accordance with (37). If $D(\nu^a_{i,k}|\nu_{i,k})$ exceeds the threshold, then the FN attack can be detected. However, we only know the PDFs of $\boldsymbol{z}_{i,k}$ and $\nu_{i,k}$ while do not know the PDFs of $\boldsymbol{z}^a_{i,k}$ and $\nu^a_{i,k}$, which means Definition 3 cannot be utilized directly. Here we assume that the random sequence $\{\nu^a_{i,h}\}^k_{h=0}$ as an independent and identically distributed (i.i.d.) random process, and $\{\nu^a_{i,h}\}^k_{h=0}$ is generated by a random variable ν^a_i. Then we consider using the information entropy of the random variable ν^a_i to replace the PDF of $\nu^a_{i,k}$ in the KLD. For ease of description, we let $\nu^a_{i,N_1,k}$ denote the sequence $\{\nu^a_{i,h}\}^k_{h=N_1}$. Based on the random sequence $\nu^a_{i,N_1,k}$, we divide the interval $[0, \max\{\nu^a_{i,N_1,k}\}]$ into $\max\{\nu^a_{i,N_1,k}\}/q$ equal parts, where q is a positive constant, and calculate of proportion in each part to substitute $p(x)$ in Definition 2. Note that the selection of q relies on k, i.e., m increases as k increases. Then according to Definition 2, we can compute the entropy $\mathcal{H}(\nu^a_{i,N_1,k})$. Then we design the novel detector as follows:

$$J_{i,k} = -\mathcal{H}(\nu^a_{i,N_1,k}) + \frac{1}{2}\bar{\nu}^a_{i,N_1,k} - \left(\frac{m}{2}-1\right)\overline{\ln\nu^a_{i,N_1,k}}, \tag{40}$$

$$\underline{J}^{th}_k \underset{H_1}{\overset{H_0}{\lessgtr}} J_{i,k} \underset{H_1}{\overset{H_0}{\lessgtr}} \bar{J}^{th}_k, \tag{41}$$

where

$$\bar{\nu}^a_{i,N_1,k} \triangleq \frac{1}{k-N_1+1}\sum_{h=N_1}^{k}\nu^a_{i,h}, \tag{42}$$

$$\overline{\ln\nu^a_{i,N_1,k}} \triangleq \frac{1}{k-N_1+1}\sum_{h=N_1}^{k}\ln\nu^a_{i,h}, \tag{43}$$

H_0 and H_1 are the two hypotheses that the target device is normal or compromised, and $\underline{J}^{th}_k$ and $\bar{J}^{th}_k$ are the lower and upper thresholds at iteration k.

Lemma 4. *The lower and upper thresholds $\underline{J}^{th}_k$ and $\bar{J}^{th}_k$ converge to $-\ln\left[2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)\right]$, i.e.,*

$$\lim_{k\to\infty}\underline{J}^{th}_k = \lim_{k\to\infty}\bar{J}^{th}_k = -\ln\left[2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)\right], \tag{44}$$

where $\Gamma(\cdot)$ denotes the Gamma function.

Proof. See Appendix C.

While the detector (38)–(43) should be used when k is not too small or too large. If k is too small, the samples of sequence $\nu^a_{i,N_1,k}$ are insufficient, and $\mathcal{H}(\nu^a_{i,N_1,k})$ cannot represent the entropy of random variable ν^a_i well. While when k goes to infinity, $\boldsymbol{Q}_{z_{i,k}}$ is a singular matrix in accordance with Lemma 3. Also, the original privacy-preserving DEMS algorithm operates within finite-time iterations and converges when the errors are below the predefined threshold in practice. Hence, we require that the detector should be used when $k \geq N_1$ is finite in practice. Meanwhile, to collect sufficient samples, we require that $k \geq N_1$, where N_1 is integer. Moreover, the threshold pair $\underline{J}^{th}_k$ and $\bar{J}^{th}_k$ is defined as

$$\sigma_{i,k} = \mathbb{P}\{J_{i,k} < \underline{J}^{th}_k \text{ or } J_{i,k} > \bar{J}^{th}_k | H_0\} < \sigma_{\max} \tag{45}$$

where $\sigma_{i,k}$ represents the false alarm rate at each iteration, and $\sigma_{\max}$ is the upper bound of the given false alarm rate. Hence, we can determine $\underline{J}^{th}_k$ and $\bar{J}^{th}_k$ by the Monte Carlo method to balance the detection success rate and false alarm rate.

4.3.2. FN attack detection algorithm

Usually, if $J_{i,k} < \underline{J}^{th}_k$ or $J_{i,k} > \bar{J}^{th}_k$, the target device may be considered malicious. However, two special cases that need to be discussed. For each device, if $\bar{J}^{th}_k \leq J_{i,k} \leq \underline{J}^{th}_k$, it can be easily judged that this device is normal. On the other hand, if $J_{i,k} < \underline{J}^{th}_k$ or $J_{i,k} > \bar{J}^{th}_k$, we should consider the cases of $\underline{P}_i = \hat{P}_{i,k+1}(t)$ or $\hat{P}_{i,k+1}(t) = \bar{P}_i$ for each device, and $\mu_{1,i,k}(t) \neq 0$ or $\mu_{2,i,k}(t) \neq 0$ for each DESD. The neighbors estimate the lower and upper bounds of power via

$$\bar{\hat{P}}_{i,k}(t) = \max\{\hat{P}_{i,h}(t),\, 0 \leq h \leq k\}, \tag{46}$$

$$\underline{\hat{P}}_{i,k}(t) = \min\{\hat{P}_{i,h}(t),\, 0 \leq h \leq k\}, \tag{47}$$

where $\underline{\hat{P}}_{i,k}(t)$ and $\bar{\hat{P}}_{i,k}(t)$ are the estimated lower and upper bounds of device i's power, respectively. If device i is a normal device, $\bar{\hat{P}}_{i,k}(t)$ and $\underline{\hat{P}}_{i,k}(t)$ should be unchanged at different iterations and time steps unless $\bar{\hat{P}}_{i,k}(t) < \hat{P}_{i,k+1}(t)$ or $\underline{\hat{P}}_{i,k}(t) > \hat{P}_{i,k+1}(t)$. At each iteration, the neighbors check if

$$\underline{\hat{P}}_{i,k}(t) < \hat{P}_{i,k+1}(t) < \bar{\hat{P}}_{i,k}(t) \tag{48}$$

or

$$\bar{\hat{P}}_{i,k}(t) < \hat{P}_{i,k+1}(t) \text{ or } \underline{\hat{P}}_{i,k}(t) > \hat{P}_{i,k+1}(t) \tag{49}$$

occur. Based on the estimated states and parameters, we analyze the possibilities according to the categories of devices.

Scenario I: $i \in \mathcal{V}\backslash\mathcal{V}_S$.

If $J_{i,k} < \underline{J}^{th}_k$ or $J_{i,k} > \bar{J}^{th}_k$ and (48) holds, we can judge that device i's behavior is suspicious; while if (49) occurs, device i's behavior may be normal.

Scenario II: $i \in \mathcal{V}_S$.

If $J_{i,k} < \underline{J}^{th}_k$ or $J_{i,k} > \bar{J}^{th}_k$ $(k > k_{i,1})$ while $\underline{J}^{th}_k \leq J_{i,k_{i,1}} \leq \bar{J}^{th}_k$ holds for the estimation process (28)–(32), and (48) holds, we need to check the results of $\boldsymbol{f}_i(\hat{\boldsymbol{x}}_{i,k})$ and $\boldsymbol{g}_i(\hat{\boldsymbol{x}}_{i,k})$. If

$$\boldsymbol{f}_i(\hat{\boldsymbol{x}}_{i,k}) + \boldsymbol{g}_i(\hat{\boldsymbol{x}}_{i,k}) \neq \boldsymbol{0}, \tag{50}$$

then device i's behavior may be normal and has reached the boundary condition (6). While if $J_{i,k} < \underline{J}^{th}_k$ or $J_{i,k} > \bar{J}^{th}_k$ for the estimation process (28)–(32), and (49) holds, device i's behavior may still be normal. Otherwise, device i is considered suspicious.

If the device i is considered malicious, all neighbors initiate a majority vote. If more than half of the neighbors conclude that device i is malicious, it is officially deemed a malicious device in the DEMS.

Remark 3. Note that the majority vote is necessary here to protect legitimate devices in various situations. For example, if a neighbor labels a legitimate device as a malicious one due to local computation error or false accusation, the majority vote can protect the legitimate device in this case. A device will be

Algorithm 1 The FNAD algorithm (for each neighbor of device $i \in \mathcal{V}$)

Input: $\hat{\xi}_{i,k}(t)$, $\hat{\lambda}_{i,k}(t)$
Initialization:
$\tilde{P}_{i,0}(t) = -\hat{\xi}_{i,0}(t)$, $\forall i \in \mathcal{V}\backslash\mathcal{V}_L$;
$\tilde{P}_{i,0}(t) = \hat{\xi}_{i,0}(t)$, $\forall i \in \mathcal{V}_L$;
$\psi_{i,0}(t) = 0$.
for $k \geq N_1$ **do**
  Updates $\tilde{P}_{i,k}(t)$ for $i \in \mathcal{V}\backslash\mathcal{V}_L$ and $i \in \mathcal{V}_L$ by (22) and (23), respectively;
  Updates $\hat{P}_{i,k}(t)$ by (28)–(32);
  Updates the $J_{i,k}$ by (38)–(43) for estimation (28)-(32);
  if $J_{i,k} < \underline{J}_k^{th}$ or $J_{i,k} > \bar{J}_k^{th}$ **then**
    Device i is normal;
  else if (49) holds **then**
    Device i may be normal and reaches the boundary condition (4);
  else
    if $i \in \mathcal{V}_S$ & $k > k_{i,1}$ & $\underline{J}_{k_1}^{th} \leq J_{i,k_1} \leq \bar{J}_{k_1}^{th}$ **then**
      Checks if (50) holds;
      if (50) holds **then**
        Device i may be normal, and reaches the boundary condition (6);
      else
        Device i is malicious, and all neighbors start a majority vote;
      end if
    else
      Device i is malicious, and all neighbors start a majority vote;
    end if
  end if
end for
Output: $J_{i,k}$

considered malicious by the DEMS only if the majority votes to do so. In other words, the majority vote enhances the robustness of DEMSs.

The computation time complexity of the proposed algorithm at each iteration k is $\mathcal{O}(1)$. By summarizing the power decision observation, state and parameter estimation, and detection process, we can conclude the process of the whole FN attack detection in the form of pseudo-codes in Algorithm 1.

Remark 4. We assume that the random sequence $\nu^a_{i,N_1,k}$ is an i.i.d. process when we design the detector, though in reality, $\nu^a_{i,N_1,k}$ may be non-i.i.d. Additionally, some literature has explored the design of non-i.i.d. attack vectors (Ren, Yang, & Zhang, 2023; Ye & Zhang, 2020). However, non-i.i.d. attack vectors do not impact the effectiveness of the proposed detector. This is because $\nu_{i,N_1,k}$ is an i.i.d. random process. From the perspective of attackers, $\nu^a_{i,N_1,k}$ should closely resemble $\nu_{i,N_1,k}$ in order to evade potential attack detectors. Therefore, unless PDFs $f_{\nu_i^a}(x)$ and $f_{\nu_i}(x)$ are sufficiently similar, the assumption of i.i.d. random process $\nu^a_{i,N_1,k}$ does not impact the detection performance.

4.4. FNAD methodology

FNAD is an FN detection method for detecting Attacks 1–3 on privacy-preserving DEMSs in a distributed manner. Also, FNAD satisfies the privacy-preserving needs in DEMSs.

The operation of FNAD has the following requirements, including data forwarding from the two-hop neighbors, parameter sharing, and zero-mean Gaussian noises. The detailed analysis of data forwarding and parameter sharing is formulated in Assumptions 1–2 and corresponding interpretation, respectively, and the necessary of zero-mean Gaussian noises is analyzed in Remark 1.

In FNAD, each device collects two-hop neighbors' information and observes the target neighbors' real-time power decisions with zero-mean Gaussian noises. When the internal updating rules of the target device is under consideration, devices are able to estimate neighbors' real-time power decisions via Kalman filter. Hence, we design the FNAD algorithm to detect misbehavior based on the estimation results. Specifically, considering the residue of Kalman filter, we design an information entropy-based detection index. Then FNAD can detect malicious device via the detection index and the boundary condition.

5. Performance of the detection method

In this section, we demonstrate the performance of FNAD by analyzing the detection process of FN attacks and discussing potential threats to FNAD.

5.1. Detection performance for FN attacks

For the aforementioned three types of FN attacks, FNAD has the detection performance as follows.

Theorem 1. *The detection index $J_{i,k}$ can detect FN attacks under $\nu_i^a = \boldsymbol{z}_{i,k}^T \boldsymbol{Q}_{z_{i,k}}^{-1} \boldsymbol{z}_{i,k} \nsim \chi^2(m)$ w.p.1.*

Proof. See Appendix D.

Theorem 1 reveals the general condition of successfully detecting FN attacks. Then we consider Attack 1–3, and analyze the detection performance.

As for Attack 1, it can be viewed as an attack that aims at the measurements of the Kalman filter. To this end, we have the following theorem:

Theorem 2. *By FNAD, as the number of iterations goes to infinity, Attack 1 is detected w.p.1.*

Proof. See Appendix E.

Theorem 3. *By FNAD, as the number of iterations goes to infinity, Attack 2 can be detected w.p.1.*

Proof. See Appendix F.

Theorem 4. *By FNAD, as the number of iterations goes to infinity, Attack 3 can be detected w.p.1.*

The proof of Theorem 4 is similar to that of Theorem 3 and is omitted herein.

Remark 5. If an attack is the combination of Attacks 1–3, it can be also detected by FNAD. For example, if the Attacks 1–2 exist, since the neighbors' prior estimation is (28), FNAD can detect if $\boldsymbol{v}_{i,k}^{\xi}$ exists by the residual $\boldsymbol{z}_{i,k}$. Another kinds of the combinations of Attacks 1–3 follow the same logic.

Remark 6. As for FN attacks, the two latest works in the literature, i.e., the attacks proposed in Duan and Chow (2019a) and Ye et al. (2023) which undermine the economics of DEMS by designing FN, can be detected by FNAD algorithm if the conditions in Theorem 1 are satisfied.

Remark 7. In Theorem 1, the necessary conditions for FDI attack detection are $\nu_i^a \nsim \chi^2(m)$. Whereas, if the Attacks 1–3 vectors satisfy $\nu_i^a \sim \chi^2(m)$, even if Algorithm 1 cannot detect them, the attacks do not deviate the DEMSs from the normal operation point. For Attack 1, if the attack vector satisfies $\nu_i^a \sim \chi^2(m)$, according to the proof of Theorem 2, the attack vector is equivalent to the artificial noise $\psi_{i,k}(t)$. Hence, the attack vector does not change the optimal operating points of DEMSs. Similarly, for Attack 2 (resp. Attack 3), according to the proof of Theorem 2, the attack vector decays to 0. In like manner, the attack vector does not change the optimal operation points of DEMSs.

The proposed method is dedicated to FN attack detection. Based on previous studies, if the attackers are detected and removed from the DEMS, then the DEMS can achieve convergence and optimal operation points (Duan & Chow, 2019b). FNAD is designed based on a distributed architecture, in which each device utilizes forwarded data from the two-hop neighbors and detects malicious neighbors. If the scale of the DEMS increases, FNAD still works effectively if each device can receive data from the two-hop neighbors. Meanwhile, the privacy of devices' power decisions is not disclosed in the state estimation process, as the Kalman filter provides only an unbiased estimation, i.e., the estimated power decisions are equal to the expectation of the power decisions but not equal to the actual power decisions.

5.2. Comparison with previous works

5.2.1. Comparison with the KLD detection method

The proposed detection method originates from the KLD detector. Here we analyze their similarities and differences.

There are two main differences between the proposed method and the KLD detector. One is about the PDFs of FDI attack vectors. For the traditional KLD-based detection methods, they need the PDFs of $z_{i,k}^a$ to calculate the KLD. However, detectors usually cannot obtain the PDFs of $z_{i,k}^a$ in advance, which limits its practical application. Unlike the KLD-based detectors, the proposed detection method can detect FDI attacks without knowing the PDFs of $z_{i,k}^a$, which is advantageous for its application. The other is about the covariance of $z_{i,k}$. Traditional KLD methods are designed based on the assumption that the prior error covariance of the Kalman filter converges, and $\boldsymbol{Q}_{z_{i,k}}$ is time-invariant. While the proposed method relaxes the restriction.

As for the similarities, the proposed method achieves almost the same detection performance as the KLD detector. When iteration k is large enough, the proposed detection method is almost equivalent to the KLD detector. Especially, as iteration k goes to infinity, the proposed detection method is equivalent to the KLD detector.

5.2.2. Comparison with other previous works

In the literature, the representative works in DEMSs are the method in Duan and Chow (2019b) and Cheng and Chow (2020), which are also reputation-based approaches. Here we make a brief comparison with the previous works. The works in Duan and Chow (2019b) and Cheng and Chow (2020) do not consider the updating process of the power decisions. Hence, the works in Duan and Chow (2019b) and Cheng and Chow (2020) cannot detect FN attacks that satisfy (10)–(13) such as the attack in Ye et al. (2023). Besides, the works in Duan and Chow (2019b) and Cheng and Chow (2020) require neighbors to offer the power decision information, which may cause privacy disclosure (Ye et al., 2021).

Compared with most existing methods in control systems designed based on the residue $\boldsymbol{z}_{i,k}$, we utilize $\nu_{i,k}$ to detect FDI attacks. Because most of the existing literature assumes that the covariance of the residue in the Kalman filter process is time-invariant, which does not hold in the estimation model (28)–(32). To this end, we design the variable $\nu_{i,k}$.

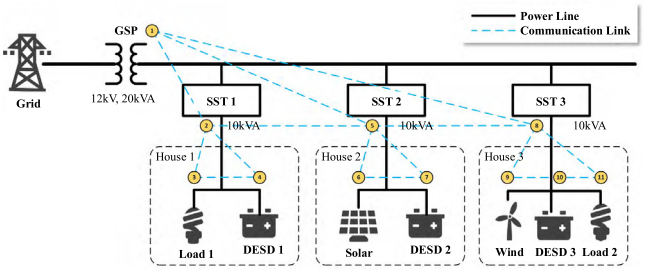


Fig. 1. The topology of the DEMS (Ye et al., 2021).

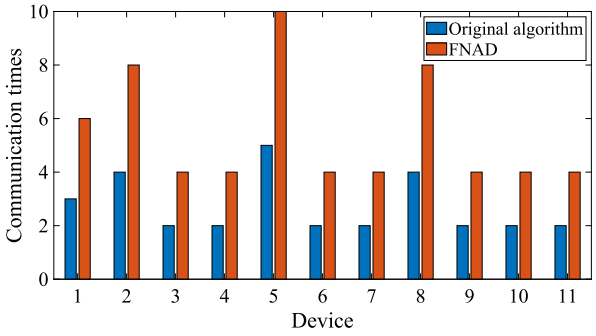


Fig. 2. Times of receiving messages at each iteration.

Table 1
Profile used in simulations.

Hour	Price (¢/kWh)	Load1 (kW)	Load2 (kW)	Wind (kW)	PV (kW)
14:00	4.33	4.3	4.4	1.8	3.8
15:00	4.24	4.3	4.35	1.9	2.5
16:00	4.22	4.2	4.35	2.1	1.3
17:00	5.76	4.4	4.35	2.1	0.4
18:00	8.34	4.8	4.7	2.2	0

Table 2
DESDs specifications and initial conditions.

	$\overline{E}_i$(kWh)	$\overline{P}_i$(kW)	$E_i(1)$(kWh)
DESD1	5	5	1
DESD2	10	5	1.5
DESD3	5	5	0.5

6. Case studies

Consider a DEMS consisting of 1 MGI, 3 DESDs, 1 PV, 1 wind turbine, and 2 loads The topology of the DEMS is shown in Fig. 1. The electricity price and the power generation or consumption for loads and RERs at each time step are shown in Table 1, and the lower and upper bounds and initial states of DESDs are shown in Table 2 (Rahbari-Asr et al., 2015). The parameters η and ρ in the distributed scheduling are set to 0.1 and 5, respectively. Other parameters are set to $w_{ij} = \frac{1}{6}$ since $\overline{d} = 5$.

Then we consider Attacks 1–3. In Attacks 1 and 2, we select Devices 1, 3, 6, and 10 as the attackers, and attack vectors are set as $\boldsymbol{v}_{i,k}^{\xi} \sim \mathbb{N}(\boldsymbol{0}, \boldsymbol{R}_1)$, $\boldsymbol{v}_{i,k}^{P} \sim \mathbb{N}(\boldsymbol{0}, \boldsymbol{R}_1)$. While in Attack 3, we select Devices 1 as the attacker, and the electricity price vector $\boldsymbol{p}$ tampers with $\boldsymbol{p} + \boldsymbol{v}_{i,k}^{p}$, where $\boldsymbol{v}_{i,k}^{p} \sim \mathbb{N}(\boldsymbol{0}, \boldsymbol{R}_1)$. Without loss of generality, we also select normal Devices 9 and 11 as reference. The parameters with regard to FNAD are set as $\rho_1 = \rho_2 = 0.9$, $\boldsymbol{R}_1 = \boldsymbol{R}_2 = diag(0.01, 0.01, \ldots, 0.01)$, $N_1 = 10$, and $N_2 = 200$. The thresholds J_k^{th} and $\bar{J}_k^{th}$ are obtained by the Monte Carlo method with the false alarm rate of 1%. For comparison, we apply χ^2

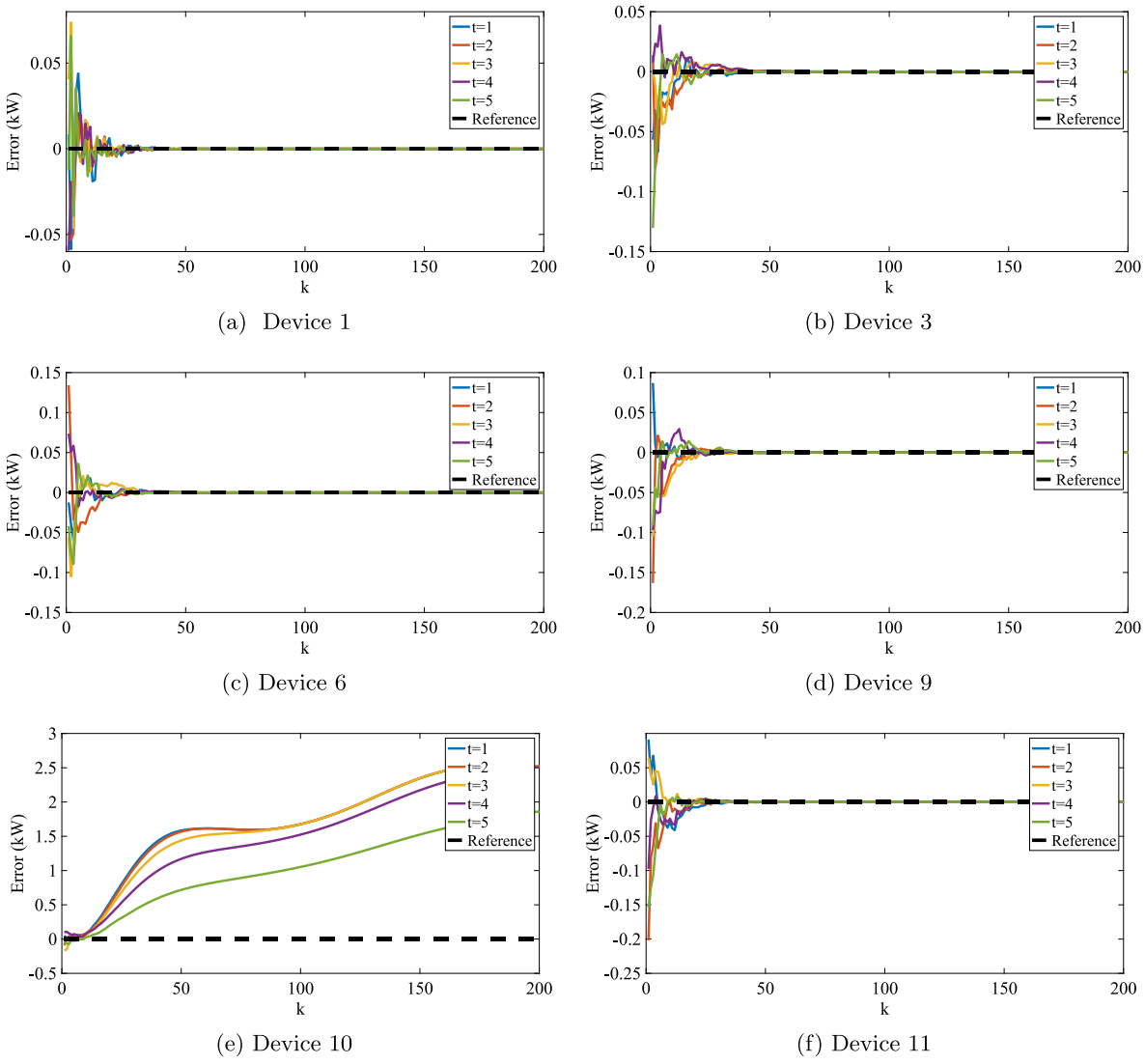


Fig. 3. Evolution of the state estimation errors without attack.

detector in the normal cases and Attacks 1–3, and parameters setting are the same as that of FNAD.

Firstly, we evaluate the communication complexity of FNAD. To save the communication costs, broadcasting is adopted in the simulation. Also, each device forwards data after receiving data from all neighbors. Hence, each device sends data only once at each iteration under the original privacy-preserving DEMS algorithm, while sends data twice per iteration under FNAD. Fig. 2 depicts the receiving message times of each device at each iteration. where the number of communications per device under the proposed algorithm is double that of the original algorithm.

Case I: Attacks do not exist in the DEMS

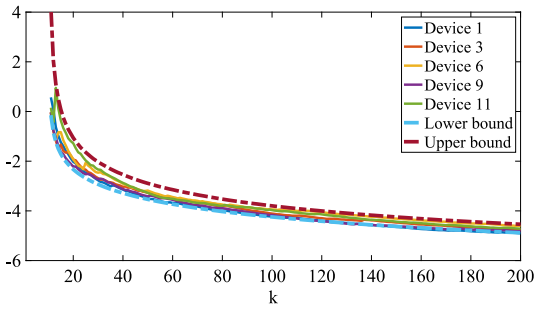
Figs. 3–5 illustrate the evolution of Devices 1, 3, 6, 9, 10, and 11 in the normal case. Fig. 3 shows the evolution of state estimation process. Obviously, as shown in Fig. 3, for Devices 1, 3, 6, 9, and 11, estimated power decisions at each time step are almost equal to the true power decisions as iteration increases, and the state estimation errors converge to 0, which verifies the unbiased estimation of the Kalman filter. While for Device 10, as shown in Fig. 3(e), the Kalman filter cannot converge after iteration k_1 as $\boldsymbol{f}_i(\hat{\boldsymbol{x}}_{i,k})+\boldsymbol{g}_i(\hat{\boldsymbol{x}}_{i,k}) \neq \boldsymbol{0}$, i.e., Device 10 charging to full or discharging to empty exists. Fig. 4 shows the evolution of the detection index $J_{i,k}$. In Fig. 4(a) $J_{i,k}$ for Devices 1, 3, 6, 9, and 11 is always within the normal range, and no false alarm occurs. However, in Fig. 4(b), $J_{i,k}$ for Device 10 exceeds the normal range as the state estimations for Device 10 contain errors. For comparison, as Fig. 5 shows, the χ^2 detector produces false alarms when detecting Devices 1, 3, 6, 9, and 11.

Case II: Attack 1 exists in the DEMS

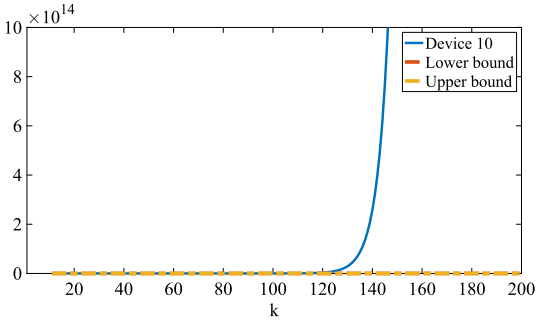
Figs. 6–8 show the evolution of selected devices under Attack 1. From Fig. 6, it is evident that the Kalman filter converges if the device is normal, while it does not converge when the device is an attacker. Fig. 7 shows the evolution of the detection index $J_{i,k}$, in which $J_{i,k}$ for attackers extremely exceed the upper bound of thresholds, while $J_{i,k}$ for normal devices remains within the lower and upper bounds. For comparison, as Fig. 8 shows, the χ^2 detector can also detect Attack 1 (shown in Fig. 8(a)), but it results in false alarm for normal devices (shown in Fig. 8(b)).

Case III: Attack 2 exists in the DEMS

Figs. 9–11 illustrate the evolution of selected devices under Attack 2. Similarly, as Fig. 9 shows, the Kalman filter converges if and only if the device is normal. Fig. 10 shows the evolution of the detection index $J_{i,k}$, in which FNAD can detect all attackers while normal devices are not affected. For comparison, as Fig. 11 shows, the χ^2 detector can also detect attacks, but it results in false alarm for normal devices.

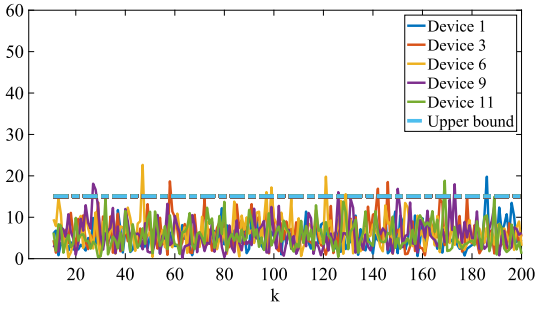


(a) Devices 1, 3, 6, 9, and 11

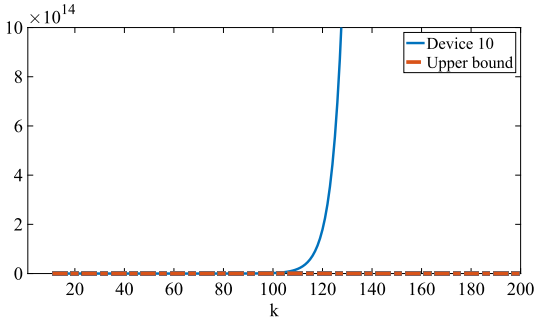


(b) Device 10

Fig. 4. Evolution of detection index without attack.



(a) Devices 1, 3, 6, 9, and 11



(b) Device 10

Fig. 5. Evolution of χ^2 detector without attack.

Case IV: Attack 3 exists in the DEMS

Figs. 12–14 show the evolution of selected devices under Attack 3. Similarly, as Fig. 12 shows, the Kalman filter converges if and only if the device is normal. Fig. 13 shows the evolution of the detection index $J_{i,k}$, in which FNAD can detect all attackers while normal devices are not affected. For comparison, as Fig. 14 shows, the attacker does not show a significant difference compared to normal devices in terms of χ^2 detector, as the frequencies and magnitudes of normal devices' false alarm χ^2 indexes are similar to those of the attacker's true alarm index.

To evaluate FNAD in terms of scalability, we also consider the DEMS topology with 5, 10, 15, 20 devices, respectively, and analyze the false alarm rate. The parameter settings of devices refer to the topology of Fig. 1. Fig. 15 shows the false alarm rate under different topologies and the comparison with the χ^2 detectors. It is obvious that FNAD has a lower false alarm rate.

In summary, FNAD can effectively detect Attacks 1–3 while not affecting normal devices. Also, FNAD has the capacity of detecting Attacks 1–3 in larger-scale topologies with different devices. Nevertheless, the χ^2 detector can detect Attacks 1 and 2 while causing false alarms, and cannot reliably detect Attack 3. Therefore, FNAD performs better than χ^2 detector in DEMS.

7. Conclusions

We have studied the FDI attacks issue in privacy-preserving distributed control of microgrids, summarized and analyzed three categories of representative FDI attacks in the literature. To this end, we proposed the FNAD algorithm to detect FDI attacks, using the power decision observation and estimation, and the information entropy-based attack detection. We have proved that FNAD can detect the aforementioned three FDI attacks. Compared with previous methods, the advantages of FNAD include zero prior knowledge to attacks and privacy preserving of the devices. In case studies, the effectiveness of FNAD is verified, and FNAD has a lower false alarm rate than the χ^2 detector. For future work, it is desirable to explore an FDI attack detection method that does not require data from two-hop neighbors, which can substantially reduce the detection cost.

Acknowledgments

This work was supported in part by the Frontier Technologies R&D Program of Jiangsu Province under Grant BF2024065, in part by the Shenzhen Science and Technology Program under Grant JCYJ20230807114609019, in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), and in part by Compute Canada.

Appendix A. Proof of Lemma 2

Proof. By rewriting (10), we can derive the power decision of device $i \in \mathcal{V}\backslash\mathcal{V}_L$ by

$$P_{i,k+1}(t) = P_{i,k}(t) + \sum_{j\in\mathcal{N}_i} w_{ij}\hat{\xi}^+_{j,k}(t) + w_{ii}\hat{\xi}^+_{i,k}(t) - \hat{\xi}_{i,k+1}(t). \tag{A.1}$$

Nevertheless, neighbors can only obtain $\hat{\xi}^+_{i,k}(t)$ instead of $\hat{\xi}_{i,k}(t)$. Then (A.1) can be rewritten as

$$\begin{aligned} P_{i,k+1}(t) =& P_{i,k}(t) + \sum_{j\in\mathcal{N}_i} w_{ij}\hat{\xi}^+_{j,k}(t) + w_{ii}\hat{\xi}^+_{i,k}(t) \\ & - \hat{\xi}^+_{i,k+1}(t) + \psi_{i,k+1}(t) \\ =& \sum_{h=0}^{k}\sum_{j\in\mathcal{N}_i\cup\{i\}} w_{ij}\hat{\xi}^+_{j,h}(t) - \sum_{h=0}^{k+1}\hat{\xi}^+_{i,h}(t) \\ & + \sum_{h=0}^{k+1}\psi^+_{i,h}(t). \end{aligned} \tag{A.2}$$

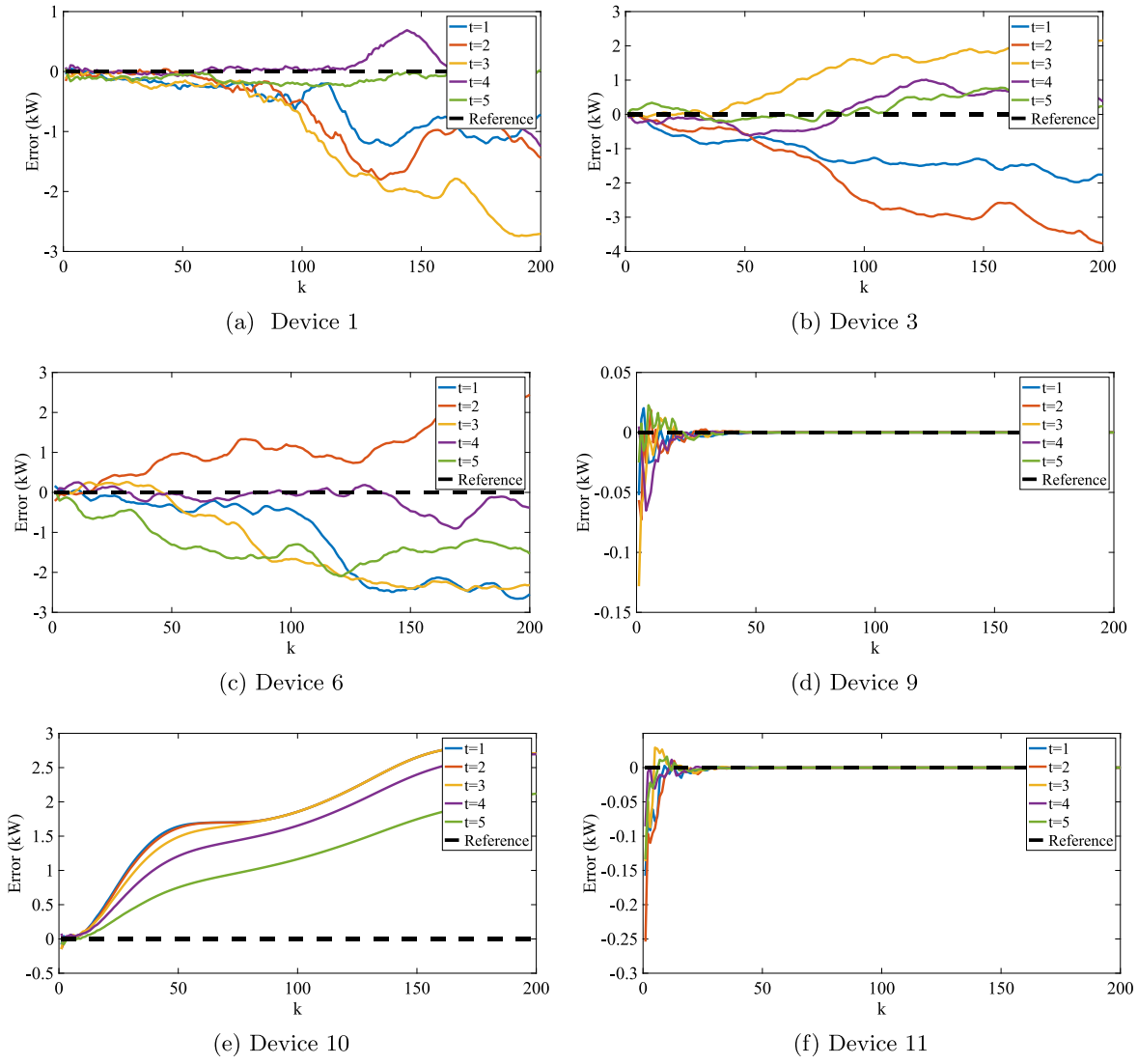


Fig. 6. Evolution of the state estimation errors under Attack 1.

Substitute (A.2) into (22):

$$\tilde{P}_{i,k}(t) = P_{i,k}(t) + \mathrm{sgn}(\tau)\rho_1^k \theta_{i,k}(t), \tag{A.3}$$

where

$$\tau = \begin{cases} 1, & \text{if } i \in \mathcal{V}_L \\ -1, & \text{otherwise.} \end{cases} \tag{A.4}$$

Since $0 < \rho_1 < 1$, one has

$$\lim_{k\to\infty} \tilde{P}_{i,k}(t) = \lim_{k\to\infty} P_{i,k}(t). \tag{A.5}$$

Hence, each device can observe the power decisions of all neighbors when DEMS converges if the neighbors obey the updating rule of (10). Therefore, the lemma is proved.

Appendix B. Proof of Lemma 3

Proof. According to (A.1)–(A.2), we can deduce that the noise $\psi_{i,k}(t)$ decreases as the updating process goes, and the measurement errors caused by the noise $\psi_{i,k}(t)$ also decrease. Hence, the results of state and parameter estimation would depend on the corresponding observation results as the iteration increases, and

$$\lim_{k\to\infty} \hat{\boldsymbol{x}}_{i,k} = \lim_{k\to\infty} \boldsymbol{x}_{i,k}. \tag{B.1}$$

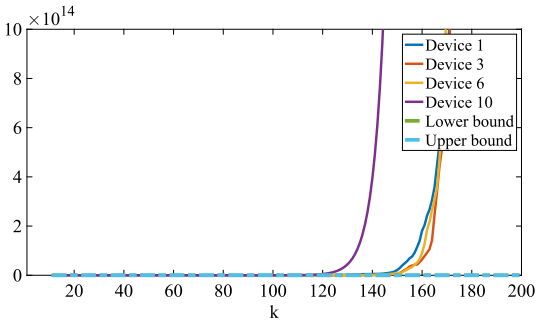
Therefore, the lemma is proved.

Appendix C. Proof of Lemma 4

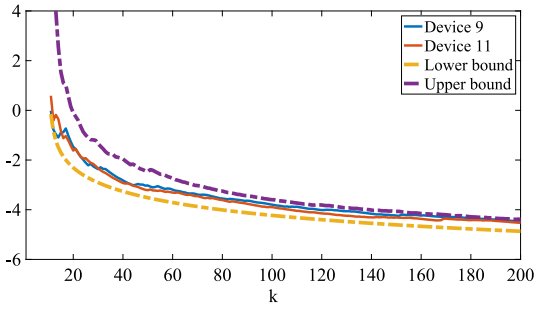
Proof. Consider the case of $k \to \infty$. If the target device is normal, since $\nu_{i,k}$ obeys $\chi^2(m)$ distribution, one has

$$\lim_{k\to\infty} J_{i,k} = \int_0^\infty f_{\nu_i}(\zeta) \ln f_{\nu_i}(\zeta)\, d\zeta + \frac{1}{2}\int_0^\infty f_{\nu_i}(\zeta)\, d\zeta - \left(\frac{m}{2} - 1\right)\int_0^\infty f_{\nu_i}(\zeta) \ln \zeta\, d\zeta, \tag{C.1}$$

where $f_{\nu_i}(\zeta)$ denotes the PDF of $\chi^2(m)$ distribution.

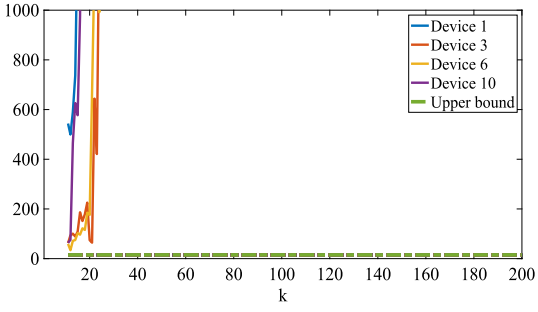


(a) Devices 1, 3, 6, and 10

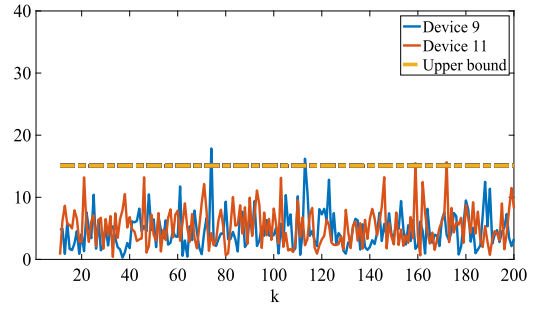


(b) Devices 9 and 11

Fig. 7. Evolution of FNAD's detection index under Attack 1.



(a) Devices 1, 3, 6, and 10



(b) Devices 9 and 11

Fig. 8. Evolution of χ^2 detector under Attack 1.

Adding $\ln\left[2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)\right]$ to both sides of (C.1), based on Definition 3, we can obtain that

$$\begin{aligned}
&\lim_{k\to\infty} J_{i,k} + \ln\left[2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)\right] \\
&= \int_0^\infty f_{\nu_i}(\zeta)\ln f_{\nu_i}(\zeta)\,d\zeta + \frac{1}{2}\int_0^\infty f_{\nu_i}(\zeta)\,d\zeta \\
&\quad - \left(\frac{m}{2}-1\right)\int_0^\infty f_{\nu_i}(\zeta)\ln\zeta\,d\zeta + \ln\left[2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)\right] \\
&= \int_0^\infty f_{\nu_i}(\zeta)\ln f_{\nu_i}(\zeta)\,d\zeta \\
&\quad - \int_0^\infty f_{\nu_i}(\zeta)\ln\left[\frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}\zeta^{\frac{m}{2}-1}e^{-\frac{\zeta}{2}}\right]d\zeta \\
&= \int_0^\infty f_{\nu_i}(\zeta)\ln\frac{f_{\nu_i}(\zeta)}{f_{\nu_i}(\zeta)}\,d\zeta \\
&= 0. \qquad \text{(C.2)}
\end{aligned}$$

Hence, when k goes to infinity, the lower and upper thresholds converge to the additive inverse of the convergence value of $J_{i,k}$, i.e.,

$$\lim_{k\to\infty} J_k^{th} = \lim_{k\to\infty} \bar{J}_k^{th} = -\lim_{k\to\infty} J_{i,k} = -\ln\left[2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)\right].$$

Therefore, the lemma is proved.

Appendix D. Proof of Theorem 1

Proof. For the random sequence $\nu^a_{i,N_1,k}$, since it is i.i.d., we can count the proportions in each sub-interval of $[0, \max\{\nu^a_{i,N_1,k}\}]$, and the proportion of $[aq, aq+q)$ can be denoted as

$$p_{\nu^a_{i,N_1,k}}([aq, aq+q)) = \frac{\text{number of samples in } [aq, aq+q)}{k+1}, \qquad \text{(D.1)}$$

where a is an integer that satisfies $0 \le a \le \frac{\max\{\nu^a_{i,N_1,k}\}}{q} - 1$. Then we can obtain that

$$\mathbb{P}\left\{\lim_{k\to\infty}\lim_{q\to 0} p_{\nu^a_{i,N_1,k}}([x, x+q)) = f_{\nu_i^a}(x)\right\} = 1 \qquad \text{(D.2)}$$

and

$$\mathbb{P}\left\{\lim_{k\to\infty} \mathcal{H}(\nu^a_{i,N_1,k}) = -\int_0^\infty f_{\nu_i^a}(x)\ln f_{\nu_i^a}(x)\,dx\right\} = 1. \qquad \text{(D.3)}$$

Similarly, in accordance with the strong law of large numbers, we have

$$\mathbb{P}\left\{\lim_{k\to\infty} \bar{\nu}^a_{i,N_1,k} = \mathbb{E}\left[\nu_i^a\right]\right\} = 1, \qquad \text{(D.4)}$$

$$\mathbb{P}\left\{\lim_{k\to\infty} \overline{\ln \nu^a_{i,N_1,k}} = \mathbb{E}\left[\ln \nu_i^a\right]\right\} = 1. \qquad \text{(D.5)}$$

Then we consider the KLD, which is the relative entropy of two random sequences. By rewriting (39) and considering ν_i and ν_i^a, we have

$$\begin{aligned}
D(\nu_i^a|\nu_i) = &-h(\nu_i^a) + \frac{1}{2}\mathbb{E}\left[\nu_i^a\right] - \left(\frac{m}{2}-1\right)\mathbb{E}\left[\ln \nu_i^a\right] \\
&+ \ln\left[2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)\right] \qquad \text{(D.6)}
\end{aligned}$$

where $f_{\nu_i}(x)$ and $f_{\nu_i^a}(x)$ are the PDFs of random variable ν_i and ν_i^a, respectively, and

$$h(\nu_i^a) = -\int_{\{x|f_{\nu_i^a}(x)>0\}} f_{\nu_i^a}(x)\ln f_{\nu_i^a}(x)\,dx \qquad \text{(D.7)}$$

is the information entropy of ν_i^a. Since ν_i is χ^2 distributed at each iteration, one has

$$f_{\nu_i}(x) = \frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}x^{\frac{m}{2}-1}e^{-\frac{x}{2}}, \qquad \text{(D.8)}$$

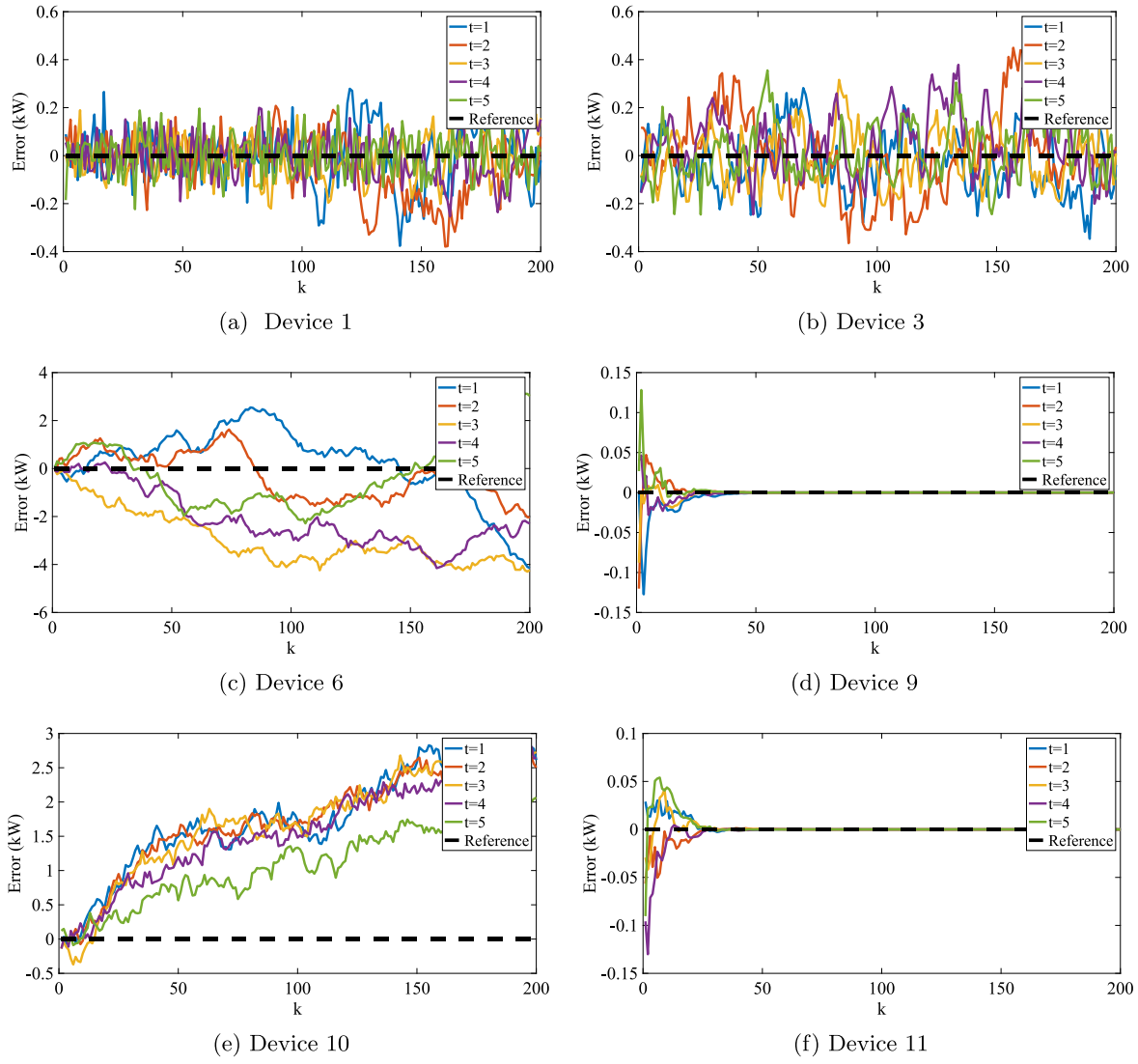


Fig. 9. Evolution of the state estimation errors under Attack 2.

then we have

$$D(\nu_i^a|\nu_i) = -h(\nu_i^a) + \ln\left[2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)\right] + \frac{1}{2}\mathbb{E}\left[\nu_i^a\right] - \left(\frac{m}{2}-1\right)\mathbb{E}\left[\ln \nu_i^a\right]. \tag{D.9}$$

While for the attacks do not satisfy $\boldsymbol{z}_{i,k}^a \sim \mathbb{N}(0, \boldsymbol{Q}_{z_{i,k}})$, we can deduce that $f_{\nu_i^a}(x) \neq f_{\nu_i}(x)$ holds. According to (40) and combining (D.2)–(D.5), we have

$$\lim_{k\to\infty} J_{i,k} = \int_0^\infty f_{\nu_i^a}(x)\ln f_{\nu_i^a}(x)\,dx + \frac{1}{2}\mathbb{E}\left[\nu_i^a\right] - \left(\frac{m}{2}-1\right)\mathbb{E}\left[\ln \nu_i^a\right] \quad \text{w.p.1.} \tag{D.10}$$

We can find that

$$\lim_{k\to\infty} J_{i,k} = D(\nu_i^a|\nu_i) - \ln\left[2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)\right] \quad \text{w.p.1.} \tag{D.11}$$

Note that $D(\nu_i^a|\nu_i)\rangle 0$ if $f_{\nu_i^a}(x) \neq f_{\nu_i}(x)$. Considering (44), we can obtain that

$$\lim_{k\to\infty} J_{i,k} > \lim_{k\to\infty} \bar{J}_k^{th} \quad \text{w.p.1} \tag{D.12}$$

for the attacks do not satisfy $\boldsymbol{z}_{i,k}^a \sim \mathbb{N}(0, \boldsymbol{Q}_{z_{i,k}})$.

Hence, $J_{i,k}$ with ν_i^a cannot be in the normal range of that with ν_i at all iterations, and the detection success rate is larger than the false alarm rate.

Thus, the proof of Theorem 1 is completed.

Appendix E. Proof of Theorem 2

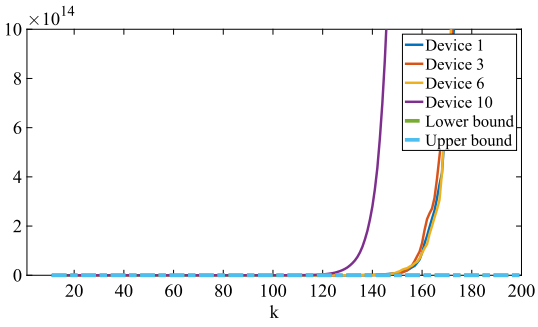
Proof. Considering the detection process under the estimation (28)–(32), we can derive the power decision observation under Attack 1 in accordance with the dynamics of (22)–(23):

$\forall i \in \mathcal{V}\backslash\mathcal{V}_L$:

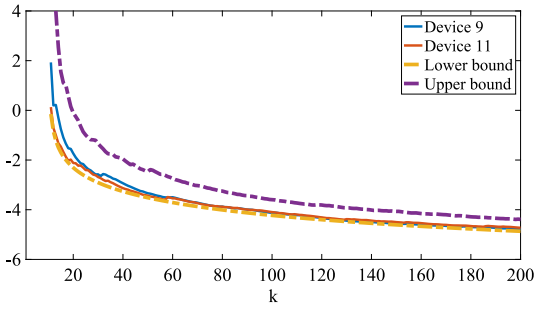
$$\tilde{P}_{i,k}^a(t) = \sum_{h=0}^{k-1}\sum_{j\in\mathcal{N}_i} w_{ij}\hat{\xi}_{j,h}^+(t) + \sum_{h=0}^{k-1} w_{ii}\hat{\xi}_{i,h}^a(t) - \sum_{h=0}^{k}\hat{\xi}_{i,h}^a(t) = P_{i,k}(t) - \sum_{h=0}^{k} v_{i,h}^{\xi}(t), \tag{E.1}$$

$\forall i \in \mathcal{V}_L$:

$$\tilde{P}_{i,k}^a(t) = \sum_{h=0}^{k}\hat{\xi}_{i,h}^a(t) - \sum_{h=0}^{k-1} w_{ii}\hat{\xi}_{i,h}^a(t) - \sum_{h=0}^{k-1}\sum_{j\in\mathcal{N}_i} w_{ij}\hat{\xi}_{j,h}^+(t) = P_{i,k}(t) + \sum_{h=0}^{k} v_{i,h}^{\xi}(t). \tag{E.2}$$

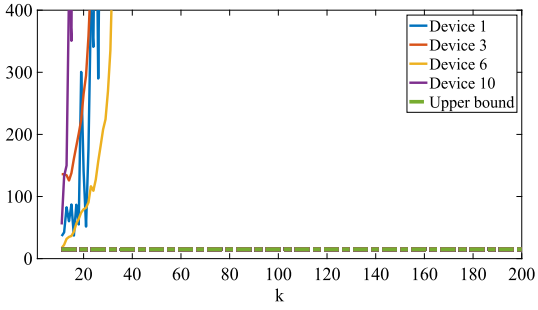


(a) Devices 1, 3, 6, and 10

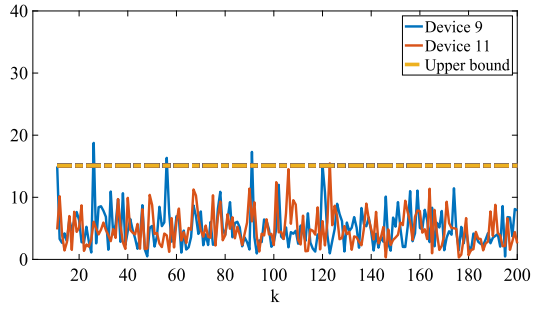


(b) Devices 9 and 11

Fig. 10. Evolution of FNAD's detection index under Attack 2.



(a) Device s 1, 3, 6, and 10

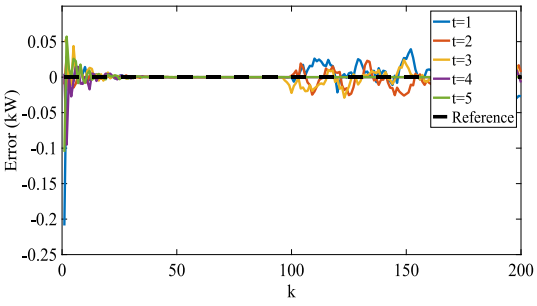


(b) Devices 9 and 11

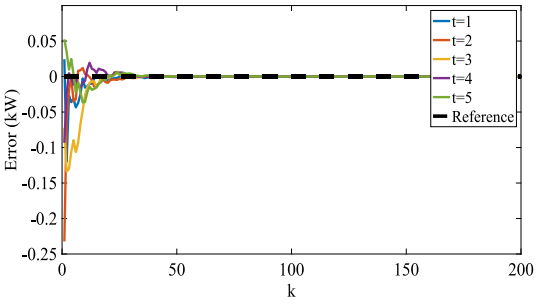
Fig. 11. Evolution of χ^2 detector under Attack 2.

Then considering (35), we can derive the residue related to the attack vector and artificial noise:

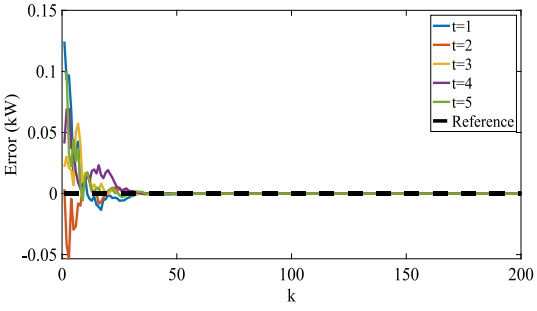
$$\begin{aligned} \boldsymbol{z}_{i,k} =& \boldsymbol{y}_{i,k} - \hat{\boldsymbol{x}}_{i,k}^{-} \\ =& \boldsymbol{x}_{i,k} - \hat{\boldsymbol{x}}_{i,k}^{-} + \operatorname{sgn}(\tau)\rho_1^k \boldsymbol{\theta}_{i,k}, \\ \boldsymbol{z}_{i,k}^a =& \boldsymbol{y}_{i,k}^a - \hat{\boldsymbol{x}}_{i,k}^{-} \end{aligned} \tag{E.3}$$



(a) Device 1



(b) Device 9



(c) Device 11

Fig. 12. Evolution of the state estimation errors under Attack 3.

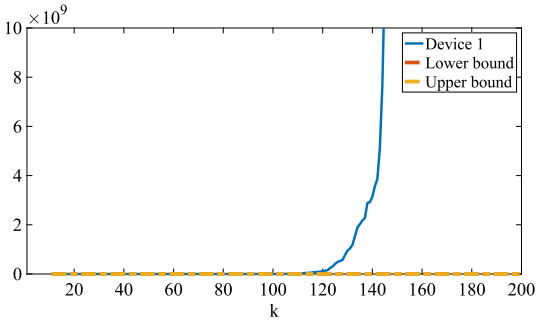
$$= \boldsymbol{x}_{i,k} - \hat{\boldsymbol{x}}_{i,k}^{-} + \operatorname{sgn}(\tau) \sum_{h=0}^{k} \boldsymbol{v}_{i,h}^{\xi}. \tag{E.4}$$

Referring to Theorem 1, when $\boldsymbol{Q}_{z_{i,k}}$ has been determined by (36), we can deduce that $\sum_{h=0}^{k} \boldsymbol{v}_{i,h}^{\xi}$ should have the same statistical characters as $\rho_1^k \boldsymbol{\theta}_{i,k}$ that $\nu_{i,k}^a$ can obtain the same statistical characters as $\nu_{i,k}$. Note that $\rho_1^k \boldsymbol{\theta}_{i,k}$ is zero-mean, then $\sum_{h=0}^{k} \boldsymbol{v}_{i,h}^{\xi}$ should be also zero-mean. Since $\boldsymbol{v}_{i,h}^{\xi}$ at iterations from 0 to $k-1$ has been decided by the attacker, the expectation of $\boldsymbol{v}_{i,k}^{\xi}$ is $-\sum_{h=0}^{k-1} \boldsymbol{v}_{i,h}^{\xi}$. Hence, if the attack vector $\boldsymbol{v}_{i,k}^{\xi}$ is generated by

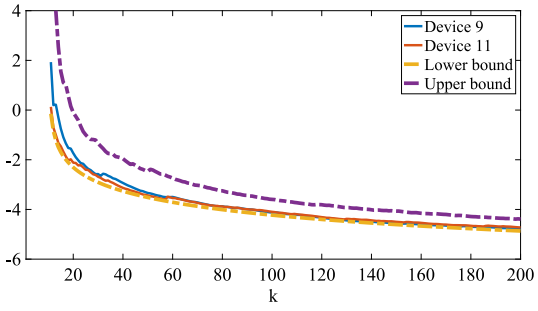
$$\boldsymbol{v}_{i,k}^{\xi} \sim \begin{cases} \mathbb{N}(\boldsymbol{0}, \boldsymbol{R}_1), & k = 0 \\ \mathbb{N}(-\sum_{h=0}^{k-1} \boldsymbol{v}_{i,h}^{\xi}, \rho_1^{2k} \boldsymbol{R}_1), & k \geq 1, \end{cases} \tag{E.5}$$

where $-\sum_{h=0}^{k-1} \boldsymbol{v}_{i,h}^{\xi}$ is known information when the attacker generates $\boldsymbol{v}_{i,k}^{\xi}$. $\sum_{h=0}^{k} \boldsymbol{v}_{i,h}^{\xi}$ can present the same probability distribution with the noise in DP, i.e.,

$$\sum_{h=0}^{k} \boldsymbol{v}_{i,h}^{\xi} \sim \mathbb{N}(\boldsymbol{0}, \rho_1^{2k} \boldsymbol{R}_1), \tag{E.6}$$

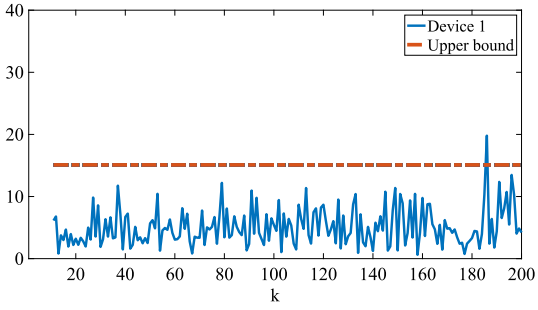


(a) Device 1

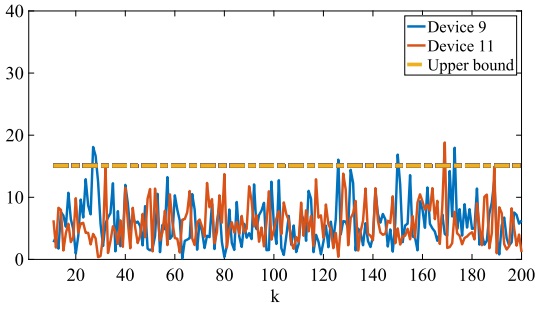


(b) Devices 9 and 11

Fig. 13. Evolution of FNAD's detection index under Attack 3.



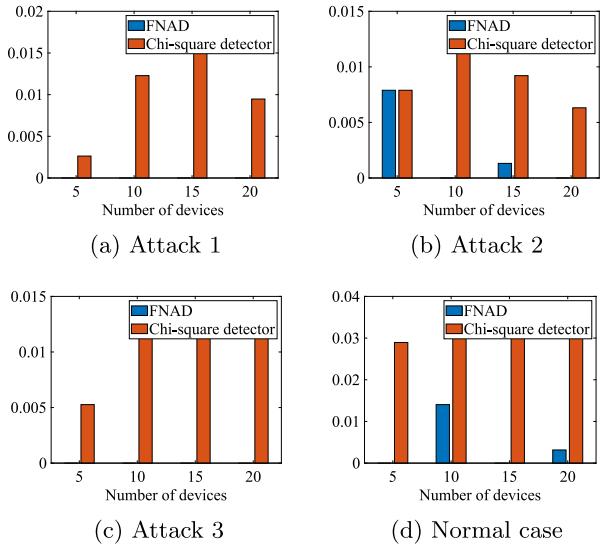
(a) Device 1



(b) Devices 9 and 11

Fig. 14. Evolution of χ^2 detector under Attack 3.

which means $\boldsymbol{v}_{i,h}^{\xi}$ generated by (E.5) cannot be detected by FNAD. However, $\boldsymbol{v}_{i,h}^{\xi}$ generated by (E.5) cannot affect the operation states of DEMS, which is meaningless for FN attacks. In other words, they are equivalent to the noise (11), which has been proved in Zhao, He, and Chen (2018), Zhao et al. (2017). Therefore, we complete the proof of Theorem 2.



(a) Attack 1 (b) Attack 2

(c) Attack 3 (d) Normal case

Fig. 15. False alarm rate comparison Between FNAD and χ^2 detector.

Appendix F. Proof of Theorem 3

Proof. The proof of Theorem 3 is similar to that of Theorem 2. Similarly, we can derive the power decision observation under Attack 2 in accordance with the dynamics of (22)–(23):

$$\forall i \in \mathcal{V}\backslash\mathcal{V}_L:$$

$$\begin{aligned}\tilde{P}_{i,k}^{a}(t) =& \sum_{h=0}^{k-1}\sum_{j\in\mathcal{N}_i} w_{ij}\hat{\xi}_{j,h}^{+}(t) + \sum_{h=0}^{k-1} w_{ii}\hat{\xi}_{i,h}(t) - \sum_{h=0}^{k}\hat{\xi}_{i,h}(t)\\ =& P_{i,k}^{a}(t) - \rho_1^k\theta_{i,k}(t)\\ =& P_{i,k}(t) + v_{i,k}^{P}(t) - \rho_1^k\theta_{i,k}(t),\end{aligned} \tag{F.1}$$

$$\forall i \in \mathcal{V}_L:$$

$$\begin{aligned}\tilde{P}_{i,k}^{a}(t) =& \sum_{h=0}^{k}\hat{\xi}_{i,h}^{a}(t) - \sum_{h=0}^{k-1} w_{ii}\hat{\xi}_{i,h}(t) - \sum_{h=0}^{k-1}\sum_{j\in\mathcal{N}_i} w_{ij}\hat{\xi}_{j,h}^{+}(t)\\ =& P_{i,k}^{a}(t) + \rho_1^k\theta_{i,k}(t)\\ =& P_{i,k}(t) + v_{i,k}^{P}(t) + \rho_1^k\theta_{i,k}(t).\end{aligned} \tag{F.2}$$

Then considering (35), we can derive the residue related to the attack vector and artificial noise:

$$\begin{aligned}\boldsymbol{z}_{i,k} =& \boldsymbol{y}_{i,k} - \hat{\boldsymbol{x}}_{i,k}^{-}\\ =& \boldsymbol{x}_{i,k} - \hat{\boldsymbol{x}}_{i,k}^{-} + \mathrm{sgn}(\tau)\rho_1^k\boldsymbol{\theta}_{i,k},\end{aligned} \tag{F.3}$$

$$\begin{aligned}\boldsymbol{z}_{i,k}^{a} =& \boldsymbol{y}_{i,k}^{a} - \hat{\boldsymbol{x}}_{i,k}^{-}\\ =& \boldsymbol{x}_{i,k} - \hat{\boldsymbol{x}}_{i,k}^{-} + \boldsymbol{v}_{i,h}^{P} + \mathrm{sgn}(\tau)\rho_1^k\boldsymbol{\theta}_{i,k}.\end{aligned} \tag{F.4}$$

And we can deduce that $\boldsymbol{v}_{i,k}^{P}$ should have the same statistical characters as $\rho_1^k\boldsymbol{\theta}_{i,k}$ that $\nu_{i,k}^{a}$ can obtain the same statistical characters as $\nu_{i,k}$, i.e., $\boldsymbol{v}_{i,k}^{P}$ satisfies

$$\boldsymbol{v}_{i,k}^{P} \sim \mathbb{N}(-\mathrm{sgn}(\tau)\rho_1^k\boldsymbol{\theta}_{i,k},\ \rho_1^{2k}\boldsymbol{R}_1) \tag{F.5}$$

for estimation (28)–(32), where $\rho_1^k\boldsymbol{\theta}_{i,k}$ is known information when the attacker generates $\boldsymbol{v}_{i,k}^{P}$. However, similarly, $\boldsymbol{v}_{i,h}^{P}$ generated by (F.5) cannot affect the operation states of DEMS. This is because the attack vector decreases as k increases, and

$$\boldsymbol{v}_{i,k}^{P} \to \boldsymbol{0}, \tag{F.6}$$

$$\lim_{k\to\infty} P_{i,k}^{a}(t) = \lim_{k\to\infty} P_{i,k}(t). \tag{F.7}$$

Hence, the attack vector does not affect the convergence value of $P_i(t)$, which is meaningless for FN attacks. Therefore, we complete the proof of Theorem 3.

References

Almehizia, A. A., Al-Masri, H. M. K., & Ehsani, M. (2019). Integration of renewable energy sources by load shifting and utilizing value storage. *IEEE Transactions on Smart Grid*, *10*(5), 4974–4984.

An, L., Duan, J., Chow, M.-Y., & Duel-Hallen, A. (2019). A distributed and resilient bargaining game for weather-predictive microgrid energy cooperation. *IEEE Transactions on Industrial Informatics*, *15*(8), 4721–4730.

Ananduta, W., & Ocampo-Martinez, C. (2021). Event-triggered partitioning for non-centralized predictive-control-based economic dispatch of interconnected microgrids. *Automatica*, *132*, Article 109829.

Chen, X., Cai, H., & Su, Y. (2023). Distributed dual objective control of an energy storage system under jointly connected switching networks. *Automatica*, *152*, Article 110979.

Chen, Y., Qi, D., Dong, H., Li, C., Li, Z., & Zhang, J. (2021). A FDI attack-resilient distributed secondary control strategy for islanded microgrids. *IEEE Transactions on Smart Grid*, *12*(3), 1929–1938.

Cheng, Z., & Chow, M.-Y. (2020). Resilient collaborative distributed energy management system framework for cyber-physical DC microgrids. *IEEE Transactions on Smart Grid*, *11*(6), 4637–4649.

Cui, L., Qu, Y., Gao, L., Xie, G., & Yu, S. (2020). Detecting false data attacks using machine learning techniques in smart grid: A survey. *Journal of Network and Computer Applications*, *170*, Article 102808.

Duan, J., & Chow, M.-Y. (2019a). A novel data integrity attack on consensus-based distributed energy management algorithm using local information. *IEEE Transactions on Industrial Informatics*, *15*(3), 1544–1553.

Duan, J., & Chow, M.-Y. (2019b). A resilient consensus-based distributed energy management algorithm against data integrity attacks. *IEEE Transactions on Smart Grid*, *10*(5), 4729–4740.

Guo, Z., Shi, D., Johansson, K. H., & Shi, L. (2018). Worst-case stealthy innovation-based linear attack on remote state estimation. *Automatica*, *89*, 117–124.

He, J., Cai, L., Cheng, P., Pan, J., & Shi, L. (2019). Distributed privacy-preserving data aggregation against dishonest nodes in network systems. *IEEE Internet of Things Journal*, *6*(2), 1462–1470.

He, J., Cai, L., & Guan, X. (2018). Preserving data-privacy with added noises: Optimal estimation and privacy analysis. *IEEE Transactions on Information Theory*, *64*(8), 5677–5690.

Higgins, M., Teng, F., & Parisini, T. (2021). Stealthy MTD against unsupervised learning-based blind FDI attacks in power systems. *IEEE Transactions on Information Forensics and Security*, *16*, 1275–1287.

Lakshminarayana, S., Kammoun, A., Debbah, M., & Poor, H. V. (2021). Data-driven false data injection attacks against power grids: A random matrix approach. *IEEE Transactions on Smart Grid*, *12*(1), 635–646.

Li, Y., Wei, X., Li, Y., Dong, Z., & Shahidehpour, M. (2022). Detection of false data injection attacks in smart grid: A secure federated deep learning approach. *IEEE Transactions on Smart Grid*, *13*(6), 4862–4872.

Li, L., Yang, H., Xia, Y., & Dai, L. (2022). Distributed filtering for nonlinear systems under false data injection attack. *Automatica*, *145*, Article 110521.

Liang, G., Weller, S. R., Zhao, J., Luo, F., & Dong, Z. Y. (2017). The 2015 Ukraine blackout: Implications for false data injection attacks. *IEEE Transactions on Power Systems*, *32*(4), 3317–3318.

Liang, G., Zhao, J., Luo, F., Weller, S. R., & Dong, Z. Y. (2017). A review of false data injection attacks against modern power systems. *IEEE Transactions on Smart Grid*, *8*(4), 1630–1638.

Liu, M., Zhao, C., Zhang, Z., & Deng, R. (2022). Explicit analysis on effectiveness and hiddenness of moving target defense in AC power systems. *IEEE Transactions on Power Systems*, *37*(6), 4732–4746.

Muhtadi, A., Pandit, D., Nguyen, N., & Mitra, J. (2021). Distributed energy resources based microgrid: Review of architecture, control, and reliability. *IEEE Transactions on Industry Applications*, *57*(3), 2223–2235.

Musleh, A. S., Chen, G., & Dong, Z. Y. (2020). A survey on the detection algorithms for false data injection attacks in smart grids. *IEEE Transactions on Smart Grid*, *11*(3), 2218–2234.

Qin, J., Wan, Y., Yu, X., Li, F., & Li, C. (2019). Consensus-based distributed coordination between economic dispatch and demand response. *IEEE Transactions on Smart Grid*, *10*(4), 3709–3719.

Rahbari-Asr, N., Zhang, Y., & Chow, M.-Y. (2015). Cooperative distributed scheduling for storage devices in microgrids using dynamic KKT multipliers and consensus networks. In *IEEE power & energy society general meeting* (pp. 1–5).

Ren, X.-X., Yang, G.-H., & Zhang, X.-G. (2023). Optimal stealthy attack with historical data on cyber–physical systems. *Automatica*, *151*, Article 110895.

Su, Q., Li, S., Gao, Y., Huang, X., & Li, J. (2021). Observer-based detection and reconstruction of dynamic load altering attack in smart grid. *Journal of the Franklin Institute*, *358*(7), 4013–4027.

Tian, J., Wang, B., Wang, Z., Cao, K., Li, J., & Ozay, M. (2022). Joint adversarial example and false data injection attacks for state estimation in power systems. *IEEE Transactions on Cybernetics*, *52*(12), 13699–13713.

Wang, R., Li, Q., Zhang, B., & Wang, L. (2019). Distributed consensus based algorithm for economic dispatch in a microgrid. *IEEE Transactions on Smart Grid*, *10*(4), 3630–3640.

Wei, F., Wan, Z., & He, H. (2020). Cyber-attack recovery strategy for smart grid based on deep reinforcement learning. *IEEE Transactions on Smart Grid*, *11*(3), 2476–2486.

Xu, W., Jaimoukha, I. M., & Teng, F. (2023). Robust moving target defence against false data injection attacks in power grids. *IEEE Transactions on Information Forensics and Security*, *18*, 29–40.

Yang, H., He, X., Wang, Z., Qiu, R. C., & Ai, Q. (2022). Blind false data injection attacks against state estimation based on matrix reconstruction. *IEEE Transactions on Smart Grid*, *13*(4), 3174–3187.

Yang, S., Tan, S., & Xu, J. (2013). Consensus based approach for economic dispatch problem in a smart grid. *IEEE Transactions on Power Systems*, *28*(4), 4416–4426.

Yang, M., & Zhai, J. (2022). Observer-based switching-like event-triggered control of nonlinear networked systems against DoS attacks. *IEEE Transactions on Control of Network Systems*, *9*(3), 1375–1384.

Yang, M., & Zhai, J. (2024). Predictor-based decentralized event-triggered secure control for nonlinear cyber-physical systems under replay attacks and time delay. *IEEE Transactions on Control of Network Systems*, *11*(1), 150–160.

Ye, F., Cao, X., Cheng, Z., & Chow, M.-Y. (2023). CASL: A novel collusion attack against distributed energy management systems. *IEEE Transactions on Smart Grid*, *14*(6), 4717–4728.

Ye, F., Cao, X., Chow, M.-Y., & Cai, L. (2024). Privacy-preserving average consensus: Fundamental analysis and a generic framework design. *IEEE Transactions on Information Theory*, *70*(4), 2870–2885.

Ye, F., Cheng, Z., Cao, X., & Chow, M.-Y. (2021). A random-weight privacy-preserving algorithm with error compensation for microgrid distributed energy management. *IEEE Transactions on Information Forensics and Security*, *16*, 4352–4362.

Ye, D., & Zhang, T.-Y. (2020). Summation detector for false data-injection attack in cyber-physical systems. *IEEE Transactions on Cybernetics*, *50*(6), 2338–2345.

Zhao, C., Chen, J., He, J., & Cheng, P. (2018). Privacy-preserving consensus-based energy management in smart grids. *IEEE Transactions on Signal Processing*, *66*(23), 6162–6176.

Zhao, C., He, J., & Chen, J. (2018). Resilient consensus with mobile detectors against malicious attacks. *IEEE Transactions on Signal and Information Processing over Networks*, *4*(1), 60–69.

Zhao, C., He, J., Cheng, P., & Chen, J. (2017). Analysis of consensus-based distributed economic dispatch under stealthy attacks. *IEEE Transactions on Industrial Electronics*, *64*(6), 5107–5117.

Zhu, M., & Martinez, S. (2012). On distributed convex optimization under inequality and equality constraints. *IEEE Transactions on Automatic Control*, *57*(1), 151–164.

Feng Ye received the B.S. degree in electrical engineering and automation from Northeastern University, Shenyang, China, in 2019, and the Ph.D. degree in control science and engineering from Southeast University, Nanjing, China, in 2024. He is currently a Postdoctoral Fellow with the Department of Electrical & Computer Engineering, University of Victoria, Victoria, BC, Canada. His current research interests include distributed coordination and optimization, smart grid, privacy preserving, and cyber security.

Xianghui Cao received the B.S. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, China, in 2006 and 2011, respectively. From 2012 to 2015, he was a Senior Research Associate with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL, USA. He is currently a Professor with the School of Automation, Southeast University, Nanjing, China. His current research interests include cyber–physical systems, wireless networked control, and network security. He was a recipient of the Best Paper Runner-Up Award from ACM MobiHoc in 2014, First Prize of Natural Science Award of Ministry of Education of China in 2017, and Second Prize of Science and Technology Award of Jiangsu Province in 2021. He serves as an Associate Editor for *ACTA Automatica Sinica* and *IEEE Transactions on Industrial Informatics*.

Lin Cai has been with the Department of Electrical & Computer Engineering at the University of Victoria since 2005, and she is currently a Professor. She is a Royal Society of Canada Fellow, an NSERC E.W.R. Steacie Memorial Fellow, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, and an IEEE Fellow. Her research interests span several areas in communications and networking, with a focus on network protocol and architecture design supporting ubiquitous intelligence. She has been elected to serve the board of the IEEE Vehicular Technology Society (2019–2027), and as its VP in Mobile Radio (2023–2025). She has been a Board Member of IEEE Women in Engineering (2022–2024) and IEEE Communications Society (2024–2026). She has served as an Associate Editor-in-Chief for *IEEE Transactions on Vehicular Technology*, and as a Distinguished Lecturer of the IEEE VTS Society and the IEEE Communications Society.

Mo-Yuen Chow received the B.S. degree in electrical and computer engineering from the University of Wisconsin–Madison in 1982 and the M.Eng. and Ph.D. degrees in electrical and computer engineering from Cornell University in 1983 and 1987, respectively. Dr. Chow has been a Professor at Shanghai Jiao Tong University since 2022. He is an Emeritus Professor in the Department of Electrical and Computer Engineering at North Carolina State University. Dr. Chow's recent research focuses on distributed control and management, smart micro-grids, batteries management, and mechatronics systems. Dr. Chow has established the Advanced Diagnosis, Automation, and Control Laboratory. He is an IEEE Fellow, the Co-Editor-in-Chief of *IEEE Transactions on Industrial Informatics* 2014–2018, Editor-in-Chief of *IEEE Transactions on Industrial Electronics* 2010–2012. He has received the IEEE Region-3 Joseph M. Biedenbach Outstanding Engineering Educator Award, the IEEE ENCS Outstanding Engineering Educator Award, the IEEE ENCS Service Award, the IEEE Industrial Electronics Society Anthony J Hornfeck Service Award, and the IEEE Industrial Electronics Society Dr.-Ing. Eugene Mittelmann Achievement Award. He is a Distinguished Lecturer of IEEE Industrial Electronics Society.