Article:
Liu, M. orcid.org/0000-0002-2663-4787, Zhang, X. orcid.org/0000-0002-6063-959X, Zhu, H. orcid.org/0009-0004-3487-7182 et al. (2 more authors) (2024) Physics-aware watermarking embedded in unknown input observers for false data injection attack detection in cyber-physical microgrids. IEEE Transactions on Information Forensics and Security, 19. pp. 7824-7840. ISSN 1556-6013

https://doi.org/10.1109/tifs.2024.3447235

Physics-aware Watermarking Embedded in Unknown Input Observers for False Data Injection Attack Detection in Cyber-Physical Microgrids

Mengxiang Liu, Xin Zhang, Hengye Zhu, Zhenyong Zhang, and Ruilong Deng

***Abstract*—The physics-aware watermarking-based detection method has shown great potential in detecting stealthy False Data Injection Attacks (FDIAs) by adding appropriate watermarks to control commands or sensor measurements, especially in industrial control systems and grid-tied Distributed Energy Resources (DERs). However, existing watermarking-based detection methods have limitations in either handling the intricate physical couplings among DERs or characterising the fast changing power electronics dynamics, and thus cannot be directly applied to microgrids. Inspired by the methodology of Unknown Input Observer (UIO), which can be employed for the distributed anomaly monitoring in microgrids but would be easily bypassed once the adversary has the knowledge of certain electrical parameters, this paper makes the first attempt to investigate the physics-aware watermarking embedded in UIOs such that the stealthy FDIAs would be intentionally disrupted by the watermarking scheme. Based on the theoretical analysis of the detection enhancement and performance degradation under watermarking-enhanced UIOs, the watermark strengths, UIO parameters, and control gains are optimally co-designed to significantly enhance the detection effectiveness while not degrading the control performance. The robustness of the watermarking-enhanced UIO to Time Synchronisation Errors (TSEs) is improved by employing a sliding time window with appropriate length. The performance of the proposed method is validated through Matlab/Simulink studies and cyber-physical co-simulation experiments, and the sensitivities of the detection latency and TSE robustness to watermark strength and detection window's length are comprehensively studied.**

***Index Terms*—False data injection attacks, microgrid, physics-aware watermarking, proactive detection, unknown input observer**

I. Introduction

Microgrids, which collect distributed energy resources (DERs) and loads in a neighboring area and can run in either the grid-tied or isolated mode, have been recognized as a promising approach to manage the massive penetration of DERs into the smart grid [1], [2]. With rapid digitalisation, the microgrid's capabilities to accurately sense DERs' states and rapidly respond to contingencies are continuously increasing [3], [4]. The widely adopted Internet-of-Things (IoT) technologies also expose critical power infrastructure to a multitude of cyber threats. In 2015's infamous BlackEnergy incident, false commands were injected into the grid to disconnect the breaker and denial-of-service (DoS) attacks were launched against customer service lines subsequently to prevent the control centre from knowing exact blackout areas, resulting in Ukraine grid's power outage affecting over the 1,400,000 customers for several hours [5]. The occurrence of the BlackEnergy incident has shown the terrifying consequences of highly coordinated attack actions on power grid and has implied the lack of effective defense strategies against them. More recently, the widely integrated and geographically dispersed DERs have also attracted the adversary's interests. In April 2022, thousands of wind turbines in Europe were forced offline due to the disconnection to the SCADA monitoring centre resulted from a satellite communication disruption attack [6]. It is thus vital to develop local monitoring module for each DER unit to work independently instead of purely relying on the SCADA centre.

Following the three essential data properties of cyber-physical microgrids consisting of availability, integrity, and confidentiality, the cyberattacks can be generally classified as DoS attacks, False Data Injection Attacks (FDIAs), and eavesdropping attacks [7]. To enhance the cyber resilience of microgrids under increasingly intelligent cyberattacks, diverse defensive strategies ranging from pre-attack protection, real-time attack detection and mitigation, as well as post-attack recovery have been investigated [8]. Motivated by the two representative incidents [5], [6], this paper focuses on the real-time detection against FDIAs, whose attack vectors can be designed carefully by the adversary to induce specific consequences while not attracting significant attentions from the system operator.

The history of Intrusion Detection Systems (IDSs) can date back to the 1970s. Since then comprehensive attention has been paid to develop host-level and network-level IDSs by identifying either known attack *signatures* or unusual behaviour *anomalies* [9]. As a suitable complement to the well-developed IDSs analysing host and network features, the newly emerging physical-level IDS that validates the correctness of physical measurements has developed rapidly. This form of IDS works as the last detection line after the host- and network-based IDSs have been bypassed or invalidated. Nevertheless, existing *passive* FDIA detection methods [10]–[15] do not take into account the attack stealthiness, under which these detectors can be easily invalidated by the intelligent adversary after obtaining enough knowledge of the system model and detection principles.

M. Liu and X. Zhang are with the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, UK (e-mails: {mengxiang.liu, xin.zhang1}@sheffield.ac.uk).

H. Zhu, Z. Zhang, and R. Deng are with the State Key Laboratory of Industrial Control Technology and the College of Control Science and Engineering, Zhejiang University, Hangzhou, China (e-mails: {zhuhycse, zhangzhenyong, dengruilong}@zju.edu.cn).

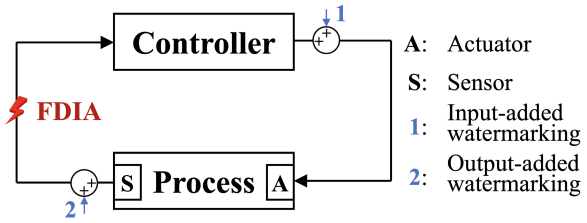


Fig. 1: The input- and output-added watermarking schemes are graphically illustrated within the general cyber-physical control loop, which are both devised against the FDIAs on sensing channels.

To further address the threat of FDIAs with strong stealthiness, the *proactive* detection method, which first actively adds perturbations to model configurations and system inputs/outputs and then observes the corresponding alternations to uncover stealthy FDIAs, has recently received much attention. By appropriately perturbing the reactance of power transmission lines, this typical proactive detection method against the FDIA affecting state estimation has been thoroughly investigated in terms of its effectiveness [16], hiddenness [17], and robustness [18]. However, the reactance perturbation method has not been extended to microgrids due to its high cost of deploying a huge number of D-FACTS devices with reactance perturbation functionality into the distribution network. Taking advantage of the re-programmable ability of converter controller, Liu *et al.* introduced the idea of proactive detection in microgrids by perturbing the control gains either in a fixed period [19] or through an event trigger [20]. Although the converter-based proactive detection method's feasibility and effectiveness have been fully demonstrated, the perturbation on the control gain is still too aggressive from the perspective of system control and operation, which may limit its applications to realistic industrial scenarios.

The watermarking-based proactive detection method is a promising solution to trade off detection and control performance, which can be classified as input-added [21]–[24] and output-added [25], [26] according to the type of objects that the watermark is added to as implied in Fig. 1. Despite this impressive progress, there still exist nontrivial research gaps, which need to be filed, before a feasible and effective watermarking scheme can be directly applied to microgrids: 1) The input-added watermarking method is usually equipped with a Kalman filter to be compatible with the χ^2 hypothesis test [21]–[24]. However, it is difficult to implement a Kalman filter in microgrids to accomplish decentralised state estimation and anomaly monitoring within each DER given the intricate physical couplings among DERs. 2) The output-added watermarking methods [25], [26] adopt a linear-prediction-based method to validate data correctness, which, however, may fail to characterise the complex and rapidly changing power electronic dynamics of DERs.

The Unknown Input Observer (UIO), which is initially designed to estimate the system states in the presence of unknown parameter uncertainties or other external disturbances for fault diagnosis [27], has been recently applied to the distributed monitoring of malicious cyber activities in microgrids [13], [14], [28]. In particular, the UIO's capability in dealing with unknown disturbances makes it possible to treat the intricate physical couplings among DERs as unknown inputs, under which the DER state can be estimated in a fully distributed manner. To improve the UIO's detection ability against replay attacks, Gallo *et al.* introduced a *special* output-added watermarking scheme that requires additional de-watermarking actions before inputting to the controller [29]. However, this special output-added watermarking scheme may not meet the hard real-time requirement of the primary controller that needs to regulates pulse-width modulation signals for power electronic devices within milliseconds. As clearly demonstrated in TABLE I, existing watermarking schemes cannot well satisfy the microgrid's requirements in terms of decentralisation scalability, modelling fidelity, and real-time guarantee. Therefore, it still requires substantial effort to facilitate the organic integration of UIOs and the output-added watermarking scheme, which shows great potentials in trading off the detection and control performance while satisfying the hard real-time requirement in microgrids.

TABLE I: Comparisons between existing literature and the proposed physics-aware watermarking scheme

Category	Literature	Decentral. scalability	Modelling fidelity	Real-time guarantee
Input-added	[21]–[24]	✗	✓	✓
Output-added	[25], [26]	✓	✗	✓
	[29]	✓	✓	✗
Output-added,	This work	✓	✓	✓

To this end, this paper aims to propose an innovative physics-aware watermarking scheme that can greatly enhance the UIO's detectability against the stealthy FDIAs in microgrids. Two fundamental research challenges centering around the theoretical analysis of detection enhancement and performance degradation as well as the optimal co-design of watermark strengths, UIO parameters, and control gains considering these two aspects will be sorted out in detail. The physics-aware characteristic of watermarking is reflected in the design of watermarking strength that incorporates the DER model and control information, which has great importance as it can better trade off the detection and control performance compared with the random watermarking scheme. The main contributions are summarised as follows:

- Physics-aware watermarking embedded in the UIO is investigated for the first time to detect highly stealthy FDIAs in microgrids, with the requirements on decentralisation scalability, modelling fidelity, and real-time guarantee being comprehensively resolved compared with the related works [21]–[26], [29].
- The detection capability enhancement and control performance degradation under the watermarking-enhanced UIO are theoretically analysed and quantified, which are then incorporated into the co-design optimisation problem of watermark strengths, UIO parameters, and control gains.
- In recognition of the voltage-regulation and current-sharing objectives of the primary and secondary con-

trollers in microgrids, the watermarks with appropriate strengths are strategically added to current or voltage measurements, respectively, to eliminate the control performance degradation, while improving the overall detection capability against the FDIAs compromising both voltage and current measurements by utilising their intra-physical couplings.
- The robustness of watermarking-enhanced UIO to Time Synchronisation Errors (TSEs) is improved by incorporating a sliding time window with appropriate length. The effectiveness and performance of the proposed method are thoroughly validated and tested through Matlab/Simulink studies and cyber-physical co-simulation experiments.

TABLE II: Illustration of Main Notations

System Model Parameters	
R_{ti}, L_{ti}, C_{ti}	RLC filter parameters
Z_{Li}, I_{Li}	Equivalent resistive and current loads
$A_{ii}, \boldsymbol{b}_i, \boldsymbol{m}_i$	System matrix, Known and unknown input parameters
$\boldsymbol{g}_i^P, g_i^I$	Primary control proportional and integral gains
T_i^p, F_i^p	System parameters of UIO$_i^p$
T_j^s, F_j^s	System parameters of UIO$_{i,j}^s$
System Model Variables	
V_i, I_{ti}	PCC voltage, Output current
$u_i(V_{ti}), \alpha_i$	Primary control input, Secondary control input
$d_i, \boldsymbol{\omega}_i, \boldsymbol{\rho}_i$	Unknown input term and system noises
$\boldsymbol{x}_i, \boldsymbol{y}_i^p, \boldsymbol{y}_{i,j}^s$	System state, Local and interaction output vectors
$\boldsymbol{\phi}_i, \boldsymbol{\phi}_{i,j}$	P-FDIA and S-FDIA vectors
$\boldsymbol{r}_i^p, \boldsymbol{r}_{i,j}^s$	Detection residuals for P-FDIA and S-FDIA
$\bar{\boldsymbol{r}}_i^p, \bar{\boldsymbol{r}}_{i,j}^s$	Detection thresholds for P-FDIA and S-FDIA
$\varrho_i^p, \varrho_{i,j}^s$	Detection alarms for P-FDIA and S-FDIA
Watermark-involved Variables	
$\boldsymbol{w}_i^p, \boldsymbol{w}_{i,j}^s$	Watermarks added to local and interaction measurements
$\boldsymbol{w}_i^{pa}, \boldsymbol{w}_{i,j}^{sa}$	Removed watermarks under P-FDIA and S-FDIA
$\boldsymbol{y}_i^{pw}/\boldsymbol{y}_{i,j}^{psw}$	Watermarked local/interaction measurements w/o attacks
$\boldsymbol{y}_i^{pwa}, \boldsymbol{y}_{i,j}^{pswa}$	Watermarked measurements under P-FDIA and S-FDIA
$\boldsymbol{y}_i^{pra}, \boldsymbol{y}_{i,j}^{psra}$	De-watermarked measurements under P-FDIA and S-FDIA
$\boldsymbol{\sigma}_i^p/\boldsymbol{\sigma}_{i,j}^s$	Watermarking strengths on local/interaction measurements

II. Related Works

Numerous *passive* attack detection methods have been proposed against the FDIA in microgrids. To detect and identify FDIAs against load frequency control, Ameli *et al.* [10] proposed a novel attack detection method by developing a stochastic unknown input estimator, which is designed to be decoupled from unknown load changes and is set with optimal gains such that the impacts of system noises could be ignored. Considering the defense cost, Yang *et al.* [11] investigated the optimal placement of phasor measurement units and proposed an effective greedy algorithm to find the minimal set of protected buses under which any FDIAs affecting smart grid state estimation could be detected. Besides the aforementioned model-based methods, Li *et al.* [12] built a dirichlet-based probabilistic model with adaptive incentive mechanism using behaviour rule specifications to detect financially-motivated opportunistic attacks in smart grid. After recognising the limitation of UIO and Luenberger-like observer in dealing with unknown disturbances resulting from parameter variations, Tan *et al.* [13], [14] applied the parity-based method to develop a robust attack detection scheme. Given the FDIA against distributed tertiary control layer, Liu *et al.* [15] devised an alternative-data-based detection scheme by identifying the consistency between system outputs and alternative communication data. Based on the noise fingerprint extracted from sensor readings in the normal operation stage, Ahmed *et al.* [30] proposed an innovative model-data-blending scheme to detect the powerful adversaries in water treatment facilities.

As for the watermarking based *proactive* detection, the input-added type which directly adds watermarks to control signals has received the most attention. Mo *et al.* [21] pioneered the input-added watermarking method by adding watermarks with specific distributions to system inputs before forwarding to actuators, and a Kalman filter was employed to validate the existence of watermark-related terms in measurement outputs. Any inconsistencies indicate that the measurement outputs may suffer from the replay attack, which is a typical stealthy FDIA. To cost-efficiently detect the transient covert attack, whose active period is substantially shorter than its sleep period, Ma *et al.* [22] proposed a event-triggered and recursive watermarking strategy that can save performance loss and guarantee detection rate. With the aim of relieving the dependence on the knowledge of noise variance and system model, Hyun *et al.* [23] integrated an auxiliary test criterion with the input-added watermarking scheme and validated its feasibility on a grid-tied photovoltaic system. Through utilisation of the cumulative sum detector in combination with dynamic watermarking, Li *et al.* successfully minimised the detection latency while ensuring low false alarms [24].

On the other hand, instead of observing the input-added watermark's impacts on measurement outputs, the output-added watermarking scheme initially adds watermarks to measurement outputs before sending out to the controller through data transmitter. After receiving the watermarked data, the data receiver will first forward it to the controller to meet the hard real-time requirement, and then implement the de-watermarking process for anomaly detection purpose [25]. When there exist FDIAs between the data transmitter and receiver, the added and removed watermarks would not be the same and thus it would be possible to reveal the *disrupted* stealthy bias injection. To balance the trade-off between detection effectiveness and control performance, Zhu *et al.* [26] investigated the optimal design of watermark strength considering system noises, signal quality, and detection latency. By adding micro-distortions into the sensor readings through either digital or physical ways, Sourav *et al.* [31] presented a mean-difference method to accurately and timely disclose the hidden attackers in industrial control systems. The idea of adding micro-distortions shares strong similarity with the output-added watermarking scheme, which both rely on the extra disturbances to reveal stealthy attack activities, but the adopted mean-difference-based attack detection method may be limited to fully address the microgrid's requirements in terms of decentralisation scalability, modelling fidelity, and real-time guarantee.

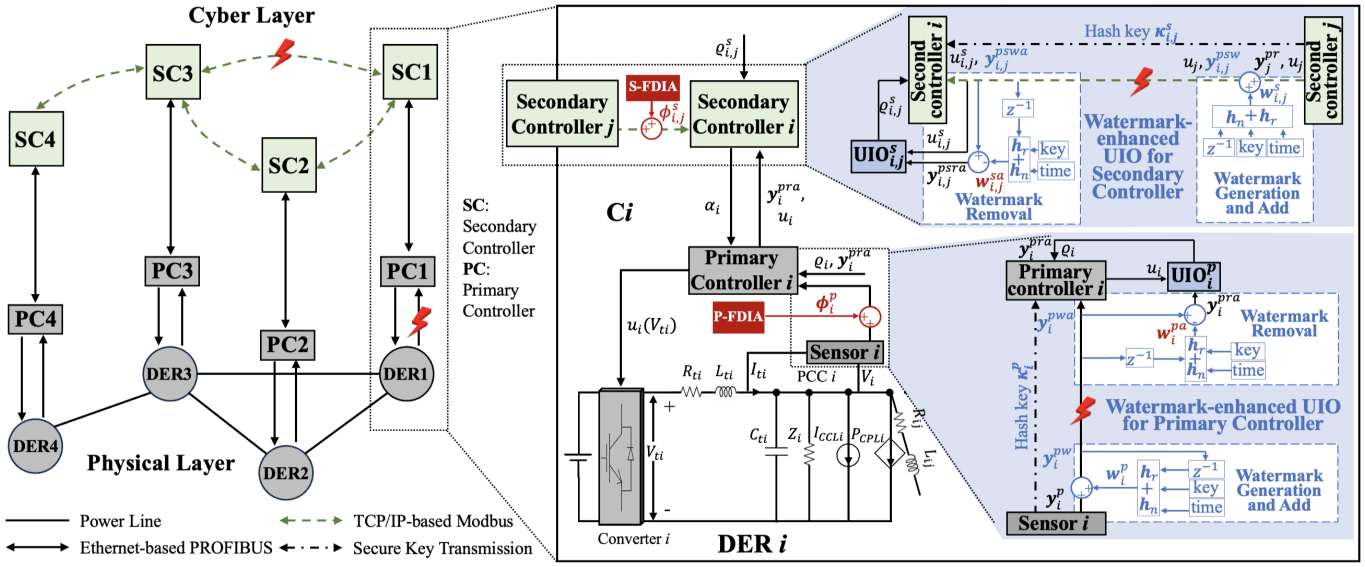


Fig. 2: In this figure, the left part depicts the cyber-physical architecture of microgrid with primary and secondary controllers, the middle part illustrates the detailed cyber-physical couplings and exposed attack surfaces, and the right part shows the integration principles of UIO and physics-aware watermarking schemes.

III. Preliminary System Models

A. Cyber-Physical Microgrid Model

We consider a isolated DC microgrid composed of $N \geq 2$ DERs, where the buck converter is commanded to supply the local ZIP load connected to the Point of Common Coupling (PCC) bus as shown in Fig. 2. The electrical network of DERs is denoted by a weighted undirected graph $\mathcal{G}_{el} = \{\mathcal{A}, \mathcal{E}_{el}\}$, where $\mathcal{A}$ is the set of DERs and $\mathcal{E}_{el}$ is the set of power lines connecting them. Specifically, DERs i and j are neighbors if power line $\{i, j\} \in \mathcal{E}_{el}$, and the set of neighbors of DER i in $\mathcal{G}_{el}$ is represented by $\mathcal{N}_i^{el}$. Moreover, the weight of $\{i, j\}$ is the branch conductance, which is denoted by $\frac{1}{R_{ij}}$.

To address the nonlinear issue induced by the introduction of Constant Power Load (CPL) P_{CPLi}, it is linearized around the nominal reference PCC voltage $V_{ref,i}$ as a negative impedance (part a) as well as a constant current (part b), i.e.,

$$I_{CPLi} = \underbrace{-\frac{P_{CPLi}}{V_{ref,i}^2} V_i}_{\text{part } a} + \underbrace{2\frac{P_{CPLi}}{V_{ref,i}}}_{\text{part } b}, \tag{1}$$

where V_i is the i-th PCC voltage and I_{CPLi} is the total current of the linearized CPL. Combined with the original Constant Impedance Load (CIL) Z_i and Constant Current Load (CCL) I_{CCLi}, the linearized equivalent load model can be obtained:

$$\begin{cases} \frac{1}{Z_{Li}} = \frac{1}{Z_i} - \frac{P_{CPLi}}{V_{ref,i}^2} \\ I_{Li} = I_{CCLi} + 2\frac{P_{CPLi}}{V_{ref,i}} \end{cases}. \tag{2}$$

By applying the Kirchhoff voltage and current laws and exploiting the quasi-stationary line approximation (i.e., $L_{ij} \approx 0$) [32], the electrical model of the RLC filter within DER i follows

$$\begin{cases} \frac{dV_i}{dt} = \frac{1}{C_{ti}} I_{ti} + \sum\limits_{j \in \mathcal{N}_i^{el}} \frac{1}{C_{ti} R_{ij}} (V_j - V_i) - \frac{1}{C_{ti}} (I_{Li} + \frac{V_i}{Z_{Li}}) \\ \frac{dI_{ti}}{dt} = -\frac{1}{L_{ti}} V_i - \frac{R_{ti}}{L_{ti}} I_{ti} + \frac{1}{L_{ti}} V_{ti} \end{cases}. \tag{3}$$

Let state $\boldsymbol{x}_i(t) = [V_i(t), I_{ti}(t)]^{\mathrm{T}}$, and then (3) can be rewritten as the state-space model:

$$\dot{\boldsymbol{x}}_i(t) = A_{ii} \boldsymbol{x}_i(t) + \boldsymbol{b}_i u_i(t) + \boldsymbol{m}_i d_i(t), \tag{4}$$

where the dynamical model parameters $A_{ii}, \boldsymbol{b}_i$, and $\boldsymbol{m}_i$ are

$$A_{ii} = \begin{bmatrix} -\frac{1}{Z_{Li} C_{ti}} - \sum\limits_{j \in \mathcal{N}_i^{el}} \frac{1}{C_{ti} R_{ij}} & \frac{1}{C_{ti}} \\ -\frac{1}{L_{ti}} & -\frac{R_{ti}}{L_{ti}} \end{bmatrix}, \boldsymbol{b}_i = \begin{bmatrix} 0 \\ \frac{1}{L_{ti}} \end{bmatrix}$$
$$\boldsymbol{m}_i = \begin{bmatrix} -\frac{1}{C_{ti}} \\ 0 \end{bmatrix}, \tag{5}$$

and the unknown input term $d_i(t) = I_{Li} + \sum_{j \in \mathcal{N}_i^{el}} -\frac{1}{R_{ij}} V_j(t)$ includes the equivalent current load as well as neighboring PCC voltages. Considering bounded system noises and fully measured system states, (4) is transformed into the following discrete-time form

$$\begin{cases} \boldsymbol{x}_i(k+1) = A_{ii}^d \boldsymbol{x}_i(k) + \boldsymbol{b}_i^d u_i(k) + \boldsymbol{m}_i^d d_i(k) + \boldsymbol{\omega}_i(k), \\ \boldsymbol{y}_i(k) = \boldsymbol{x}_i(k) + \boldsymbol{\rho}_i(k), \end{cases} \tag{6}$$

where $A_{ii}^d, \boldsymbol{b}_i^d$, and $\boldsymbol{m}_i^d$ are discrete-time system parameters and their relations with the continuous-time system parameters are

$$A_{ii}^d = e^{A_{ii} T_{samp}}, Y_{ii}^d = (A_{ii})^{-1} (A_{ii}^d - \mathrm{I}^2),$$
$$\boldsymbol{b}_i^d = Y_{ii}^d \boldsymbol{b}_i, \boldsymbol{m}_i^d = Y_{ii}^d \boldsymbol{m}_i \tag{7}$$

where T_{samp} is the sampling time. The bounded process and measurement noises satisfy $\boldsymbol{\omega}_i(k) \leq \bar{\boldsymbol{\omega}}_i$ and $\boldsymbol{\rho}_i(k) \leq \bar{\boldsymbol{\rho}}_i$, respectively.

B. Primary and Secondary Controllers

For the local measurements received from sensor, we have $\boldsymbol{y}_i^p(k) = \boldsymbol{y}_i(k)$ in the attack-free case. A typical Proportional-Integral (PI) control strategy is adopted to compute the primary control input as

$$u_i(k) = (\boldsymbol{g}_i^P)^{\mathrm{T}} \boldsymbol{y}_i^p(k) + g_i^I \sum_{l=0}^{k} \left(V_{ref,i} + \alpha_i(k) - \boldsymbol{\iota}^{\mathrm{T}} \boldsymbol{y}_i^p(l) \right), \quad (8)$$

where constant vector $\boldsymbol{\iota} = [1, 0]^{\mathrm{T}}$. Similarly, for the interaction measurements from neighboring DER j, we have $\boldsymbol{y}_{i,j}^s(k) = \boldsymbol{y}_j(k)$ in the normal case. To achieve load sharing among DERs, a consensus scheme is employed to calculate the secondary control input

$$\alpha_i(k) = \boldsymbol{\gamma}_i^{\mathrm{T}} \sum_{l=0}^{k} \sum_{j \in \mathcal{N}_i^c} a_{ij}^c \left(\frac{\boldsymbol{y}_{i,j}^s(l)}{I_{tj}^s} - \frac{\boldsymbol{y}_i^p(l)}{I_{ti}^s} \right), \quad (9)$$

where $\boldsymbol{\gamma}_i = [0, \gamma_i]^{\mathrm{T}}$ is the consensus gain, and $I_{ti}^s > 0$ and $I_{tj}^s > 0$ are the rated currents corresponding to DERs i and j, respectively. The communication network among DERs is denoted by a weighted undirected graph $\mathcal{G}_c = \{\mathcal{A}, \mathcal{E}_c\}$, where the set $\mathcal{E}_c$ collects all communication links and the weight of link $\{i, j\} \in \mathcal{G}_c$ is denoted by a_{ij}^c. The set $\mathcal{N}_i^c$ includes the neighbors of DER i interconnected in the communication network.

C. False Data Injection Attack Model

The wide spread adoption of IoT technologies has significantly advanced the digitalisation of cyber-physical microgrids, thus enabling the real-time and accurate control actions, which, however, also exposes each DER's local measurements and the interacted data among DERs to cyber threats. The system operator may bring smart phones, tablets, and mobile devices to the field and connect to primary controllers as shown in Fig. 3 [25]. Intelligent edge-sensors are also expected to play important roles like uploading local measurements to industrial cloud to enable big data based applications [33]. Based on the disclosed vulnerabilities of Ethernet-based field communication protocols including Profinet [34] and Modbus [35], a potential P-FDIA path is specified in Fig. 3.

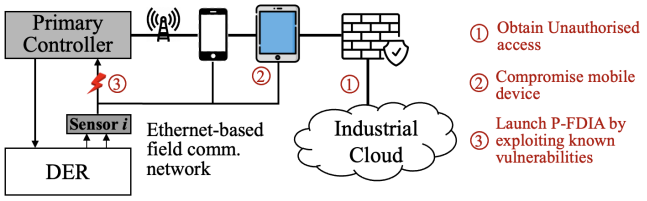


Fig. 3: A potential P-FDIA path is illustrated in the highly digitalised microgrid by following ① Obtain unauthorised access via industrial cloud, ② Compromise mobile devices, and ③ Launch P-FDIA by exploiting known vulnerabilities.

In this paper, we assume that the adversary has gained access to the microgrid's field control network and is able to inject desired biases into the communication packets to affect the primary and secondary control functions. The man-in-the-middle attack is a typical way to implement bias injection into communication data packets [36]. Despite the existence of mature detection schemes from Information Technology (IT) domain against man-in-the-middle attacks, such as authentication [37], these schemes could be ineffective when the adversary is able to spoof certificates [38] or the system operator does not apply security configurations as expected to meet the real-time control requirement in the scale of milliseconds in microgrids. Hence, we consider the worst case where all IT security mechanisms have been bypassed, under which it would rely purely on the information and knowledge from operational technology to complete the data integrity check.

In the primary control loop, the received local measurements from sensors under attacks are modelled as

$$\boldsymbol{y}_i^p(k) = \boldsymbol{y}_i(k) + \boldsymbol{\phi}_i(k), \quad (10)$$

where $\boldsymbol{\phi}_i(k)$ signifies the Primary FDIA (P-FDIA) vector arbitrarily designed by the adversary. In the secondary control layer, the interactive measurement output received from the neighboring DER under attacks is modelled as

$$\boldsymbol{y}_{i,j}^s(k) = \boldsymbol{y}_j(k) + \boldsymbol{\phi}_{i,j}(k), \quad (11)$$

where $\boldsymbol{\phi}_{i,j}(k)$ is the designed Secondary FDIA (S-FDIA) vector. Although the P-FDIA (10) and S-FDIA (11) have similar forms, the reasons for considering them independently are three-fold: 1) The requirement of implementing P-FDIA is higher than S-FDIA as it requires access to the network closer to the field level, 2) As implied by the demonstrative example's results in Fig. 4, the S-FDIA can directly destabilise the voltage, which can quickly trigger the action of protection device. By contrary, the impact of the P-FDIA, i.e., steady-state voltage and current deviations, is more imperceptible and can be even more severe in the long-term, 3) The design of the watermarking-based P-FDIA and S-FDIA defense schemes is different as the primary controller (8) is voltage-based while the secondary controller (9) is current-oriented, and 4) From the perspective of attack detection, the co-existence of P-FDIA and S-FDIA could be significantly different from the single attack type as demonstrated in Section VI-A-5).

D. UIO-based Detector

To perceive P-FDIAs and S-FDIAs, two kinds of detectors will be deployed in the primary and secondary control layers based on the UIO methodology. The UIO methodology was first proposed to estimate system states in the presence of unknown disturbances such as faults [27], and has been recently extended to large-scale interconnected systems [28] such as microgrids, where the current loads within DERs and the physical couplings with neighboring DERs are typically unknown, for distributed attack detection as indicated by (6). The principle of utilising UIOs for attack detection is first to estimate the system state from the observed measurements $(\boldsymbol{y}_i^p, \boldsymbol{y}_{i,j}^s)$, based on which detection residuals are then calculated by comparing the observed measurements with the estimated ones $(\hat{\boldsymbol{x}}_i^p, \hat{\boldsymbol{x}}_{i,j}^s)$. Any unexpected high detection residual indicates that the observed data does not conform to the physical dynamics and an anomaly may exist.

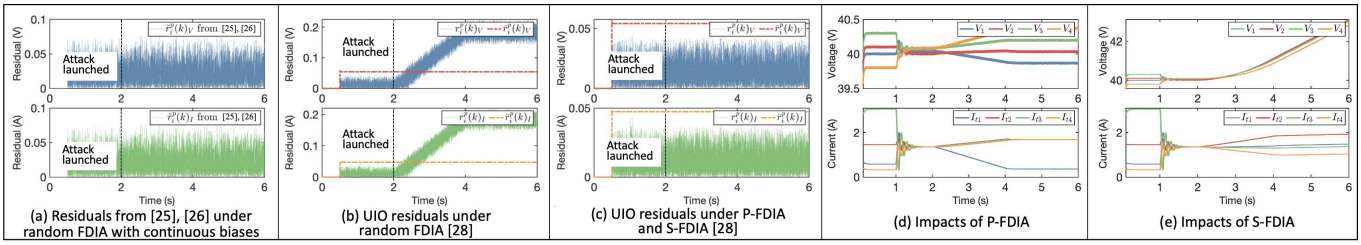


Fig. 4: The deficiency of linear-prediction-based and UIO-based detectors are showcased through examples, where sub-figure (a) shows that random FDIA can easily deceive the linear-prediction-based detector utilised in [25], [26], sub-figure (b) indicates that the UIO-based detector can perceive random FDIAs, sub-figure (c) illustrates the stealthiness of P-FDIA and S-FDIA under legacy UIO-based detectors [28], and sub-figures (d), (e) demonstrate these stealthy attacks' impacts on PCC voltages and output currents.

In the primary control layer, considering the DER dynamics (6), UIO_i^p is employed, as shown in Fig. 2, to validate the integrity of $\boldsymbol{y}_i^p$ and is constructed as

$$\text{UIO}_i^p \begin{cases} \boldsymbol{z}_i^p(k+1) = F_i^p \boldsymbol{z}_i^p(k) + T_i^p \boldsymbol{b}_i^d u_i(k) + \hat{K}_i^p \boldsymbol{y}_i^p(k) \\ \hat{\boldsymbol{x}}_i^p(k+1) = \boldsymbol{z}_i^p(k+1) + H_i^p \boldsymbol{y}_i^p(k+1) \end{cases}, \quad (12)$$

where $\boldsymbol{z}_i^p(k)$ is UIO_i^p's internal state, and the UIO parameters $F_i^p, \hat{K}_i^p$, and H_i^p should satisfy (13)-(17), i.e.,

$$T_i^p = \mathrm{I}^2 - H_i^p, \quad (13)$$
$$T_i^p \boldsymbol{m}_i^d = \mathbb{0}^{3\times 1}, \quad (14)$$
$$\hat{K}_i^p = K_{i1}^p + K_{i2}^p \quad (15)$$
$$F_i^p = T_i^p A_{ii}^d - K_{i1}^p, \quad (16)$$
$$K_{i2}^p = F_i^p H_i^p, \quad (17)$$

to make the estimated state $\hat{\boldsymbol{x}}_i^p$ converge to the actual one $\boldsymbol{x}_i$. The anomaly is reflected on the detection residual $\boldsymbol{r}_i^p(k) = \boldsymbol{y}_i^p(k) - \hat{\boldsymbol{x}}_i^p(k)$, which, in the attack-free case, satisfies

$$\boldsymbol{r}_i^p(k) = (F_i^p)^k \big(\boldsymbol{\epsilon}_i^p(0) + H_i^p \boldsymbol{\rho}_i(0)\big) + T_i^p \boldsymbol{\rho}_i(k) \sum_{l=0}^{k-1} \big(F_i^p\big)^{k-1-l} \big(T_i^p \boldsymbol{\omega}_i(l) - \hat{K}_i^p \boldsymbol{\rho}_i(l)\big), \quad (18)$$

where $\boldsymbol{\epsilon}_i^p(0) = \boldsymbol{x}_i(0) - \hat{\boldsymbol{x}}_i^p(0)$ denotes the initial state estimation error, which can be bounded by the measurement noise's bound, i.e., $|\boldsymbol{\epsilon}_i^p(0)| \leq \bar{\boldsymbol{\rho}}_i$, once $\boldsymbol{z}_i^p(0)$ is set as $T_i^p \boldsymbol{y}_i^p(0)$. Since matrix F_i^p is stable, i.e., all of its eigenvalues lie inside the unit circle, it is possible to find positive scalars ν_i^p and $0 < \varsigma_i^p < 1$ such that $||(F_i^p)^k|| \leq \nu_i^p (\varsigma_i^p)^k$. Hence, the detection residual's upper bound $\bar{\boldsymbol{r}}_i^p(k)$ in the attack-free case can be written as

$$|\boldsymbol{r}_i^p(k)| \leq \bar{\boldsymbol{r}}_i^p(k) = \nu_i^p (\varsigma_i^p)^k |\mathrm{I}^2 + H_i^p| \bar{\boldsymbol{\rho}}_i + |T_i^p| \bar{\boldsymbol{\rho}}_i \sum_{l=0}^{k-1} \nu_i^p (\varsigma_i^p)^{k-1-l} \big(|T_i^p| \bar{\boldsymbol{\omega}}_i + |\hat{K}_i^p| \bar{\boldsymbol{\rho}}_i\big). \quad (19)$$

In the secondary control layer, to validate the integrity of $\boldsymbol{y}_{i,j}^s(k)$ through $\text{UIO}_{i,j}^s$ as shown in Fig. 2, DER j needs to transmit its primary control input to DER i denoted by $u_{i,j}^s(k)$, in addition to $\boldsymbol{y}_{i,j}^s(k)$. Similarly, $\text{UIO}_{i,j}^s$ is constructed as

$$\text{UIO}_{i,j}^s \begin{cases} \boldsymbol{z}_{i,j}^s(k+1) = F_j^s \boldsymbol{z}_{i,j}^s(k) + T_j^s \boldsymbol{b}_j^d u_{i,j}^s(k) + \hat{K}_j^s \boldsymbol{y}_{i,j}^s(k) \\ \hat{\boldsymbol{x}}_{i,j}^s(k+1) = \boldsymbol{z}_{i,j}^s(k+1) + H_j^s \boldsymbol{y}_{i,j}^s(k+1) \end{cases}, \quad (20)$$

where $\boldsymbol{z}_{i,j}^s(k)$ is $\text{UIO}_{i,j}^s$'s internal state, and UIO parameters $F_j^s, T_j^s, \hat{K}_j^s$, and H_j^s need to similarly satisfy (21)-(25), i.e.,

$$T_j^s = \mathrm{I}^2 - H_j^s, \quad (21)$$
$$T_j^s \boldsymbol{m}_j^d = \mathbb{0}^{3\times 1}, \quad (22)$$
$$\hat{K}_j^s = K_{j1}^s + K_{j2}^s \quad (23)$$
$$F_j^s = T_j^s A_{jj}^d - K_{j1}^s, \quad (24)$$
$$K_{j2}^s = F_j^s H_j^s, \quad (25)$$

to converge the state estimation error. The detection residual $\boldsymbol{r}_{i,j}^s(k) = \boldsymbol{y}_{i,j}^s(k) - \hat{\boldsymbol{x}}_{i,j}^s(k)$ and the corresponding threshold $\bar{\boldsymbol{r}}_{i,j}^s(k)$ can be computed in a similar way to (18) and (19), respectively, but is omitted here for simplification. The two kinds of UIOs' alarming principles are synthesised as

$$\text{UIO}_i^p : \boldsymbol{r}_i^p(k) > \bar{\boldsymbol{r}}_i^p(k) \Rightarrow \varrho_i^p = 1, \quad (26)$$
$$\text{UIO}_{i,j}^s : \boldsymbol{r}_{i,j}^s(k) > \bar{\boldsymbol{r}}_{i,j}^s(k) \Rightarrow \varrho_{i,j}^s = 1, \quad (27)$$

where the two alarming signals of ϱ_i^p and $\varrho_{i,j}^s$ are sent to the primary and secondary controllers, respectively, for the following-up impact mitigation action.

IV. Deficiency of Existing Anomaly Detectors and Watermarking Enabled Enhancement

A. Deficiency of Linear-prediction- and UIO-based Detectors

Considering the decentralisation scalability requirement of microgrids, this paper mainly analyses the linear-prediction-based [25], [26] and UIO-based [28] detectors. The linear-prediction-based detector directly predicts the next state through the historical linear relation, i.e.,

$$\hat{\hat{\boldsymbol{x}}}_i^p(k+1) = \boldsymbol{y}_i^p(k) + \boldsymbol{y}_i^p(k) - \boldsymbol{y}_i^p(k-1), \quad (28)$$

where $\hat{\hat{\boldsymbol{x}}}_i^p$ denotes the predicted local system states and the associated residual is $\tilde{\boldsymbol{r}}_i^p(k) = \boldsymbol{y}_i^p(k) - \hat{\hat{\boldsymbol{x}}}_i^p(k)$. Despite the linear-prediction process (28)'s ultra-low complexity, its fatal deficiency comes from the self that its simplicity makes it unable to capture the spatial relations between voltage and current variables that reflect the rapidly changing power electronic dynamics. In such case, any random continuous bias injection can easily deceive the linear-prediction-based detector.

Although the legacy UIO-based detectors UIO_i^p and $\text{UIO}_{i,j}^s$ are able to perceive most simplistic cyberattacks, such as random ones, it has been disclosed recently that these purely

model-based detectors can be easily bypassed once the adversary has access to the state-space model parameters $A_{ii}^d, \boldsymbol{m}_i^d$ and $A_{jj}^d, \boldsymbol{m}_j^d$ determined by electrical parameters [39]. In particular, these stealthy P-FDIA and S-FDIA vectors are constructed as

$$\begin{cases} \boldsymbol{\phi}_i^p(k+1) = A_{ii}^d \boldsymbol{\phi}_i^p(k) + \boldsymbol{m}_i^d d_i^a(k) \\ \boldsymbol{\phi}_i^p(k_i^p) = \mathbb{0}^{2\times 1}, k \geq k_i^p \end{cases}, \tag{29}$$

and

$$\begin{cases} \boldsymbol{\phi}_{i,j}^s(k+1) = A_{jj}^d \boldsymbol{\phi}_{i,j}^s(k) + \boldsymbol{m}_j^d d_{i,j}^a(k) \\ \boldsymbol{\phi}_{i,j}^s(k_{i,j}^s) = \mathbb{0}^{2\times 1}, k \geq k_{i,j}^s \end{cases}, \tag{30}$$

respectively, where $d_i^a(k)$ and $d_{i,j}^a$ denote the faked unknown inputs and k_i^p and $k_{i,j}^s$ signify the attack launch time instants. The intuitive idea of constructing FDIAs as (29)-(30) is to mimic the DER dynamics (6) such that UIO_i^p and $\text{UIO}_{i,j}^s$ cannot distinguish the injected biases from normal data and the resulted detection residuals would be the same as those under normal operations.

Demonstrative examples are given in Fig. 4 to illustrate the above deficiencies of existing anomaly detectors. In the 4-DER microgrid with its cyber and physical topologies being setup according to Fig. 2, the linear-prediction-based detector [25], [26] and legacy UIO-based detector [28] are activated at $t = 0.5$s, the secondary controllers (9) are started at $t = 1$s, and the P-FDIA against DER 1's local measurements and the S-FDIA against DER 1's transmitted measurements to DER 2 are launched at $t = 2$s. When the random attack with ramp bias $\phi_1(k) = 0.75(kT_{samp} - 2), kT_{samp} \in [2, 4]$ is launched, the linear-prediction-based detector cannot observe any anomaly while the UIO-based detector can present the abnormal residuals that rise over the detection thresholds as shown in sub-figures (a) and (b), respectively. After the adversary has some understanding of the DER dynamics, the FDIA vectors can be constructed as (29) and (30) by mimicking the DER dynamics, under which the detection residuals will remain unaltered as those under normal operations as shown in sub-figure (c). Since the P-FDIA compromises both primary and secondary controllers, it can deceive other DERs to undertake more loads while DER 1 only supplies a small portion of loads. As shown in sub-figure (d), after the attack, DER 1 only supplies about 0.35A load while the remaining DERs need to undertake about 1.7A load each. The S-FDIA only compromises the secondary controller's input and directly makes the consensus of current sharing not achievable, which can quickly drive the microgrid's operation status to an unsafe zone and easily trigger the actions of protection devices as in sub-figure (e). The results indicate that by launching P-FDIAs, the adversary can gain accurate and specific profits in an unobtrusive manner, while the S-FDIA can quickly destabilise the microgrid.

B. Physics-Aware Watermarking Enabled Enhancement

Based on the above exemplifying results, the standalone linear-prediction-based detector cannot characterise the complex and rapidly changing power electronic dynamics. Therefore, in this paper's subsequent parts, the focus is on the UIO-based detector. By strategically integrating the output-added Recursive Watermarking (RWM) scheme [25], it is possible to significantly enhance the UIO's detection capability against stealthy P-FDIAs and S-FDIAs, while ensuring real-time control adherence in microgrids. Specifically, as shown in Fig. 2, the watermark generated at the data transmitter will be added to the original measurement signal before it is sent out through communication channels. After receiving the watermarked data, it will first be forwarded to controllers to meet the real-time requirement and then the watermark will be removed to retain the original signal such that its integrity can be validated by UIOs. The key advantage of RWM scheme is that it strategically integrates the signal dynamics into watermarks, under which the existence of bias injection would make the added and removed watermarks inconsistent if the adversary does not perfectly know the watermark generation scheme. Then, the communication data will not be tampered with in the expected stealthy manner such that the UIO-based detector may be able to disclose P-FDIAs and S-FDIAs. While in the normal case, the added and removed watermarks are the same and no alarm will be triggered.

1) Watermark Generation and Addition: Since the communicated data $\boldsymbol{y}_i^p$ and $\boldsymbol{y}_{i,j}^s$ have the same structure and both contain voltage and current information, the following analysis only shows the watermarking process of $\boldsymbol{y}_i^p$. Let $\boldsymbol{w}_i^p$, $\boldsymbol{y}_i^{pw}$, and $\boldsymbol{y}_i^{pr}$ denote the watermark, watermarked signal, and de-watermarked signals, respectively. To adhere to the real-time requirement, the watermarks are generated using hashing functions $\boldsymbol{h}_n(\cdot), \boldsymbol{h}_r(\cdot)$ with related hash tables being established in advance and key $\boldsymbol{\kappa}_i^p$ being transmitted through secure channels, i.e.,

$$\boldsymbol{w}_i^p(k) = \begin{cases} \boldsymbol{h}_n(k, \boldsymbol{\kappa}_i^p) + \boldsymbol{h}_r\big(\boldsymbol{y}_i^{pw}(k-1), \boldsymbol{\kappa}_i^p\big), k \geq k_i^{pw} + 1, \\ \mathbb{0}^{2\times 1}, k = k_i^{pw} \end{cases} \tag{31}$$

where $\boldsymbol{h}_n(\cdot)$ and $\boldsymbol{h}_r(\cdot)$ map integer k and real numbers $\boldsymbol{y}_i^{pw}(k-1)$ to real numbers within the range of $[-\frac{\boldsymbol{\sigma}_i^p}{2}, \frac{\boldsymbol{\sigma}_i^p}{2}]$ and are both based on a mature and quick pseudorandom integer sequence-generation algorithm [40] with appropriate inputs/outputs normalization, and k_i^{pw} denotes RWM's activation time. The recursive nature of the watermarking process is demonstrated by including the watermarked data from the previous time step in the calculation of the current watermark, which can introduce internal signal dynamics into the watermark and eliminate the necessity for directly transmitting it. Thus, the generated watermarks can be regarded as a sequence of white noise in the range $[-\boldsymbol{\sigma}_i^p, \boldsymbol{\sigma}_i^p]$ with strength specified by $\boldsymbol{\sigma}_i^p$.

The watermarked signals can be obtained by adding the generated watermarks to the original signal, i.e.,

$$\boldsymbol{y}_i^{pw}(k) = \boldsymbol{y}_i^p(k) + \boldsymbol{w}_i^p(k), k \geq k_i^{pw} + 1. \tag{32}$$

2) Watermark Removal: In the normal case, the added watermark can be successfully removed by the data receiver and the original data will be recovered as

$$\begin{aligned} \boldsymbol{y}_i^{pr}(k) &= \boldsymbol{y}_i^{pw}(k) - \boldsymbol{h}_n(k, \boldsymbol{\kappa}_i^p) - \boldsymbol{h}_r\big(\boldsymbol{y}_i^{pw}(k-1)\big), \boldsymbol{\kappa}_i^p) \\ &= \boldsymbol{y}_i^{pw}(k) - \boldsymbol{w}_i^p(k) = \boldsymbol{y}_i^p(k). \end{aligned} \tag{33}$$

Thus, the adoption of RWM will not falsely trigger the alarm of UIOs. In the presence of P-FDIAs, the removed watermarks

are not consistent with the added ones and thus the recovered signals would not be equal to the sum of the original data and the injected bias, i.e.,

$$\begin{aligned}\boldsymbol{y}_i^{pra}(k) &= \underbrace{\boldsymbol{y}_i^{pw}(k)+\boldsymbol{\phi}_i^p(k)}_{\boldsymbol{y}_i^{pwa}(k)} - \underbrace{\big(\boldsymbol{h}_n(k,\boldsymbol{\kappa}_i^p)+\boldsymbol{h}_r\big(\boldsymbol{y}_i^{pwa}(k-1),\boldsymbol{\kappa}_i^p\big)\big)}_{\boldsymbol{w}_i^{pa}(k)} \\ &\neq \boldsymbol{y}_i^p(k)+\boldsymbol{\phi}_i^p(k). \end{aligned} \tag{34}$$

Considering the RWM's recursive nature, once the watermarked signals at $k-1$ is compromised, i.e., $\boldsymbol{y}_i^{pwa}(k-1) \neq \boldsymbol{y}_i(k-1)+\boldsymbol{w}_i^p(k-1)$, then the reconstructed watermark at the data receiver will not be the same as the generated one, i.e., $\boldsymbol{w}_i^{pa}(k) \neq \boldsymbol{w}_i^p(k)$. Hence, we know that under the RWM, the de-watermarked signals $\boldsymbol{y}_i^{pr}(k)$ may not be able to include the carefully designed P-FDIA vector $\boldsymbol{\phi}_i^p(k)$ completely, under which the deployed UIO_i^p is likely to perceive these incomplete bias injection. Under the extreme case when the adversary knows the generation scheme of RWMs, it is still difficult to totally eliminate the impact of inaccurately reconstructed watermarks once the adversary does not exactly know the hash key $\boldsymbol{\kappa}_i^p$.

Since the primary and secondary controllers have different functionalities and are designed separately in principle, the design of watermarks added to these two controllers are also decoupled. When watermarking the interactive measurements from DERs j to i, the recovered local measurements $\boldsymbol{y}_j^{pr}$ will be utilised as the original data. Then, a process similar to the watermarking process of local measurements can be established for the interactive measurements among DERs, which is clarified in Fig. 2 with the watermark-related variables being fully explained in TABLE II. Nevertheless, the detection of P-FDIAs and S-FDIAs cannot be completely decoupled since the recovered local measurements $\boldsymbol{y}_j^{pra}$ will also trigger the alarm of $\text{UIO}_{i,j}^s$. In other words, the P-FDIAs against DER j will hide the impact of the S-FDIAs against communication links from DER j to its neighbors, and it would be hard to judge the existence of these S-FDIAs when both of the two types of FDIAs exist as validated in Section VI-A-5).

C. Problem Formulation

The main challenge of adopting the watermark-enhanced UIO is to balance the trade-off between detection accuracy and control performance. Normally, when generating the watermark sequence with a larger strength $\boldsymbol{\sigma}_i^p$ or $\boldsymbol{\sigma}_{i,j}^s$, it would be much easier for UIO_i^p or $\text{UIO}_{i,j}^s$ to detect P-FDIAs or S-FDIAs since the inconsistently generated watermarks by the data transmitter and receiver can have a large impact on the injected biases. However, when the watermark signal is strong, it can deteriorate the control performance or even destabilize the microgrid if it is not deigned appropriately. Hence, in this paper, we aim to address the optimal co-design problem of watermark strengths, UIO parameters, and control gains such that the watermark-enhanced UIO's detection capability against P-FDIAs and S-FDIAs can be significantly enhanced while making the resulting control performance degradation negligible.

V. Optimal Co-design Considering Control and Detection Trade-off

This section introduces the optimal co-design of watermark strengths, UIO parameters, and control gains that trade off the watermark-enhanced UIO's detection capability and the resulted control performance degradation. The following subsections will first theoretically analyse the detection capability enhancement and performance degradation elimination, and then synthesise the quantified metrics and related results into an co-design optimisation problem.

A. Detection Capability Enhancement

Since the added and removed watermarks are uniformly distributed within $[-\boldsymbol{\sigma}_i^p, \boldsymbol{\sigma}_i^p]$, we have that, according to (31), the alterations caused by P-FDIAs on the injected biases, i.e., $\Delta\boldsymbol{w}_i^p = \boldsymbol{w}_i^p(k) - \boldsymbol{w}_i^{pa}(k)$, would satisfy $\Delta\boldsymbol{w}_i^p \in [-\boldsymbol{\sigma}_i^p, \boldsymbol{\sigma}_i^p]$ with $\boldsymbol{\sigma}_i^p$ being its reachable bound. The reachable upper bound of resulted residual variation $\Delta\boldsymbol{r}_i^{pwa}(k)$ can be computed by considering the bounded alternation $\Delta\boldsymbol{w}_i^p$.

Lemma 1: Given DER dynamics (6) and the UIO constructed as (12), when DER i's local measurements $\boldsymbol{y}_i^p$ are under the P-FDIA (10), the corresponding residual variation $|\Delta\boldsymbol{r}_i^{pwa}(k)|$ resulted from watermarks will be bounded by $\boldsymbol{\sigma}_i^{prs}$ and satisfy

$$\boldsymbol{\sigma}_i^{prs} > \underline{\boldsymbol{\sigma}}_i^{prs} = \Big(|T_i^p| + \big|F_i^p\hat{K}_i^p\big| + \big|\hat{K}_i^p\big|\Big)\boldsymbol{\sigma}_i^p, \tag{35}$$

where $\underline{\boldsymbol{\sigma}}_i^{prs}$ is regarded as a lower estimation of $\boldsymbol{\sigma}_i^{prs}$.

Proof: Given the superposition principle of linear system dynamics (6) and observer dynamics (12), the residual variations under watermarking alternations $\Delta\boldsymbol{w}_i^p$ can be computed as

$$\Delta\boldsymbol{r}_i^{pwa}(k) = T_i^p\Delta\boldsymbol{w}_i^p(k) - \sum\nolimits_{l=k_i^p}^{k-1}(F_i^p)^{k-1-l}\hat{K}_i^p\Delta\boldsymbol{w}_i^p(l) \tag{36}$$

for $k \geq k_i^p$. Since $\Delta\boldsymbol{w}_i^p(k) \in [-\boldsymbol{\sigma}_i^p, \boldsymbol{\sigma}_i^p]$, the reachable upper bound of $\Delta\boldsymbol{r}_i^{pwa}(k)$ can be derived as

$$\begin{aligned}\boldsymbol{\sigma}_i^{prs} &= \Big(|T_i^p| + \sum\nolimits_{l=k_i^p}^{k-1}\big|(F_i^p)^{k-1-l}\hat{K}_i^p\big|\Big)\boldsymbol{\sigma}_i^p \\ &\geq \Big(|T_i^p| + |F_i^p\hat{K}_i^p| + |\hat{K}_i^p|\Big)\boldsymbol{\sigma}_i^p. \end{aligned} \tag{37}$$

which completes the proof. ■

Based on (35), the sufficient condition on detecting P-FDIAs can be formalised as the following.

Theorem 1: If the UIO parameters and watermark strengths are designed such that

$$\underline{\sigma}_{i,V}^{prs} > \bar{r}_{i,V}^{p\infty} \quad \text{or} \quad \underline{\sigma}_{i,I}^{prs} > \bar{r}_{i,I}^{p\infty}, \tag{38}$$

where $\underline{\boldsymbol{\sigma}}_i^{prs} = [\underline{\sigma}_{i,V}^{prs}, \underline{\sigma}_{i,I}^{prs}]^{\mathrm{T}}$, $\bar{\boldsymbol{r}}_i^p(\infty) = [\bar{r}_{i,V}^{p\infty}, \bar{r}_{i,I}^{p\infty}]^{\mathrm{T}}$, then let $\bar{\boldsymbol{r}}_i^p(\infty)$ be the steady-state value of $\bar{\boldsymbol{r}}_i^p(k)$ as $k \to \infty$ and it can be computed as

$$\bar{\boldsymbol{r}}_i^p(\infty) = |\mathrm{I}^2 + H_i^p|\bar{\boldsymbol{\rho}}_i + \frac{\nu_i^p}{1-\varsigma_i^p}\big(|T_i^p|\bar{\boldsymbol{\omega}}_i + |\hat{K}_i^p|\bar{\boldsymbol{\rho}}_i\big) \tag{39}$$

Proof: The detection residuals under P-FDIAs and watermarks can be obtained as

$$\boldsymbol{r}_i^{pwa}(k) = \boldsymbol{r}_i^p(k) + \Delta\boldsymbol{r}_i^{pwa}(k). \tag{40}$$

According to (18) and (19), when the initial state estimation error term converge to zero with $k \to \infty$, we know that $\boldsymbol{r}_i^p(\infty)$ is mainly determined by system noises and is independent from watermarks. Thus, the upper reachable bound of $|\boldsymbol{r}_i^{pwa}(\infty)|$, denoted by $\boldsymbol{\sigma}_i^{pwa}$, will satisfy

$$\boldsymbol{\sigma}_i^{pwa} > \boldsymbol{\sigma}_i^{prs}. \tag{41}$$

If (38) is satisfied, we know that, for $k \to \infty$, the lower estimation of $|\boldsymbol{r}_i^{pwa}(\infty)|$'s upper reachable bound is larger than the corresponding steady-state detection threshold, i.e.,

$$\sigma_{i,V}^{pwa} > \bar{r}_{i,V}^{p\infty} \text{ or } \sigma_{i,I}^{pwa} > \bar{r}_{i,I}^{p\infty}, \tag{42}$$

where $\boldsymbol{\sigma}_{i,V}^{pwa} = [\sigma_{i,V}^{pwa}, \sigma_{i,I}^{pwa}]^{\mathrm{T}}$. Therefore, the P-FDIA will be detected for sure and the proof is completed. ■

Satisfying condition (38) can guarantee the detection capability of P-FDIAs and a larger $\underline{\boldsymbol{\sigma}}_i^{prs}$ can make the P-FDIA detected with a smaller detection latency. More specifically, as $\underline{\boldsymbol{\sigma}}_i^{prs}$ is increased, the residual part exceeding the threshold $\bar{\boldsymbol{r}}_i^p(\infty)$ will hold a larger percentage of the residual range. Thus, it would be much quicker on average to reach the attack detection condition.

By following a similar process, we have that, under the S-FDIA (11), the watermarks' impact on the injected bias, i.e., $\Delta\boldsymbol{w}_{i,j}^s$, satisfies $\Delta\boldsymbol{w}_{i,j}^s \in [-\boldsymbol{\sigma}_{i,j}^s, \boldsymbol{\sigma}_{i,j}^s]$ with $\boldsymbol{\sigma}_{i,j}^s$ being the strengths of the added watermarks to $\boldsymbol{y}_j^{pr}$. Thus, the corresponding residual fluctuation $|\Delta\boldsymbol{r}_{i,j}^{swa}|$ has upper reachable bound $\boldsymbol{\sigma}_{i,j}^{srs}$, whose lower estimation, denoted by $\underline{\boldsymbol{\sigma}}_{i,j}^{srs}$, can be derived as

$$\boldsymbol{\sigma}_{i,j}^{srs} \geq \underline{\boldsymbol{\sigma}}_{i,j}^{srs} = \Big(|T_j^s| + \big|F_j^s \hat{K}_j^s\big| + \big|\hat{K}_j^s\big|\Big) 2\boldsymbol{\sigma}_{i,j}^s. \tag{43}$$

The sufficient attack detection condition can be obtained by

$$\underline{\sigma}_{i,j,V}^{srs} > \bar{r}_{i,j,V}^{s\infty} \quad \text{or} \quad \underline{\sigma}_{i,j,I}^{srs} > \bar{r}_{i,j,I}^{s\infty}, \tag{44}$$

where $\underline{\boldsymbol{\sigma}}_{i,j}^{srs} = [\underline{\sigma}_{i,j,V}^{srs}, \underline{\sigma}_{i,j,I}^{srs}]^{\mathrm{T}}$, $\bar{\boldsymbol{r}}_{i,j}^s(\infty) = [\bar{r}_{i,j,V}^{s\infty}, \bar{r}_{i,j,I}^{s\infty}]^{\mathrm{T}}$, and

$$\bar{\boldsymbol{r}}_{i,j}^s(\infty) = |\mathrm{I}^2 + H_j^s|\bar{\boldsymbol{\rho}}_j + \frac{\nu_j^s}{1-\varsigma_j^s}\big(|T_j^s|\bar{\boldsymbol{\omega}}_j + |\hat{K}_j^s|\bar{\boldsymbol{\rho}}_j\big).$$

B. Elimination of Control Performance Degradation

Adding watermarks to the inputs of primary and secondary controllers would inevitably degrade the control performance. In the attack-free case, the primary and secondary control inputs after incorporating the watermarked measurements are

$$u_i^w(k) = (\boldsymbol{g}_i^P)^{\mathrm{T}}\boldsymbol{y}_i^{pw}(k) + g_i^I \sum_{l=0}^{k}\big(V_{ref,i} + \alpha_i(k) - \boldsymbol{\iota}^{\mathrm{T}}\boldsymbol{y}_i^{pw}(l)\big), \tag{45}$$

and

$$\alpha_i^w(k) = \boldsymbol{\kappa}^{\mathrm{T}} \sum_{l=0}^{k}\sum_{j\in\mathcal{N}_i^c} a_{ij}^c\Big(\frac{\boldsymbol{y}_{i,j}^{psw}(l)}{I_{tj}^s} - \frac{\boldsymbol{y}_i^{pr}(l)}{I_{ti}^s}\Big), \tag{46}$$

It is necessary to analyse these watermarks' impacts on the control performance and try to eliminate any degradation. Since the primary and secondary controllers have different objectives in voltage regulation and current sharing, respectively, their control performance degradation under watermarks is discussed separately as follows:

1) Voltage-based Primary Controller: For primary controller (45), the variations of u_i^w resulted from watermarks, denoted by Δu_i^w, are extracted as

$$\Delta u_i^w(k) = (\boldsymbol{g}_i^P)^{\mathrm{T}}\boldsymbol{w}_i^p(k) - g_i^I \sum_{l=k_i^{pw}}^{k} \boldsymbol{\iota}^{\mathrm{T}}\boldsymbol{w}_i^p(l), \tag{47}$$

from which the upper reachable bound of $|\Delta u_i^w(k)|$, denoted by σ_i^{pu}, can be characterised as

$$\sigma_i^{pu} = \big|(\boldsymbol{g}_i^P)^{\mathrm{T}}\big|\boldsymbol{\sigma}_i^p + (k-k_i^{pw})g_i^I\boldsymbol{\iota}^{\mathrm{T}}\boldsymbol{\sigma}_i^p, k \geq k_i^{pw}, \tag{48}$$

which is proportional to the increase of time instant k due to the integral term of voltage tracking errors. To eliminate the linearly increasing control input disturbance, the watermark strength should be designed such that

$$\boldsymbol{\iota}^{\mathrm{T}}\boldsymbol{\sigma}_i^p = 0, \tag{49}$$

which means that the watermark strength added to the local voltage measurement has to be zero and the enhanced detection capability totally relies on the watermarks introduced into the local current measurement. To further eliminate the watermark's negative impacts on the primary control performance, the control signal's variation magnitude after incorporating (49) needs to also satisfy

$$\sigma_i^{pu} = \big|(\boldsymbol{g}_i^P)^{\mathrm{T}}\big|\boldsymbol{\sigma}_i^p = 0. \tag{50}$$

2) Current-oriented Secondary Controller: For the secondary controller (46), the variation of α_i^w under watermark $\boldsymbol{w}_{i,j}^s$, denoted by $\Delta\alpha_i^w$, can be calculated as

$$\Delta\alpha_i^w = \boldsymbol{\kappa}^{\mathrm{T}} \sum_{j\in\mathcal{N}_i^c}\sum_{l=k_{i,j}^{sw}}^{k} \frac{a_{ij}^c}{I_{tj}^s}\boldsymbol{w}_{i,j}^s(l), \tag{51}$$

whose upper reachable bound $\sigma_i^{s\alpha}$ can be calculated as

$$\sigma_i^{s\alpha} = \sum_{j\in\mathcal{N}_i^c}\big|(k-k_{i,j}^{sw})\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\sigma}_{i,j}^s\big|, \tag{52}$$

which is mainly induced by the integral term of current sharing errors, and can be eliminated by designing watermarks satisfying

$$\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\sigma}_{i,j}^s = 0. \tag{53}$$

With $\boldsymbol{\kappa} = [0,1]^{\mathrm{T}}$, we know that the watermark added to the interactive current measurement among DERs needs to be zero to eliminate the performance degradation of the secondary controller. The UIO's detection capability enhancement is achieved by the watermark introduced into the interactive voltage measurement. Although the watermarks are added to either voltage or current measurements, the FDIAs compromising both measurements can potentially enlarge the detection residuals as indicated by (35), and thus may trigger the alarm of the UIO detector, which incorporates the DER's power electronics dynamics that closely couple voltages and currents.

TABLE III: Co-design optimisation problems for watermarking strengths, UIO parameters, and control gains under the P-FDIA and S-FDIA

	P-FDIA	S-FDIA
Obj.	min Watermark strength (54)	min Watermark strength (60)
Dec. Vars.	Watermark strengths, UIO parameters, Control gains	Watermark strengths, UIO parameters
Constrs.	1) Attack detection (38) 2) Ctrl. performance (49),(50) 3) UIO design (13)-(17), (55) 4) Control stability (56), (57)	1) Attack detection (44) 2) Control performance (53) 3) UIO design (21)-(25), (61)

C. Optimal Co-Design Problems

Based on the above analysis, the controller and UIO parameters need to be co-designed with the watermark strengths to minimise the performance degradation of the controller while guaranteeing that the P-FDIA and S-FDIA are detectable. The basic structures of the optimisation co-design problems are shown in Fig. III. The objectives are to minimise the watermark strength such that the addition of watermarks to measurements can be easily hidden from the adversary, and the decision variables compromise watermark strength, UIO parameters, and primary control gains. The attack detection conditions and control performance requirements are included in the constraints to trade off the detection capability against control performance using appropriately chosen weight parameters. Moreover, the design principles of the UIO parameters and the stability constraints on the control gains are incorporated.

For P-FDIAs, the co-design problem is formulated as

$$\min_{\boldsymbol{\sigma}_i^p, T_i^p, F_i^p, \boldsymbol{g}_i^P, g_i^I} \quad \left|\boldsymbol{\sigma}_{\boldsymbol{i}}^{\boldsymbol{p}}\right| \tag{54}$$

$$\begin{aligned} \text{s.t.} \quad & (38), (49), (50), \\ & (13)-(17), \\ & |\lambda(F_i^p)| < 1, \qquad (55) \\ & |\boldsymbol{g}_i^P| \le [1, R_{ti}]^{\mathrm{T}}, \qquad (56) \\ & 0 < g_i^I < \bar{g}_i^I, \qquad (57) \end{aligned}$$

The objective function is to minimise the watermark strength to keep it hidden from the adversary. Inequality (38) imposes the attack detection condition and equalities (49) and (50) eliminate the linearly increasing watermark-induced impact on the primary control input. Equations (13)-(17) include the design principles of the UIO parameters and (55) converges the UIO's state estimation error, where the function $\lambda(\cdot)$ calculates the matrix eigenvalues. Inequalities (56)-(57) are explicit constraints for the primary control gains that guarantee exponential stability [41], where $\bar{g}_i^I = \frac{1}{L_{ti}}\big(g_i^P(1)-1)(g_i^P(2)-R_{ti})\big)$.

To handle the intractable matrix eigenvalue function, the matrix F_i^p is considered to be symmetric and thus can be decomposed as

$$F_i^p = [\boldsymbol{\beta}_i^{p1}, \boldsymbol{\beta}_i^{p2}] \begin{bmatrix} \lambda_i^{p1}, 0 \\ 0, \lambda_i^{p2} \end{bmatrix} \begin{bmatrix} (\boldsymbol{\beta}_i^{p1})^{\mathrm{T}} \\ (\boldsymbol{\beta}_i^{p2})^{\mathrm{T}} \end{bmatrix}, \tag{58}$$

where $\boldsymbol{\beta}_i^{p1}$ and $\boldsymbol{\beta}_i^{p2}$ are orthogonal unit eigenvectors corresponding to eigenvalues λ_i^{p1} and λ_i^{p2}, respectively. Hence, constraint (55) can be replaced with (58) and

$$|\lambda_i^{p1}| < 1, |\lambda_i^{p2}| < 1. \tag{59}$$

Finally, the optimization problem can be transformed into a polynomial optimization problem, which is non-convex but can be solved by mature solvers such as the *fmincon* function in Matlab. It is possible to find a feasible sub-optimum with appropriately chosen initial points. It is noted that problem (54) can be also solved in other environments such as AMPL, however the *fmincon* function is chosen here merely due to Matlab's compatibility with OPAL-RT simulator. Another direction to solve (54) involves relaxing (54) to a convex problem such that the optimum can be quickly obtained, which, however, requires substantial effort and is left to future work.

For S-FDIAs, the co-design problem is given by

$$\min_{\boldsymbol{\sigma}_{i,j}^s, T_j^s, F_j^s} \quad \left|\boldsymbol{\sigma}_{\boldsymbol{i,j}}^{\boldsymbol{s}}\right| \tag{60}$$

$$\begin{aligned} \text{s.t.} \quad & (44), (53), \\ & (21)-(25), \\ & |\lambda(F_j^s)| < 1, \qquad (61) \end{aligned}$$

which, compared with the optimization problem (54), is much simpler since no control stability constraint needs to be considered. Inequality (44) imposes the detection condition for S-FDIAs and equality (53) eliminates the watermarks' impacts on the secondary control input. Equations (21)-(25) and inequality (61) set the constraints for the UIO parameters. By following the similar transformation principle, the optimization problem (60) can be converted to a polynomial optimization problem and solved by existing solvers.

D. Robustness to Time Synchronisation Error

According to (31), the generated watermark signals are closely related to the time synchronisation between the data transmitter and receiver. In the ideal case, when the data transmitter's and receiver's time instants perfectly match, the generated watermark would be equal to the removed watermark. However, considering the hard real-time requirement of the primary controller (millisecond), it is difficult to always guarantee accurate time synchronisation between local sensors and primary controllers especially under extreme operating environments and varying communication delays [42]. Moreover, the inconsistency of generated and removed watermarks resulting from the TSEs is indistinguishable from that caused by the FDIAs.

To enhance the robustness of the watermarking scheme to TSEs, the single-moment alarming principle (26) is changed to a multiple-moments alarming principle. That is, for the detection metric of P-FDIAs denoted by $\tilde{\boldsymbol{r}}_i^p(k)$, it will capture the accumulated $|\boldsymbol{r}_i^p(l)|$ within a sliding time window with length $l \in [k - L_p + 1, k]$, i.e.,

$$\tilde{\boldsymbol{r}}_i^p = \frac{1}{L_p} \sum_{l=k-L_p+1}^{k} |\boldsymbol{r}_i^p(l)|. \tag{62}$$

The updated detection metric $\tilde{\boldsymbol{r}}^s_{i,j}$ for S-FDIAs can be similarly obtained as

$$\tilde{\boldsymbol{r}}^s_{i,j} = \frac{1}{L_p} \sum_{l=k-L_p+1}^{k} |\boldsymbol{r}^s_{i,j}(l)|. \tag{63}$$

The detection thresholds and alarming principles are kept the same as the previous ones. The updated detection metrics ensure that once the operator notices the TSEs and correct them rapidly, then false alarms will not be triggered, thus making the proposed watermarking-enhanced UIOs robust to certain TSEs. The larger T_p can definitely improve this robustness but will also decrease the detector's sensitivity to attacks and induce higher detection latency. In this case, the attack detection conditions for the P-FDIAs and S-FDIAs used in optimization problems (54) and (60) should be adjusted as

$$\underline{\sigma}^{pr}_{i,V} > \xi^p_{i,V}\bar{r}^{p\infty}_{i,V} \text{ or } \underline{\sigma}^{pr}_{i,I} > \xi^p_{i,I}\bar{r}^{p\infty}_{i,I}, \tag{64}$$

and

$$\underline{\sigma}^{sr}_{i,j,V} > \xi^s_{i,j,V}\bar{r}^{s\infty}_{i,j,V} \text{ or } \underline{\sigma}^{sr}_{i,j,I} > \xi^s_{i,j,I}\bar{r}^{s\infty}_{i,j,I}, \tag{65}$$

respectively, with weight parameters $\xi^p_{i,V}, \xi^p_{i,I} > 1$ and $\xi^s_{i,j,V}, \xi^s_{i,j,I} > 1$ to amplify the detection residuals under attacks. Choosing an appropriate T_p is important to trade off the detection performance and synchronisation error robustness, which is comprehensively demonstrated in the simulation study.

VI. Matlab/Simulink Studies and Experimental Validation

In this section, the effectiveness of the proposed watermarking-enhanced UIOs is fully verified and the method's detection latency and TSE robustness under different parameter settings are carefully investigated through Matlab/Simulink studies. To validate the proposed method's applicability to industrial scenarios, extensive experimental studies are conducted in a cyber-physical co-simulation microgrid testbed and Raspberry Pi to demonstrate its effectiveness and low complexity, respectively.

A. Matlab/Simulink Studies

A 4-DER microgrid with cyber and physical topology from Fig. 2 is established in Matlab/Simulink, where the converters are simplified as controllable voltage sources for demonstrative purposes and the related electrical parameters of DER i are configured as in [19], [20]. The reference PCC voltages are $V_{ref,1} = 40$V, $V_{ref,2} = 40.1$V, $V_{ref,3} = 40.3$V, and $V_{ref,4} = 39.8$V, and DERs' resistive loads $Z_1 = 30\Omega$, $Z_2 = 25\Omega$, $Z_3 = 35\Omega$, and $Z_4 = 30\Omega$ are considered. The initial primary control gains are $\boldsymbol{g}^P_i = [0.1, 0.1]^{\mathrm{T}}$ and $g^I_i = 50$ and the secondary link weights are $a^c_{ij} = 2$. The system noises bounds are $\bar{\boldsymbol{\rho}}_i = 0.01 * [1, 1]^{\mathrm{T}}$ and $\bar{\boldsymbol{\omega}}_i = 0.01 * [1, 1]^{\mathrm{T}}$. The continuous system is discretized with sampling time $T = 500\mu$s.

1) Trade-off between Detection Effectiveness and Control Performance: In this part, the trade-off between detection effectiveness and control performance of the carefully designed watermarks is demonstrated by conducting comparative studies with randomly chosen watermarks. The time window length for the updated residual is $L_p = 200$, which is equivalent to 0.1s with sampling time 0.5ms. The watermarking scheme is activated at $t = 2$s, and the FDIA is launched at $t = 3$s. The cases under P-FDIAs and S-FDIAs are shown separately as follows:

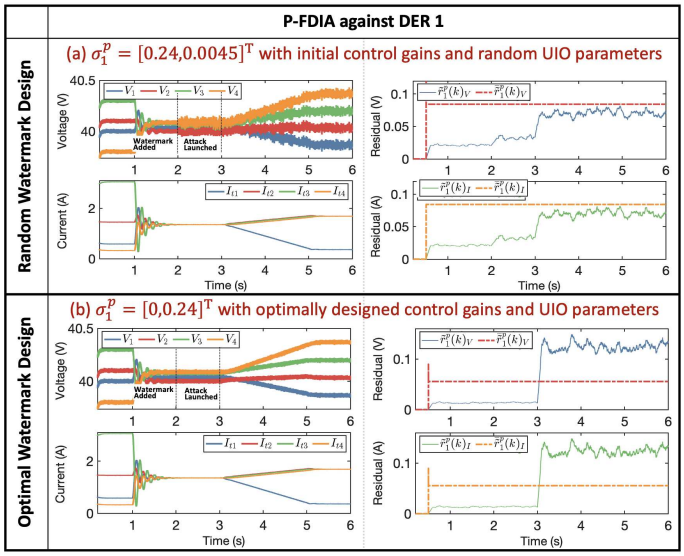


Fig. 5: The superiority of optimally designed watermarks in detecting P-FDIAs and avoiding excessive control performance degradation compared with the randomly chosen watermarks is demonstrated, where sub-figures (a) and (b) denote the cases of random and optimal watermarks, respectively.

Case I: This case shows the results under the P-FDIA against DER 1, where the attack vector is constructed as (29) with fake unknown input $d^a_1(k) = (kT_{samp} - 3), kT_{samp} \in [3, 5]$. As shown in sub-figure (a) of Fig. 5, when the watermark strength $\boldsymbol{\sigma}^p_i$ is randomly chosen as $[0.24, 4.5\mathrm{e}-3]^{\mathrm{T}}$, primary control gains are kept unaltered, and UIO parameters are chosen randomly, the added watermarks will induce serious oscillations on PCC voltages and only introduce limited residual improvements, where no residual is larger than the threshold. By solving the formulated optimization problem (54), the watermark strength is designed as $\boldsymbol{\sigma}^p_i = [0, 0.24]^{\mathrm{T}}$, whose sum is smaller than that of the randomly chosen case, and the primary control gains are adjusted as $\boldsymbol{g}^P_i = [-0.38, 0]^{\mathrm{T}}$ and $g^I_i = 44.63$ with optimally designed UIO parameters. The results in sub-figure (b) of Fig. 5 indicate that the negative impacts of watermarks on control performance are fully eliminated with significantly improved residuals, where almost 100% residuals can exceed the threshold. Therefore, the trade-off between detection and control can be perfectly balanced by designing watermark strengths using (54). The amplification gains of attack detection conditions (64) are set as $\xi^p_{i,V} = \xi^p_{i,I} = 9.6$ to effectively detect attacks via updated residual $\tilde{\boldsymbol{r}}^p_i$. Compared with the legacy UIO-based detection scheme (12) without embedded watermarks, the proposed detection method can achieve satisfactory detectability against P-FDIAs.

Case II: This case illustrates the results under the S-FDIA against the link from DERs 1 to 2, where the attack vector is constructed as (30) with the same fake unknown input

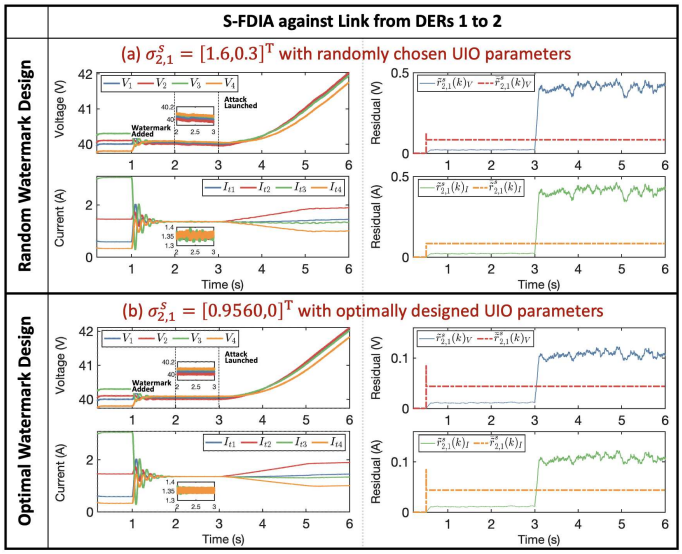


Fig. 6: The superiority of optimally designed watermarks in detecting S-FDIAs and avoiding excessive control performance degradation compared with the randomly chosen watermarks is demonstrated, where sub-figures (a) and (b) denote the cases of random and optimal watermarks, respectively.

$d^a_{2,1} = d^a_1$. In sub-figure (a) of Fig. 6, when the watermark strength is randomly chosen as $\boldsymbol{\sigma}^s_{i,j} = [1.6, 0.3]^{\mathrm{T}}$, unexpected fluctuations will appear on both voltages and currents due to the non-trivial watermarks added to current measurements, which directly affect the secondary control input (46). Although the resulted detection residuals are far larger than the threshold, the severely degraded secondary control performance may also make the watermarking scheme unacceptable. The optimal watermark strength $\boldsymbol{\sigma}^s_{i,j} = [0.9560, 0]^{\mathrm{T}}$ and UIO parameters obtained by solving the optimization problem (60) can well balance the trade-off between detection and control. As illustrated in sub-figure (b) of Fig. 6, the resulting residuals are all larger than the threshold while the control performance degradation can be totally eliminated. To meet the attack detection conditions under updated residual $\tilde{\boldsymbol{r}}^s_{i,j}$, the amplifications gains in (65) are set as $\xi^s_{i,j,V} = \xi^s_{i,j,I} = 11$.

2) Robustness to Time Synchronisation Error: In this part, the robustness of updated residual residuals to TSEs are investigated. Since the results for primary and secondary control loops are similar, only the primary control's case is demonstrated. The optimal designed watermark strength, primary control gains, and UIO parameters are adopted in this study. We consider three typical synchronisation error types, i.e., continuous, intermittent, and random, and the results shown in Fig. 7 with time window length $L_p = 200$ are explained as follows:

a) The continuous TSE (CTSE) exists for a fixed time duration. When introducing constant TSE into the primary control loop, the residual increases as the TSE duration and the threshold can tolerate up to 25 sampling points' continuous TSEs.
b) The intermittent TSE occurs with a certain interval. When the interval between TSE events is set to L_p, the maximal tolerable TSE event length is 20 sampling points.
c) The random TSE happens with a predefined probability. The maximal TSE probability that can be tolerated by the threshold is 3%.

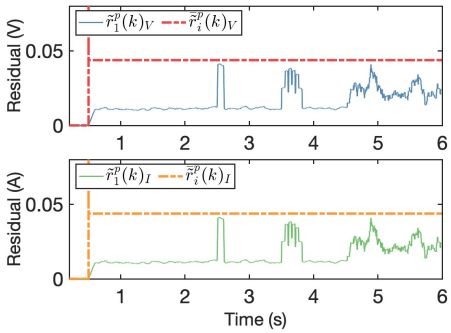
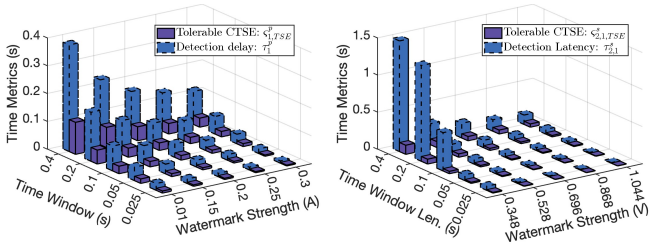


Fig. 7: The robustness of updated residual $\tilde{\boldsymbol{r}}^p_1$ to the time synchronisation errors existing in primary control loop is validated.

From the analysis above, we have two observations: 1) The tolerable TSE length is much smaller than the time window length. The significantly reduced length of tolerable TSEs is because that the data alternation caused by TSEs is very large, almost twice of the watermark strength according to the watermarking generation scheme (31). 2) The updated residual has highest robustness to continuous TSEs while has lowest robustness to random TSEs. This phenomenon is caused by the memory property of the UIO detection residual, i.e., the impact of previous data alternation will be memorized but vanish exponentially with time. Hence, the probability would become low when it is required to keep TSE-free for a long time.

3) Sensitivity of Detection Latency and TSE Robustness: The detection latency and tolerable CTSE, denoted by τ^p_i, ς^p_i and $\tau^s_{i,j}, \varsigma^s_{i,j}$ under P-FDIAs and S-FDIAs, respectively, can be affected by both watermark strengths and the time window L_p. To illustrate their connection, five different watermark strengths obtained by solving optimization problems (54) and (60) with amplification gains $\xi^p_{i,V}, \xi^p_{i,I}$ and $\xi^s_{i,j,V}, \xi^s_{i,j,I}$ varying among $\{2, 4, 6, 8, 10, 12\}$ and five time window lengths $L_p \in \{50, 100, 200, 400, 800\}$ are chosen, under which the detection latency and tolerable CTSEs are shown in Fig. 8. Two clear observations can be obtained from the results: 1) The detection latency will increase as the time window length, and the watermark strength can significantly decrease the detection latency. 2) The tolerable CTSE is proportional to the time window length, and the increased watermark strength will decrease the tolerable CTSE. In general, to balance the trade-off between detection latency and tolerable CTSEs, the time window length and watermark strength need to be appropriately chosen as a whole.

4) Comparison with State-of-the-Art Watermarking Schemes: The comparative study with the state-of-the-art watermarking scheme for microgrids [29] illustrating the proposed watermarking scheme's advantage in decreasing the time delay within control loops is described in Fig. 9. In particular, the special output-added watermarking scheme adopted in [29] requires de-watermarking the signal before forwarding it to the controller, which will introduce double time delays into the control loop compared with the



(a) The case under the P-FDIA against DER 1

(b) The case under the S-FDIA against the link from DERs 1 to 2

Fig. 8: The detection latency $\tau^p_i, \tau^s_{i,j}$ and tolerable CTSE $\varsigma^p_{i,TSE}, \varsigma^s_{i,j,TSE}$ under the variations of time window L_pT_{samp} and watermark strength $\boldsymbol{\sigma}^p_i, \boldsymbol{\sigma}^s_{i,j}$ are illustrated.

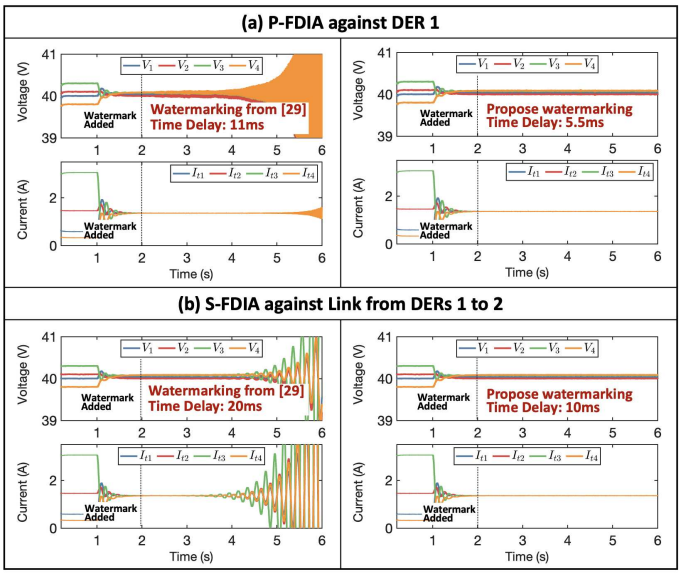


Fig. 9: This figure verifies the proposed watermarking scheme's advantage in decreasing the time delay within primary and secondary control loops as shown in sub-figures (a) and (b), respectively, which is especially important for the time-critical control operations of microgrids, compared with the state-of-the-art watermarking scheme [29].

watermarking scheme proposed in this work. Although the conventional watermarking scheme can effectively enhance the UIO's detectability against stealthy FDIAs, the additional control delay may destabilise the voltage and current since the primary and secondary control layers in microgrids are time-critical. As implied by Fig. 9, in the P-FDIA case, when the watermarking scheme from [29] induce a 11ms delay into the primary and secondary control loops, the voltage will be destabilised first with the cascading impact being then propagated to the secondary control layer. Under the same circumstance, the proposed watermarking scheme will only bring in 5.5ms time delay, which is validated to be tolerable by the controllers' essential robustness. The similar results occur in the S-FDIA case, where the difference in tolerable time delay is due to the fact that only the secondary control loop is affected under this scenario. In practice, the actual time delay introduced by watermarking/de-watermarking actions can vary depending on the resources of the edge-devices. However, through this

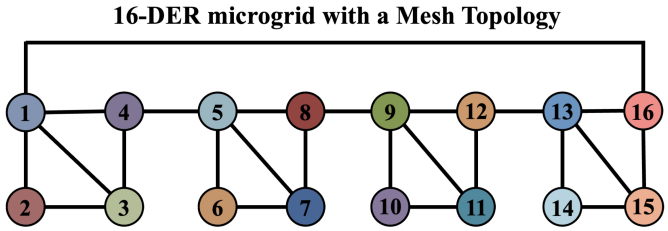
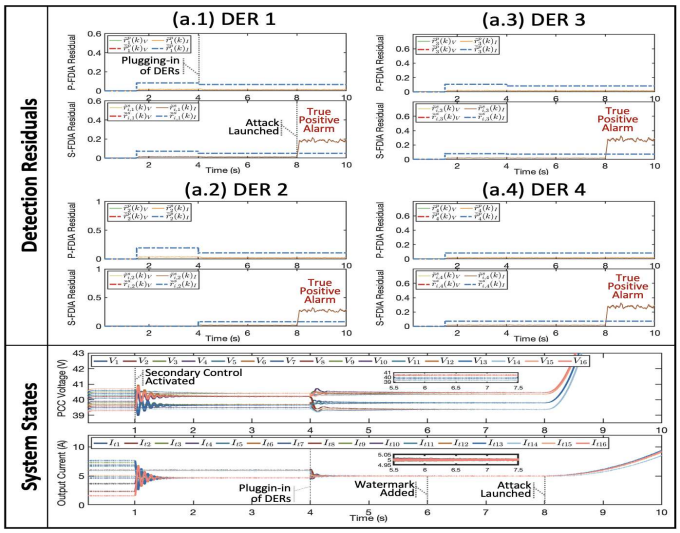


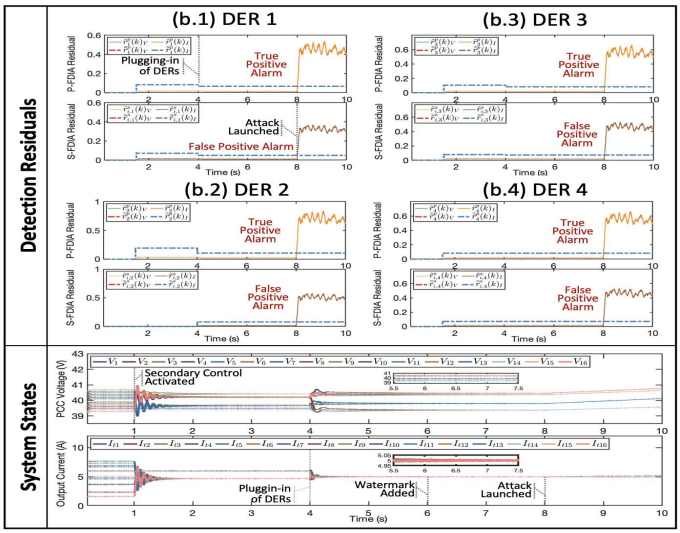
Fig. 10: This figure pictures the communication and electrical topology networks of 16-DER microgrid.

comparative study, it is clear that the proposed watermarking scheme has unique advantages and can significantly decrease the requirements on the computation capability of sensors and controllers. Moreover, after appropriately co-designing the control gains and watermarking strengths, the control performance degradation resulting from the added watermarks can be negligible.

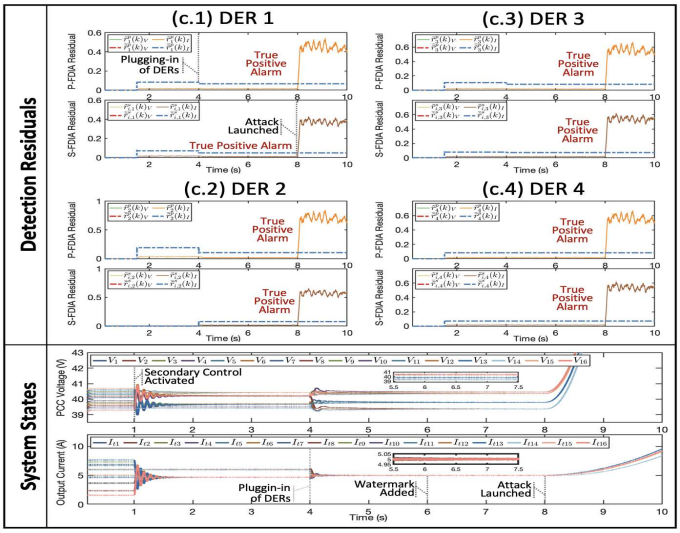
5) Scalability to 16-DER Microgrid: When the cyber-physical microgrid scale increases to 16 DERs, the proposed method would still be effective as validated in this part. A 16-DER microgrid with mesh communication and electrical topology networks as shown in Fig. 10 is established, where the electrical parameters are set according to [19]. In particular, the plugging-in of DERs 2, 6, 10, 14 is enabled at $t = 4$s, the watermark addition and removal are activated at $t = 6$s, and the P-FDIAs and S-FDIAs are launched at $t = 8$s to affect all primary and secondary controllers. The detection results related to four representative DERs, i.e., DERs 1, 2, 3, 4, under three attack cases including (a) S-FDIA only, (b) P-FDIA only, and (c) Both S-FDIA and P-FDIA and the associated system states are showcased in Fig. 11. When the plugging-in of the DERs is enabled, the UIO parameters of the electrically connected DERs, such as DERs 1, 2, 3, are updated accordingly, under which the corresponding detection residuals and thresholds have slight alterations. But there is no detection anomaly due to the plugging-in event as the UIO-based detector is designed to be robust to the cyber-physical topology change caused by the plugging-in of DERs in microgrids. As validated in sub-figures (a)-(c) of Fig. 11, the watermarks with optimised strengths will not cause significant impact on the system states. However, the watermark-enhanced UIOs perform differently under the three attack cases. In the S-FDIA only case, the attack alarm will only be flagged by the UIOs designed for S-FDIAs. When it comes to the P-FDIA only case, the UIOs designed for both P-FDIAs and S-FDIAs give attack alarms, among which the latter are deemed to be false positive alarms. This result is due to the fact that the incorrectly removed watermarks caused by the P-FDIAs will also affect the data communicated to the neighboring secondary controllers even if there is no S-FDIA present. Although, in the case consisting of both S-FDIA and P-FDIA, all UIOs can generate attack alarms, it is difficult to distinguish the attack case (c) from (b) as their alarming results seem to be similar. In other words, when the P-FDIA exists in a DER, it is hard to judge if there are S-FDIAs against the communication links connecting it to the neighboring DERs.



(a) S-FDIA only case

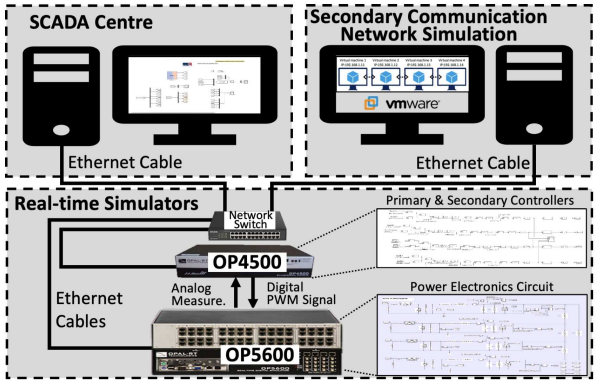


(b) P-FDIA only case

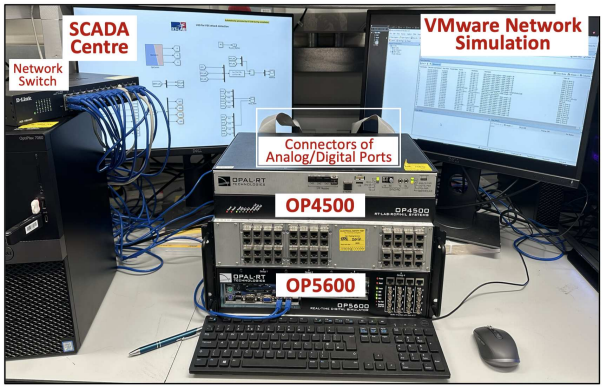


(c) S-FDIA and P-FDIA case

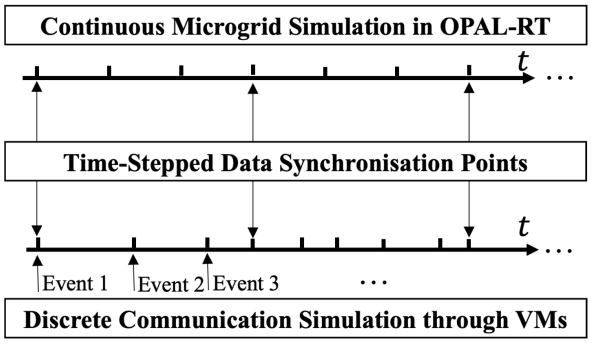
Fig. 11: The detection results against three attack cases including (a) S-FDIA only, (b) P-FDIA only, and (c) Both S-FDIA and P-FDIA, in the presence of DERs' plugging-in, as well as the associated system states under watermarks and attacks are illustrated.



(a) Overview of cyber-physical co-simulation testbed setup



(b) Connection of real devices in Lab



(c) Time-stepped data synchronisation scheme

Fig. 12: This figure pictures the setup of the cyber-physical co-simulation microgrid testbed, where the OP5600 emulates the 4-DER's power electronic dynamics, the OP4500 implements primary and secondary controllers and interacts with OP5600 through analog and digital ports, a host simulates the secondary communication network utilising VMs and exchanges data with OP4500 via a time-stepped data synchronisation interface, and the SCADA centre running in an independent host collects real-time operation data from OP4500 through an Ethernet cable.

Therefore, there still requires some future efforts to distinguish the existence of S-FDIA from P-SFDIA when these two types of attacks occur at the same time.

B. Cyber-Physical Co-Simulation Experimental Validation

1) Effectiveness of Proposed Method: The effectiveness of the proposed watermark design method is validated in a cyber-physical co-simulation testbed as shown in Fig. 12. Two high-

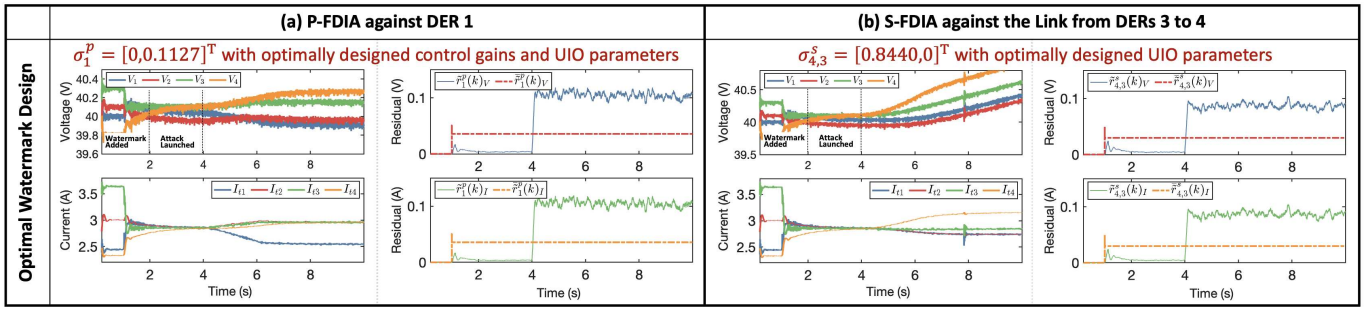


Fig. 13: The effectiveness of designed optimal watermarks in enhancing attack detection performance while preserving control performance is validated in the cyber-physical co-simulation microgrid testbed, where sub-figures (a) and (b) represent the cases of P-FDIA and S-FDIA, respectively.

fidelity real-time simulators including OP5600 and OP4500 are used to emulate the power electronics circuits of 4-DER microgrid and develop primary and secondary controllers, wherein the state measurement and PWM control signal are exchanged through two connectors of analog and digital ports, respectively. The communication network that transmits data between secondary controllers is simulated in VMware running in a high-performance computer equipped with 13-th Intel i9 CPU and 96GB RAM. Four Virtual Machines (VMs) are established in a local area network and are configured to interact data via TCP Modbus communication protocol. The data synchronisation between the continuous microgrid simulation and discrete communication simulation is implemented in a time-stepped scheme [43], where the two simulations run independently and halt at fixed synchronisation points to exchange data. The SCADA centre running in another host collects real-time operation data from OP4500 via an Ethernet cable, which can be used for monitoring purpose.

The power and communication typologies, electrical parameters, and control gains are the same as the ones in the Matlab/Simulink studies. The bounds of system noises are set by $\bar{\boldsymbol{\rho}}_i = 0.005 * [1,1]^{\mathrm{T}}$ and $\bar{\boldsymbol{\omega}}_i = 0.005 * [1,1]^{\mathrm{T}}$. In addition to the resistive loads, extra 1A current loads are added to all DERs. The secondary control is activated at $t = 1$s, the watermarks with designed strengths are integrated at $t = 2$s, and the P-FDIA and S-FDIA are launched at $t = 4$s. Similarly, two cases are studied to demonstrate the proposed scheme's effectiveness under P-FDIAs and S-FDIAs.

Case I: This case shows the results under the P-FDIA against DER 1, where the fake unknown input is $d_1^a(k) = 0.25 * (kT_{samp} - 4), kT_{samp} \in [4,6]$. The designed watermark strength $\boldsymbol{\sigma}_1^p = [0, 0.1127]^{\mathrm{T}}$ is smaller than that of Case I in simulation studies due to the decreased noise bounds. According to sub-figure (a) of Fig. 13, the results follow those from Case I in Section VI-A-1), where the detection capability against P-FDIAs is significantly enhanced without affecting the control performance.

Case II: This case illustrates the results under the S-FDIA against the link from DERs 3 to 4, where the attack vector is the same as that of the P-FDIA. The designed watermark strength $\boldsymbol{\sigma}_{4,3}^s = [0.8440, 0]^{\mathrm{T}}$ can effectively enhance the detection capability while not causing obvious control performance degradation. The results follow the those from Case II in Section VI-A-1) besides some additional fluctuations on

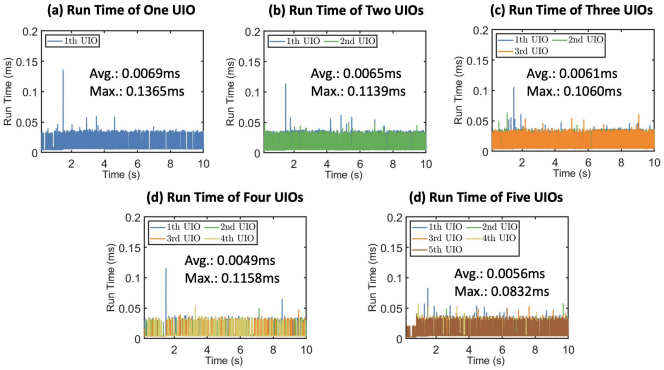


Fig. 14: The run time of the proposed method in a Raspberry Pi equipped with 8GB RAM is demonstrated, where the number of watermark-enhanced UIOs varies from 1 to 5 corresponding to the sub-figures from (a) to (d).

system states caused by disruptive attack vectors.

2) Time and Space Complexity of Proposed Method: The proposed watermarking-enhanced UIO-based detector is fully distributed and requires only the local information within each DER for P-FDIAs and neighboring DER's information for S-FDIAs, which will not be significantly affected by the growth of the microgrid's scale. As discussed in Section VI-A-5), in the presence of plugging-in of DERs, the UIO-based detector requires merely to update UIO parameters and no extra computation burden will be resulted. In particular, the computational process of the proposed method involves mainly three components, i.e., watermark generation (31), UIO-based detector (12) or (20), and residual smoothing (62) or (63). The watermark generation module takes the local or neighboring information as input and look ups the hashed value in the corresponding hashing table established in advance. Similarly, the UIO-based detector uses the local or neighboring information as an input to calculate the detection residual, followed by the residual smoothing process to get the average residual within a sliding time window that has fixed length. Since the size of the local and neighboring DER's information is fixed, i.e., a 3×1 vector, the computation time and memory usage of these three computational components should be constant from the average sense and will not be affected by the size of the microgrid. Therefore, the time and space complexity of proposed method are both $\mathcal{O}(1)$. To validate the statement, the run time of the proposed method is tested in a Raspberry Pi equipped with 8GB RAM and the results are demonstrated in

Fig. 14, where the number of watermark-enhanced UIOs varies from 1 to 5. As the number of UIOs grows, the run time does not significantly increase. For example, when five watermark-enhanced UIOs run simultaneously in the Raspberry Pi, the average run time is still less than 0.01ms with the maximal run time smaller than 0.1ms, which are both significantly away from the sampling time 0.5ms. Therefore, the proposed watermark-enhanced UIO has negligible computation burden and can be seamlessly applied to industrial scenarios without degrading the performance of primary and secondary controllers. The maximal number of UIOs deployed within a DER is set as 5 considering that in practice, the microgrid scale and associated network density would not be unlimited in order to satisfy the cost-efficiency requirement such as the 16-DER microgrid shown in Fig. 10.

3) Discussion on the Implementation of Proposed Method in Real-World Microgrids: Since the complexity of watermarking-enhanced UIO is lightweight enough to be integrated into the primary and secondary controllers, the key issue of implementing the proposed method in real-world microgriods is the addition of watermarks to original measurements. For the S-FDIA case, the designed watermark can be added to the measurement by secondary controller before transmitting it to the neighboring DERs, which only requires to update the code inside controller and is thus seamlessly applicable to realistic scenarios. For the P-FDIA case, the addition of watermarks to sensor measurements depends on the programming flexibility of the deployed sensors: 1) If the sensor's firmware is re-programmable [31], then the watermark-addition functionality may be included in a upgraded sensor firmware, which requires significant programming efforts and close collaboration with the sensor manufacturer. 2) Otherwise, a bump-in-the-wire device is needed behind the sensor to add watermarks to sensor readings before sending them to the primary controller. Although this newly deployed device may introduce additional attack surfaces, it turns out to be a cost-friendly option for the system operator as the other components, including the sensor and controller, do not require replacement. Moreover, following the attack model in Section III-C, the newly deployed device would likely not be the target as the adversary aims to launch man-in-the-middle attacks to inject biases into the data packets transmitted within the primary and secondary control loops.

VII. Conclusion

In this paper, an innovative physics-aware watermarking embedded in UIOs is proposed to proactively detect the P-FDIA and S-FDIA in microgrids, where random and bounded watermark noises are employed to strategically perturb the physical measurements such that the UIO's detection capability can be significantly enhanced while not severely degrading the control performance. The watermark strengths, UIO parameters, and control gains are optimally co-designed to balance the trade off between detection effectiveness and control performance. The watermarking-enhanced UIO's robustness to TSEs is improved by adopting a sliding time window with appropriate length. The results of Matlab/Simulink studies and cyber-physical co-simulation experiments indicate that the proposed method can rapidly detect P-FDIAs and S-FDIAs without degrading the control performance, and the watermark strength and detection window's length need to be carefully designed to obtain acceptable detection latency and desired TSE robustness. Future perspectives include: 1) Distinguishing S-FDIAs from P-FDIAs when both types of attacks exist, 2) Extending the watermarking-enhanced UIO-based detector to wider industrial scenarios like AC micorgrids, 3) Investigating detection-triggered mitigation schemes to reduce or eliminate the impact from malicious bias injection attacks.

Acknowledgement

This work was supported in part by the UK Research and Innovation Future Leaders Fellowship entitled 'Digitalisation of Electrical Power and Energy Systems Operation' under Grant MR/W011360/2. The authors would like to thank anonymous reviewers very much for their fruitful and insightful comments during the revision of this manuscript. The authors would also like to thank Dr. Ross Drummond from University of Sheffield for his valuable suggestions during the proofreading of this manuscript.

References

[1] J. M. Guerrero, J. C. Vasquez, J. Matas, L. G. de Vicuna, and M. Castilla, "Hierarchical control of droop-controlled ac and dc microgrids—a general approach toward standardization," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 1, pp. 158–172, Jan 2011.

[2] F. R. Badal, P. Das, S. K. Sarker, and S. K. Das, "A survey on control issues in renewable energy integration and microgrid," *Protection and Control of Modern Power Systems*, vol. 4, no. 1, pp. 1–27, 2019.

[3] Z. Cheng *et al.*, "To centralize or to distribute: That is the question: A comparison of advanced microgrid management systems," *IEEE Industrial Electronics Magazine*, vol. 12, no. 1, pp. 6–24, 2018.

[4] M. Zhang, Y. Han, Y. Liu, A. S. Zalhaf, E. Zhao, K. Mahmoud, M. M. F. Darwish, and F. Blaabjerg, "Multi-timescale modeling and dynamic stability analysis for sustainable microgrids: State-of-the-art and perspectives," *Protection and Control of Modern Power Systems*, vol. 9, no. 3, pp. 1–35, 2024.

[5] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 ukraine blackout: Implications for false data injection attacks," *IEEE transactions on power systems*, vol. 32, no. 4, pp. 3317–3318, 2016.

[6] M. Egan, "A retrospective on 2022 cyber incidents in the wind energy sector and building future cyber resilience," Master's thesis, Boise State University, 2022.

[7] S. Tan, J. M. Guerrero, P. Xie, R. Han, and J. C. Vasquez, "Brief survey on attack detection methods for cyber-physical systems," *IEEE Systems Journal*, vol. 14, no. 4, pp. 5329–5339, 2020.

[8] M. Liu, F. Teng, Z. Zhang, P. Ge, M. Sun, R. Deng, P. Cheng, and J. Chen, "Enhancing cyber-resiliency of der-based smart grid: A survey," *IEEE Transactions on Smart Grid*, pp. 1–1, 2024.

[9] H.-J. Liao, C.-H. R. Lin, Y.-C. Lin, and K.-Y. Tung, "Intrusion detection system: A comprehensive review," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 16–24, 2013.

[10] A. Ameli, A. Hooshyar, A. H. Yazdavar, E. F. El-Saadany, and A. Youssef, "Attack detection for load frequency control systems using stochastic unknown input estimators," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 10, pp. 2575–2590, 2018.

[11] Q. Yang *et al.*, "On optimal pmu placement-based defense against data integrity attacks in smart grid," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 7, pp. 1735–1750, 2017.

[12] B. Li, R. Lu, W. Wang, and K.-K. R. Choo, "Ddoa: A dirichlet-based detection scheme for opporftunistic attacks in smart grid cyber-physical system," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 11, pp. 2415–2425, 2016.

[13] S. Tan, P. Xie, J. M. Guerrero, and J. C. Vasquez, "False data injection cyber-attacks detection for multiple dc microgrid clusters," *Applied Energy*, vol. 310, p. 118425, 2022.

[14] S. Tan, P. Xie, J. M. Guerrero, J. C. Vasquez, and R. Han, "Cyberattack detection for converter-based distributed dc microgrids: Observer-based approaches," *IEEE Industrial Electronics Magazine*, vol. 16, no. 3, pp. 67–77, 2021.

[15] S. Liu, G. Hu, M. Xia, Q. Zhang, W. Fang, and X. Liu, "Detection and mitigation via alternative data for false data injection attacks in dc microgrid clusters," *CPSS Transactions on Power Electronics and Applications*, 2023.

[16] Z. Zhang, R. Deng, D. K. Yau, P. Cheng, and J. Chen, "Analysis of moving target defense against false data injection attacks on power grid," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 2320–2335, 2019.

[17] M. Liu, C. Zhao, Z. Zhang, and R. Deng, "Explicit analysis on effectiveness and hiddenness of moving target defense in ac power systems," *IEEE Transactions on Power Systems*, vol. 37, no. 6, pp. 4732–4746, 2022.

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

[19] M. Liu *et al.*, "Converter-based moving target defense against deception attacks in dc microgrids," *IEEE Transactions on Smart Grid*, vol. 13, no. 5, pp. 3984–3996, 2022.

[20] ——, "Pddl: Proactive distributed detection and localization against stealthy deception attacks in dc microgrids," *IEEE Transactions on Smart Grid*, vol. 14, no. 1, pp. 714–731, 2023.

[21] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *2009 47th annual Allerton conference on communication, control, and computing (Allerton)*. IEEE, 2009, pp. 911–918.

[22] L. Ma, Z. Chu, C. Yang, G. Wang, and W. Dai, "Recursive watermarking-based transient covert attack detection for the industrial cps," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 1709–1719, 2023.

[23] W.-H. Ko, J. A. Ramos-Ruiz, T. Huang, J. Kim, H. Ibrahim, P. N. Enjeti, P. R. Kumar, and L. Xie, "Robust dynamic watermarking for cyber-physical security of inverter-based resources in power distribution systems," *IEEE Transactions on Industrial Electronics*, pp. 1–11, 2023.

[24] Y. Li, N. Lin, J. Wu, Y. Pan, and Y. Zhao, "Low latency attack detection with dynamic watermarking for grid-connected photovoltaic systems," *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, 2023.

[25] Z. Song, A. Skuric, and K. Ji, "A recursive watermark method for hard real-time industrial control system cyber-resilience enhancement," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 2, pp. 1030–1043, 2020.

[26] H. Zhu, M. Liu, C. Fang, R. Deng, and P. Cheng, "Detection-performance tradeoff for watermarking in industrial control systems," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 2780–2793, 2023.

[27] J. Chen, R. J. Patton, and H.-Y. Zhang, "Design of unknown input observers and robust fault detection filters," *International Journal of Control*, vol. 63, no. 1, pp. 85–105, 1996.

[28] A. J. Gallo, M. S. Turan, F. Boem, T. Parisini, and G. Ferrari-Trecate, "A distributed cyber-attack detection scheme with application to dc microgrids," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3800–3815, 2020.

[29] A. J. Gallo, M. S. Turan, F. Boem, G. Ferrari-Trecate, and T. Parisini, "Distributed watermarking for secure control of microgrids under replay attacks," *IFAC-PapersOnLine*, vol. 51, no. 23, pp. 182–187, 2018.

[30] C. M. Ahmed, J. Zhou, and A. P. Mathur, "Noise matters: Using sensor and process noise fingerprint to detect stealthy cyber attacks and authenticate sensors in cps," in *Proceedings of the 34th Annual Computer Security Applications Conference*, 2018, pp. 566–581.

[31] S. Sourav and B. Chen, "Exposing hidden attackers in industrial control systems using micro-distortions," *IEEE Transactions on Smart Grid*, vol. 15, no. 2, pp. 2089–2101, 2024.

[32] M. Tucci, S. Riverso, J. C. Vasquez, J. M. Guerrero, and G. Ferrari-Trecate, "A decentralized scalable approach to voltage control of DC islanded microgrids," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 6, pp. 1965–1979, 2016.

[33] J. Li, C. Gu, Y. Xiang, and F. Li, "Edge-cloud computing systems for smart grid: state-of-the-art, architecture, and applications," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 4, pp. 805–817, 2022.

[34] M. Baud and M. Felser, "Profinet io-device emulator based on the man-in-the-middle attack," in *2006 IEEE Conference on Emerging Technologies and Factory Automation*, 2006, pp. 437–440.

[35] P. Huitsing, R. Chandia, M. Papa, and S. Shenoi, "Attack taxonomies for the modbus protocols," *International Journal of Critical Infrastructure Protection*, vol. 1, pp. 37–44, 2008.

[36] H. Pu, L. He, P. Cheng, J. Chen, and Y. Sun, "Cormand2: A deception attack against industrial robots," *Engineering*, vol. 32, pp. 186–201, 2024.

[37] R. C. Merkle, "Secure communications over insecure channels," *Communications of the ACM*, vol. 21, no. 4, pp. 294–299, 1978.

[38] F. Callegati, W. Cerroni, and M. Ramilli, "Man-in-the-middle attack to the https protocol," *IEEE Security & Privacy*, vol. 7, no. 1, pp. 78–81, 2009.

[39] M. Liu, C. Zhao, R. Deng, P. Cheng, and J. Chen, "False data injection attacks and the distributed countermeasure in dc microgrids," *IEEE Transactions on Control of Network Systems*, pp. 1–12, 2022.

[40] F. J. MacWilliams and N. J. Sloane, "Pseudo-random sequences and arrays," *Proceedings of the IEEE*, vol. 64, no. 12, pp. 1715–1729, 1976.

[41] P. Nahata, R. Soloperto, M. Tucci, A. Martinelli, and G. Ferrari-Trecate, "A passivity-based approach to voltage stabilization in dc microgrids with zip loads," *Automatica*, vol. 113, p. 108770, 2020.

[42] J. He, P. Cheng, L. Shi, J. Chen, and Y. Sun, "Time synchronization in wsns: A maximum-value-based consensus approach," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 660–675, 2013.

[43] A. T. Al-Hammouri, "A comprehensive co-simulation platform for cyber-physical systems," *Computer Communications*, vol. 36, no. 1, pp. 8–19, 2012.