This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

A Distributed Cyber-attack Detection Scheme with Application to DC Microgrids

Alexander J. Gallo, *Student Member, IEEE*, Mustafa S. Turan, Francesca Boem, *Member, IEEE*, Thomas Parisini, *Fellow, IEEE*, Giancarlo Ferrari-Trecate, *Senior Member, IEEE*

***Abstract*—DC microgrids often present a hierarchical control architecture, requiring integration of communication layers. This leads to the possibility of malicious attackers disrupting the overall system. Motivated by this application, in this paper we present a distributed monitoring scheme to provide attack-detection capabilities for linear Large-Scale Systems. The proposed architecture relies on a Luenberger observer together with a bank of Unknown-Intput Observers (UIOs) at each subsystem, providing attack detection capabilities. We describe the architecture and analyze conditions under which attacks are guaranteed to be detected, and, conversely, when they are *stealthy*. Our analysis shows that some classes of attacks cannot be detected using either module independently; rather, by exploiting both modules simultaneously, we are able to improve the detection properties of the diagnostic tool as a whole. Theoretical results are backed up by simulations, where our method is applied to a realistic model of a low-voltage DC microgrid under attack.**

***Index Terms*—Electrical Power Systems, fault detection, cooperative control, estimation, cyber-attack detection**

I. Introduction and Problem Statement

A. Objectives and Contributions

Hierarchical control architectures are an established solution for the regulation of DC microgrids (DCmGs) [1], allowing for local stabilization, as well as cooperation among subsystems, for the achievement of global control objectives. In this scenario, coordination is enabled through the introduction of a communication network, enabling information transfer between Distributed Generation Units (DGUs). This in turn leads to the possibility of malicious agents interfering with transmitted data, altering the behavior of the overall system.

This work has been partially supported by European Union's Horizon 2020 research and innovation programme under grant agreement No 739551 (KIOS CoE) and by the Italian Ministry for Research in the framework of the 2017 Program for Research Projects of National Interest (PRIN), Grant no. 2017YKXYXJ. This work has also been conducted as part of the Swiss National Science Foundation under the COFLEX project (grant number 200021_169906).

A. J. Gallo is with the Department of Electrical and Electronic Engineering at the Imperial College London, UK. Email: alexander.gallo12@imperial.ac.uk

M. S. Turan, G. Ferrari-Trecate are with the Automatic Control Laboratory, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. Email: {mustafa.turan, giancarlo.ferraritrecate}@epfl.ch

F. Boem is with the Department of Electronic and Electrical Engineering, University College London, UK. Email: f.boem@ucl.ac.uk

T. Parisini is with the Department of Electrical and Electronic Engineering at the Imperial College London, UK, the KIOS Research and Innovation Centre of Excellence, University of Cyprus, and the Department of Engineering and Architecture at University of Trieste, Italy. Email: t.parisini@gmail.com

DC microgrids can be seen as a typical case of Large-Scale and Cyber-Physical System (LSS and CPS respectively), given both their size and the integration of information technology resources required to effectively achieve control (e.g. an information network). These systems can be decomposed into multiple interconnected units [2] interacting both through physical and cyber links. Similarly to DC microgrids, motivated by size, complexity, and the need to embed *scalability* in the control architecture, distributed control has been widely used to design scalable regulation schemes for LSSs.

The main objective of this paper is to design a model-based attack detection strategy for LSSs that is *distributed* and *scalable*. Specifically, we require that each subsystem be equipped with its own local diagnoser, and that the information needed for the design and operation of the monitor be limited to a subset of the LSS. Note that in some works on secure estimation and detection (e.g. [3], [4], [5]) the term "distributed" is used with a different meaning. Indeed, it refers to the scenario where multiple sensors observe the same system and each local estimator aims at reconstructing the global state of the system.

In this paper, we consider linear LSSs and propose a novel distributed monitoring architecture devoted to the timely detection of attacks on the information network connecting subsystems of linear LSS, relying on two modules, exploiting:

- a bank of Unknown-Input Observers (UIO);
- a distributed Luenberger observer,

as will be further illustrated in Section III. These two modules exploit different sets of relations and different model knowledge to perform detection, thus compensating each other's vulnerabilities, and reducing the number of attacks that are stealthy. In fact, while the Luenberger observer of the local state exploits analytical relations from the physical interconnection between subsystems to perform detection, the UIOs estimating the neighbors' states exploit knowledge of the model of the neighbors themselves. This difference proves critical in the analysis of the properties of each module, as it determines both the classes of attacks that are guaranteed to be detected and, more importantly, the classes of attacks that cannot be detected by each module independently. Indeed, the simultaneous use of both modules reduces the classes of attacks that are stealthy to each local detector.

The main contributions of this work are:

a. to design a *local* monitoring unit $\mathcal{D}_i$ for the i-th subsystem, to detect attacks on the communication network;
b. to propose a distributed and scalable design technique in which the synthesis of $\mathcal{D}_i$ requires at most information from

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

neighbors of subsystem i;

c. to provide theoretical results on detectability and stealthiness properties of the proposed attack detection scheme, given bounds on unknown disturbances influencing both subsystem dynamics and measurements;
d. to introduce a state augmentation technique to improve the detection capabilities of the UIO-based module;
e. to validate the monitoring scheme through analysis and simulations using a realistic model of a DCmG.

In the following, we present the model of interconnected subsystems composing an LSS. Then, the detection problem is formally given, and the state of the art on security in distributed control systems is summarized (with an emphasis on smart grids and power networks).

Some preliminary and partial works have been presented in the conference papers [6], [7]. In this work, we design a novel monitoring scheme extending the estimation framework where two estimators exploiting different sets of information are used to detect cyber attacks. We also provide thorough analysis as far as *detectable* and *stealthy* attacks are concerned. Furthermore, the improvements in terms of detectability that are achieved using the two modules simultaneously in a *stacked* configuration are demonstrated both analytically and through simulations.

B. Problem formulation

1) Large-Scale Sytems: Motivated by the example of DCmGs, which is structured as a set of interconnected DGUs, we model an LSS as a network of N subsystems $\mathcal{S}_i$, each coupled with a set of *neighbors* $\mathcal{N}_i \subseteq \mathcal{N} \triangleq \{1, \dots, N\}$, $N_i \triangleq |\mathcal{N}_i|$. The dynamics of each subsystem can be written as:

$$\mathcal{S}_i : \begin{cases} \dot{x}_{[i]} = A_{ii}x_{[i]} + B_i u_{[i]} + M_i d_{[i]} + \xi_{[i]} + w_{[i]} \\ y_{[i]} = C_i x_{[i]} + \rho_{[i]} \end{cases}, \quad (1)$$

where $x_{[i]} \in \mathbb{R}^{n_i}$, $u_{[i]} \in \mathbb{R}^{m_i}$, $d_{[i]} \in \mathbb{R}^{g_i}$, $y_{[i]} \in \mathbb{R}^{p_i}$ are respectively the subsystem state, control and exogenous input, and output; $\xi_{[i]} \in \mathbb{R}^{n_i}$ represents the physical interconnection between subsystems, defined as $\xi_{[i]} \triangleq \sum_{j \in \mathcal{N}_i} A_{ij} x_{[j]}$, while $w_{[i]} \in \mathbb{R}^{n_i}$ and $\rho_{[i]} \in \mathbb{R}^{p_i}$ model process and measurement disturbances. In Section II we show how the dynamics of DC microgrids can be modeled as in (1) [8].

Assumption 1. *For all $\mathcal{S}_i$, the pair (C_i, A_{ii}) is observable.* ◁

Assumption 2. *Process and measurement disturbances $w_{[i]}(t)$ and $\rho_{[i]}(t)$ are unknown but bounded, i.e.*

$$|w_{[i]}(t)| \leq \bar{w}_{[i]}, \quad |\rho_{[i]}(t)| \leq \bar{\rho}_{[i]}, \quad (2)$$

for all $t \geq 0$, where $\bar{w}_{[i]}, \bar{\rho}_{[i]} > 0, \forall i \in \mathcal{N}$, are known. ◁

We consider the control input $u_{[i]}$ to be the result of a *distributed* control architecture, depending directly on communicated variables $y^c_{[j,i]}$ that $\mathcal{S}_i$ *receives* from its neighbors. Here $y^c_{[j,i]}$ is used to differentiate the output $y_{[j]}$ *locally* available to $\mathcal{S}_j$ from the information that $\mathcal{S}_i$ receives. We assume that the communication network shares the same topology of the LSS, and we consider that it is *ideal*, i.e. that it is not affected by non-idealities, such as delays and packet drops, among others.

2) Model of cyber-attack: The necessity of integrating a communication network in the control architecture of a LSS may expose the system to cyber-security threats [9]. The information received by $\mathcal{S}_i$ from $\mathcal{S}_j$ is written as:

$$y^c_{[j,i]}(t) \triangleq y_{[j]}(t) + \beta_{j,i}(t - T_a^{j,i})\phi_{j,i}(t), \forall t \geq 0 \quad (3)$$

where $\beta_{j,i}(t)$ is an activation function, $\phi_{j,i}(t)$ is an attack function, as defined by the attacker to achieve some unknown objective, and $T_a^{j,i} > 0$ is the unknown initial time of attack. The activation function can be any function of time satisfying $\beta_{j,i}(t) = 0$, $\forall t < 0$ and $\beta_{j,i}(t) \neq 0$, $\forall t \geq 0$. Readers are referred to [6] for possible choices of this function. Note that, in nominal conditions (i.e. for $t < T_a^{i,j}$), the information received by $\mathcal{S}_i$ from $\mathcal{S}_j$ is the exact measurement vector, i.e. $y^c_{[j,i]}(t) = y_{[j]}(t)$.

Assumption 3. *Each edge $(i,j), \forall i, j \in \mathcal{N}$ is affected by at most one attack, and $T_a^{j,i} > 0, \forall i, j \in \mathcal{N}$.* ◁

Remark 1. Assumption 3 is not very restrictive, as it does not exclude the occurrence of complex attacks targeting multiple lines simultaneously. ◁

Through appropriate definition of $\phi_{j,i}(t)$ in (3), it is possible to model different types of attacks [10], such as: *false data injection attacks*, where $\phi_{j,i} : \mathbb{R} \to \mathbb{R}^{n_i}$ is any attacker-defined function of time; *covert attacks*, where an attack of the form $\phi_{j,i}(t) \triangleq -y_{[j]}(t) + y^a_{[j]}(t)$ replaces the transmitted information with the output $y^a_{[j]}(t)$ of a *simulated system* with the same dynamics as $\mathcal{S}_j$; *replay attacks*, where transmitted information $y_{[j]}(t)$ is stored and then *replayed* periodically by the attacker, hiding any changes in operating condition of $\mathcal{S}_i$, and where $\phi_{j,i}(t) \triangleq -y_{[j]}(t) + y_{[j]}(t - nT)$, with $n \in \mathbb{N}$ modeling the periodicity of the attack.

Remark 2. In the context of this work, differently from others in the literature, we only consider attacks on the variables which are communicated between subsystems. Thus, both the local measurement $y_{[i]}$ and the control input $u_{[i]}$ are considered to be secure. This is motivated by the DC microgrid application, where controllers are colocated with the sensors and actuators interfacing the system. ◁

3) Attack detection: We now formulate the problem of *attack detection*. We define the activation time of the first attack on the incoming communication channels of a subsystem:

$$\check{T}_a^i \triangleq \min_{j \in \mathcal{N}_i} T_a^{j,i}.$$

Problem 1 (Attack Detection)**.** Design, for each subsystem, an attack detector $\mathcal{D}_i$ to verify the null hypothesis at time t:

$$\mathcal{H}_i^0(t) : \{y^c_{[j,i]}(t) = y_{[j]}(t), \forall j \in \mathcal{N}_i\}, \quad (4)$$

i.e. the received communication is not under attack. ▶

C. State of the art

The design and analysis of monitoring schemes to detect cyber-attacks for CPSs have attracted great interest in the literature, as demonstrated by the recent special issue [11], as well as the surveys [12], [13] and references cited therein. This

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

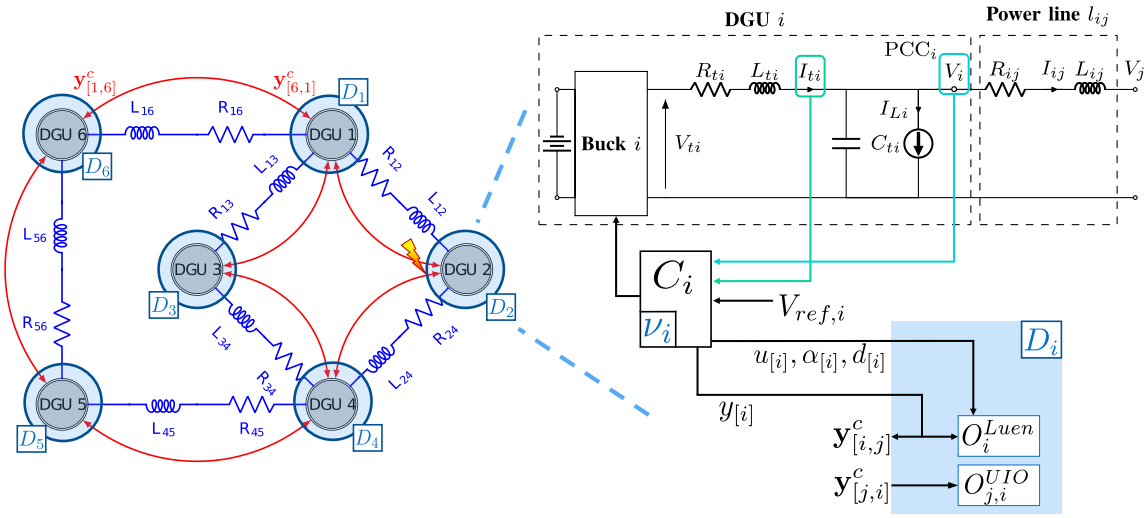


Fig. 1: Diagram representation of the DC microgrid. On the left, the graph representing the DCmG; the physical interconnections are shown as the blue power lines, and the communication topology appears as the red arrows. Cyber attacks are directed at the communication lines. On the right, the circuit diagram of a DGU, together with the information structure of the detector D_i.

is due to the fact that modern control systems are evermore exposed to cyber-attacks, given the increasing integration of physical and *cyber* resources in CPS control loops [14]. An area that has received specific attention because of its criticality has been the secure control and estimation of power networks, with specific focus on smart grids [9], [15], [16], [17], and *microgrids* [18], [19], [20], [21]. Among the works addressing the security problem in microgrids, [20], [21] offer techniques to detect cyber-attacks in DCmGs. In particular, the authors of [20] exploit *Signal Temporal Logic* (STL) to detect whether an attack is present, by verifying whether given STL requirements are violated. In [21], on the other hand, the authors consider "balanced" attacks, and define a Cooperative Vulnerability Factor (CVF) for detection, exploiting secure knowledge of control inputs of neighboring DGUs.

In the context of secure control, as highlighted recently in [22], attack detection and resilience schemes can be often divided in *data driven* and *knowledge-based* approaches. We here focus on the latter, without the pretence of providing an exhaustive survey of the literature, as it is out of the scope of this paper. Many knowledge-based techniques are available, most of which have focused on centralized architectures to detect malicious intrusion and tampering of the communication between plants and controllers [23], [24], [25], [26]. However, centralized methodologies are known to be undesirable in the context of microgrids, as they are not scalable and cannot easily incorporate addition, removal, and replacement of DGUs.

Although the limitations to centralized architectures for CPS are well known, few works propose distributed methods, of which [4], [5], [27], [28], [29] are examples, but often requiring additional assumptions. For instance, [28], [29] suppose secure communication between different monitoring units. In [27] the differences between centralized and decentralized architectures in cyber-attack detection are analyzed in the context of stochastic interconnected systems. Finally, [4], [5] present *distributed* detection methods in which locally available information is exploited to estimate the global state of the system. These approaches share similarities with methods proposed for secure distributed state estimation, such as [3], where the global state of an LSS is reconstructed from partial measurements in the presence of cyber attacks. Differently to the objectives of the present paper, all of these methods, while referred to as "distributed", require the knowledge of the whole dynamics of the CPS for design and implementation. Finally, note that, apart from [27], none of the previously mentioned works include system and measurement disturbances in their modeling.

It is worth noting that attack detection methods can be inspired by Fault Detection and Isolation (FDI) algorithms, for which distributed solutions have indeed been recently proposed [30], [31], [32], [33], [34], [35]. Of these, [32] proposes an FDI architecture based on a bank of UIOs to detect faults on either subsystems or interconnections. An analysis of the differences between fault and cyber-attack detection is provided in [36].

As anticipated, the main objective of this paper is to provide a scalable design procedure for a novel distributed attack detection scheme solving Problem 1.

D. Organization of the Paper

In Section II we present the model of a low-voltage islanded DC mircogrid. In Section III, we illustrate the attack detection architecture, in which $\mathcal{D}_i$ utilizes parallel modules to solve Problem 1. In Sections IV and V we analyze the properties of the modules individually, in terms of *detectable* and *stealthy* attacks. In Section VI we evaluate the detectability properties of $\mathcal{D}_i$ as a whole, thus showing the benefits of combining the two modules. In Section VII, extensive results from numerical simulations using realistic dynamics of a DC microgrid are given, and the effectiveness of the strategy demonstrated.

Notation

In the paper, the operator $|\cdot|$ applied to a set determines its cardinality, while used with matrices or vectors it defines

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

their component-by-component absolute value. The operator $\|\cdot\|$ is used to define the matrix 2-norm. In general, in this paper inequalities are considered component-by-component. $\mathbf{I}$ and $\mathbf{0}$ represent the identity matrix and a matrix or vector of zeros, each of the appropriate dimensions. For two matrices A and B with the same dimensions, $A \geq B$ indicates the element-wise inequality; the same is considered for vectors. With $\mathrm{col}(\cdot)$, $\mathrm{diag}(\cdot)$, and $\ker(\cdot)$ we define the column and diagonal concatenation of vectors or matrices, and the null-space of a matrix. For a matrix A, $A^\dagger$ denotes its right inverse.

II. Low-Voltage Islanded DC Microgrids

A. Modeling low-voltage DC Microgrids

Microgrids, both AC and DC, are a promising technology for future power networks, as they offer the possibility of merging distributed energy generation, consumption and storage. This is important, given the growing penetration of renewable energy sources in electrical grids. We focus here on low-voltage islanded DC microgrids, that provide an attractive solution for energy distribution, as many renewable energy sources, energy storage technologies, and loads are inherently DC [1]. Nowadays, DCmGs find applications in e.g. data centers, smart houses, and electric vehicle charging stations.

As shown in Figure 1, a low-voltage DCmG can be represented as a network of N interconnected DGUs, each composed of a Buck converter, interfacing a variable DC voltage source with the rest of the network through an RLC filter. We assume that loads are connected to the DGU terminals[1], and DGUs are coupled through resistive lines. The interconnected dynamics of DGU i can be written as in (1), with state $x_{[i]} \triangleq [V_i, I_{ti}, \nu_i]^\top$ (where ν_i is an integrator state internal to the controller, used for reference voltage tracking), exogenous input $d_{[i]} \triangleq I_{Li}$, and input $u_{[i]} = [V_{ti}, \Delta V_i]^\top$, where ΔV_i is the result of a secondary control layer (e.g. a consensus protocol) used for current sharing across the network, and V_{ti} is the switching terminal voltage of the Buck converter. The specific definitions of the matrices in (1) can be found in Appendix A, and the interested reader is referred to [8] for further details.

Remark 3. In the literature, the design of controllers for DC microgrids with DC-DC converters often relies on the so-called *state-space averaging method*, to disregard the switching behavior of the terminal input [37]. It is therefore possible to define an average control input $V_{ti}^{avg} \triangleq \delta_i V_{si}$, where $\delta_i \in [0,1]$ is the duty cycle of the Buck converter and $V_{si} \in \mathbb{R}$ is the voltage of its power source. In this paper we suppose V_{si} is sufficiently large to avoid saturation of δ_i. ◁

Assumption 4. *For every DGU $i \in \mathcal{N}$, $C_i = \mathbf{I}$ and the measurement is affected by an unknown disturbance $\rho_{[i]}$.* ◁

Assumption 4 is not restrictive as V_i and I_{ti} can be measured within the DGU, and ν_i is an internal state of the controller.

[1]If load buses appear elsewhere, they can be mapped to the output terminals of DGUs using Kron reduction [8].

B. Controller architecture

The control strategies proposed for islanded DCmGs are often designed in the context of hierarchical architectures (see the review [1], and the references cited therein), where primary controllers within the DGUs guarantee global stability [8], [38], while secondary and tertiary controllers achieve different operational objectives, such as current and power sharing, microgrid synchronization, and overall energy management [39], [38], [40], [41]. In this paper, we consider that each DGU is controlled by primary and secondary controllers defined as in [8] and [39] respectively. Our choice is motivated by the fact that these controllers can be designed in a scalable fashion while providing stability of the whole DCmG.

Specifically, the schemes presented in [8], [39] define control laws to respectively compute the average terminal voltage V_{ti}^{avg} (and thus δ_i) to obtain global voltage stability, and the secondary control input ΔV_i, to achieve current sharing, by employing a consensus protocol reliant on neighbors' communicated outputs (3). To achieve coordination across the whole DCmG, reliable communication between DGUs is necessary. Thus, cyber-attacks can easily alter the operating point of the DCmG as a whole.

We note that in this paper we consider the case of *islanded* DCmGs. In the case of grid connection, DCmGs provide ancillary services to the main grid, typically through the use of an energy management system (EMS). In recent years, distributed optimization methods have been presented for distributed EMSs [42], which may be tackled with the distributed detection scheme here proposed.

III. Attack detector $\mathcal{D}_i$ – Detection architecture

As previously mentioned, the proposed detection architecture, illustrated in Figure 1, relies on two modules simultaneously estimating the state of the local subsystem (through a Luenberger observer) and the states of the neighboring subsystems (with a bank of N_i UIOs). The bank of UIOs compute an estimate $\hat{\mathbf{x}}_{[j,i]}(t)$ of a suitably defined augmented state $\mathbf{x}_{[j]}$ for each neighbor of $\mathcal{S}_i$, whilst the Luenberger-observer-based module generates an estimate $\hat{x}_{[i]}(t)$ of its state $x_{[i]}(t)$. The augmented state $\mathbf{x}_{[j]}$ and communicated output measurement $\mathbf{y}^c_{[j,i]}$ required for the design of the UIO-based modules in $\mathcal{D}_i$ are introduced in Section IV. The output estimates are then compared respectively to $\mathbf{y}^c_{[j,i]}$ and $y_{[i]}$, and the resulting residual is then used to detect the presence of an attack, by evaluating the following inequalities:

$$\underbrace{\left|\mathbf{y}^c_{[j,i]}(t) - \mathbf{C}_j \hat{\mathbf{x}}_{[j,i]}(t)\right|}_{|r_{[j,i]}(t)|} \leq \bar{r}_{[j,i]}(t), \quad \forall j \in \mathcal{N}_i \tag{5a}$$

$$\underbrace{\left|y_{[i]}(t) - C_i \hat{x}_{[i]}(t)\right|}_{|r_{[i]}(t)|} \leq \bar{r}_{[i]}(t) \tag{5b}$$

where matrix $\mathbf{C}_j$ is defined in (11) and the thresholds $\bar{r}_{[j,i]}(t)$ and $\bar{r}_{[i]}(t)$ are defined appropriately to prevent false alarms, based on knowledge of the disturbance bounds in (2). This design choice, albeit guaranteeing that the process will not be interrupted without a certified threat, also implies that the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

TABLE I: Information required for design of $\mathcal{D}_i$ and attack detection

	UIO$(j,i), \forall j \in \mathcal{N}_i$	**Luenberger observer**
Offline Information	Matrices $\mathbf{A}_{jj}$,$\mathbf{E}_j$,$\mathbf{C}_j$, bounds $\bar{w}_{[j]}$ and $\bar{\rho}_{[j]}$	Model of $\mathcal{S}_i$, matrix $\mathbf{C}_j$, and bounds $\bar{w}_{[i]}$ and $\bar{\rho}_{[k]}$, for all $k \in \{i\} \cup \mathcal{N}_i$
Online Information	Communicated measurements $\mathbf{y}^c_{[j,i]}(t)$	Local measurements and inputs, and communicated measurements $\mathbf{y}^c_{[j,i]}(t)$

Algorithm 1 Attack detection and isolation at time t

while $\mathcal{S}_i$ *online* $\forall i \in \mathcal{N}$ **do**
  Update estimates $\hat{\mathbf{x}}_{[j,i]}(t), \forall j \in \mathcal{N}_i$ and $\hat{x}_{[i]}(t)$;
  Update bounds $\bar{r}_{[j,i]}(t), \forall j \in \mathcal{N}_i$ and $\bar{r}_{[i]}(t)$;
  Compute residuals $r_{[j,i]}(t), \forall j \in \mathcal{N}_i$ and $r_{[i]}(t)$;
  Evaluate (5a) and (5b)
  if (5a) and (5b) hold **then**
    No attack is detected at time t
  else
    if $|r_{[j,i]}(t)| > \bar{r}_{[j,i]}(t)$ for any $j \in \mathcal{N}_i$ **then**
      Attack detected on link (j,i)
    else
      Attack detected, no link is isolated
    end if
  end if
end while

thresholds are possibly conservative. If at any time $t > \check{T}^i_a$ either of the inequalities in (5) is violated, an attack is detected by $\mathcal{D}_i$. Moreover, if (5a) is violated, the attacked communication line is also isolated. The operation of the detection logic is summarized in Algorithm 1, while in Table I we highlight the information required by $\mathcal{D}_i$ at design time (*offline*), and during normal operations (*online*).

As shown in Table I, the two modules exploit different model knowledge to detect the presence of cyber-attacks. Specifically, each UIO exploits knowledge of augmented dynamics of $\mathcal{S}_j$ (i.e. matrices $\mathbf{A}_{jj}, \mathbf{E}_j, \mathbf{C}_j$) to estimate the state of each of its neighbors $\mathcal{S}_j, j \in \mathcal{N}_i$ from $\mathbf{y}^c_{[j,i]}$. This allows for detection of false data injection attacks, while being vulnerable to replay and covert attacks. On the other hand, the Luenberger-observer-based detection module uses knowledge of dynamics of $\mathcal{S}_i$ (1) to exploit the physical interconnections between subsystems, thus detecting attacks with analytical relations to the local dynamics.

The detector $\mathcal{D}_i$, by combining the two modules in the same *stacked* architecture and having them run simultaneously, as illustrated in Algorithm 1, is capable of detecting attacks that would be stealthy to either of the modules independently as will be analytically presented in Section VI.

We now focus on the appropriate design of the two observer-based modules, the definition of thresholds $\bar{r}_{[j,i]}(t)$ and $\bar{r}_{[i]}(t)$, and analyze their individual properties.

Remark 4. As can be seen from Algorithm 1 and Table I, the design and operation of $\mathcal{D}_i$ rely at most on information from the set of neighbors $\mathcal{N}_i$, and are therefore distributed, as well as scalable with the number of subsystems in the network. ◁

IV. Bank of Unknown Input Observers

A. Design of the detection module

We first focus on the design and properties of $\mathcal{O}^{UIO}_{j,i}$, the UIO-based detection modules estimating the state of neighboring subsystems. UIOs are a class of observers designed to algebraically decouple the residual error from a vector of unknown inputs [43]. This proves fundamental for $\mathcal{D}_i$ to estimate the state $x_{[j]}$, $j \in \mathcal{N}_i$, as $\mathcal{S}_i$ does not have access to the inputs affecting the dynamics of its neighbors. To design the UIOs we rewrite the dynamics of $\mathcal{S}_j$ in (1) as:

$$\begin{aligned} \dot{x}_{[j]} &= A_{jj}x_{[j]} + \bar{E}_j\bar{d}_{[j]} + w_{[j]} \\ y_{[j]} &= C_j x_{[j]} + \rho_{[j]} \end{aligned}, \tag{6}$$

where $\bar{E}_j\bar{d}_j = \xi_{[j]} + B_j u_{[j]} + M_j d_{[j]}$ represents the effect of the unknown inputs on $x_{[j]}$. The matrix $\bar{E}_j \in \mathbb{R}^{n_j \times q_j}, q_j \leq n_j$ links the unknown inputs to the dynamics of $\mathcal{S}_j$, its columns consisting of a basis of the range of matrix $E_j \triangleq \begin{bmatrix} A_{jk_1}, \dots, A_{jk_{N_j}}, B_j, M_j \end{bmatrix}$, where $\{k_1, \dots, k_{N_j}\} = \mathcal{N}_j$ are the indices of the neighbors of $\mathcal{S}_j$. This definition ensures that $\bar{E}_j$ is full column rank, as required by [43]. The term $\bar{d}_{[j]}(t) \triangleq \hat{E}_j\hat{d}_{[j]}$ is a linear combination of $\hat{d}_{[j]}$, defined as

$$\hat{d}_{[j]} \triangleq \begin{bmatrix} x^\top_{[k_1]}, \dots, x^\top_{[k_{N_j}]}, u^\top_{[j]}, d^\top_{[j]} \end{bmatrix}^\top, \tag{7}$$

i.e. the vector containing all inputs to $\mathcal{S}_j$ unknown to $\mathcal{D}_i$. Matrix $\hat{E}_j$ is derived, following the choice of $\bar{E}_j$, such that $\bar{E}_j\hat{E}_j = E_j$, and is not relevant to the design of the UIOs.

The full-order UIO state and state estimate of $\mathcal{S}_j$ can be defined as follows [43]:

$$\begin{aligned} \dot{z}_{[j,i]}(t) &= F_j z_{[j,i]}(t) + \widehat{K}_j y^c_{[j,i]}(t) \\ \hat{x}_{[j,i]}(t) &= z_{[j,i]}(t) + H_j y^c_{[j,i]}(t) \\ \hat{y}_{[j,i]}(t) &= C_j \hat{x}_{[j,i]}(t) \end{aligned}, \tag{8}$$

where the matrices are defined as in [43] and are such that:

$$(H_jC_j - \mathbf{I})\bar{E}_j = \mathbf{0} \tag{9a}$$
$$S_j = \mathbf{I} - H_jC_j \tag{9b}$$
$$F_j = S_jA_{jj} - \widetilde{K}_jC_j \tag{9c}$$
$$\bar{K}_j = F_jH_j \tag{9d}$$
$$\widehat{K}_j = \widetilde{K}_j + \bar{K}_j \tag{9e}$$

The definition of S_j through design of H_j (9a)-(9b) decouples the residual error $r_{[j,i]}(t) \triangleq y^c_{[j,i]}(t) - \hat{y}_{[j,i]}(t)$ from the unknown input vector $\bar{d}_{[j]}$, while matrix $\widetilde{K}_j$ is such that F_j in (9c) is Hurwitz stable. The following necessary and sufficient conditions are given in [43] to verify the possibility of designing the UIO (8):

$$\text{rank}(C_j\bar{E}_j) = \text{rank}(\bar{E}_j); \tag{C1}$$
$$\text{the pair } (C_j, S_jA_{jj}) \text{ is detectable.} \tag{C2}$$

These two conditions need to be satisfied for a generic system of the form (6) in order to employ the proposed detection methodology. The following guarantees that these conditions are met in the special case of microgrids:

Remark 5. Given Assumption 4, $C_j = \mathbf{I}$, and thus conditions (C1) and (C2) are satisfied. ◁

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

Lemma 1. *Consider a subsystem with dynamics in* (6) *such that* (C1) *and* (C2) *hold, and a UIO with dynamics as in* (8). *If rank*$(C_j) =$ *rank*$(\bar{E}_j) = q_j$ *the residual* $r_{[j,i]} = y^c_{[j,i]} - \hat{y}_{[j,i]}$ *is independent of the attack function* $\phi_{[j,i]} \neq \mathbf{0}$ *at all times.* □

Proof. The proof is provided in Appendix B ■

Given the results stated in Lemma 1, in order to design an attack detection architecture, it is necessary either to reduce the number of unknown inputs (which may not be feasible), or increase the output information transmitted. To address the latter, additional sensors providing independent measurements could be added, although this may not be possible depending on the application. Rather, here we augment the transmitted information such that the original output $y_{[j]}$ can be reconstructed – as it is necessary for control purposes – and it represents the output of a dynamical system known to $\mathcal{O}^{UIO}_{i,j}$.

Let us hence introduce the following augmented state variable $\mathbf{x}_{[j]} = \left[x^{art}_{[j]}, e^{art}_{[j]}\right]$, with $x^{art}_{[j]}$ some *artificial* state the dynamics of which is known to $\mathcal{O}^{UIO}_{i,j}, \forall i \in \mathcal{N}_j$ and simulated by $\mathcal{S}_j$, and $e^{art}_{[j]} \triangleq x_{[j]} - x^{art}_{[j]}$. By construction,

$$\begin{aligned} x_{[j]} &= \begin{bmatrix} \mathbf{I} & \mathbf{I} \end{bmatrix} \mathbf{x}_{[j]} \\ y_{[j]} &= \begin{bmatrix} C_j & C_j \end{bmatrix} \mathbf{x}_{[j]} + \rho_{[j]} \end{aligned}, \tag{10}$$

allowing for reconstruction of $y_{[j]}$. Let us define the dynamics of $\mathbf{x}_{[j]}$, and hence $x^{art}_{[j]}$, as:

$$\begin{aligned} \dot{\mathbf{x}}_{[j]} &= \begin{bmatrix} A^{art}_{jj} & \mathbf{0} \\ A_{jj} - A^{art}_{jj} & A_{jj} \end{bmatrix} \mathbf{x}_{[j]} + \begin{bmatrix} \mathbf{E}_{j,1} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{j,2} \end{bmatrix} \begin{bmatrix} \mathbf{d}_{[j,1]} \\ \mathbf{d}_{[j,2]} \end{bmatrix} + \\ &\quad + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} w_{[j]} = \mathbf{A}_{jj}\mathbf{x}_{[j]} + \mathbf{E}_j \mathbf{d}_{[j]} + \tilde{\mathbf{w}}_{[j]} \\ \mathbf{y}_{[j]} &= \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ C_j & C_j \end{bmatrix} \mathbf{x}_{[j]} + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \rho_{[j]} = \mathbf{C}_j \mathbf{x}_{[j]} + \boldsymbol{\rho}_{[j]} = \begin{bmatrix} x^{art}_{[j]} \\ y_{[j]} \end{bmatrix}. \end{aligned} \tag{11}$$

where $A^{art}_{jj} \in \mathbb{R}^{n_j \times n_j}$ is any Hurwitz stable matrix. Nonzero matrices $\mathbf{E}_{j,1}$ and $\mathbf{E}_{j,2}$ are constructed such that $[\mathbf{E}_{j,1}, \mathbf{E}_{j,2}] = \bar{E}_j$, up to column permutations, and unknown input vectors $\mathbf{d}_{[j,1]}$ and $\mathbf{d}_{[j,2]}$ satisfy $[\mathbf{E}_{j,1}, \mathbf{E}_{j,2}]\,[\mathbf{d}^\top_{[j,1]}, \mathbf{d}^\top_{[j,2]}]^\top = \bar{E}_j \bar{d}_j$. Additionally, the following hold by construction: $\text{rank}(\mathbf{E}_{j,1}) < n_j$, $\text{rank}(\mathbf{E}_{j,2}) < q_j$, and $\text{Im}(\mathbf{E}_{j,1}) \subset \text{Im}(\bar{E}_j)$, $\text{Im}(\mathbf{E}_{j,2}) \subset \text{Im}(\bar{E}_j)$. Finally note that, as $x^{art}_{[j]}$ is *simulated* by $\mathcal{S}_j$, it is fully available and therefore appears in $\mathbf{y}_{[j]}$. We then redefine the communicated measurement in (3) as

$$\mathbf{y}^c_{[j,i]}(t) \triangleq \mathbf{y}_{[j]}(t) + \beta_{j,i}(t - T^{j,i}_a)\boldsymbol{\phi}_{j,i}(t) \tag{12}$$

with $\boldsymbol{\phi}_{j,i}(t) \triangleq [\varphi^\top_{j,i}(t), \phi^\top_{j,i}(t)]^\top \in \mathbb{R}^{n_j + p_j}$, where $\varphi_{j,i}(t)$ is the attack influencing the communicated artificial state. We note that the transmitted information, as seen in (12), is redefined to include both the output measurements, and the artificial state. In the following, we show how through state and output augmentation (11) necessary condition in Lemma 1 is satisfied.

Lemma 2. *If* (C1) *and* (C2) *hold for* $(A_{jj}, C_j, \bar{E}_j)$, *then they are also satisfied for* $(\mathbf{A}_{jj}, \mathbf{C}_j, \mathbf{E}_j)$. *If, additionally,* $\text{rank}(C_j) = \text{rank}(\bar{E}_j)$, *then* $\text{rank}(\mathbf{C}_j) > \text{rank}(\mathbf{E}_j)$. □

Proof. Condition (C1) holds given definitions of $\mathbf{C}_j$ and $\mathbf{E}_j$:

$$\mathbf{C}_j \mathbf{E}_j = \begin{bmatrix} \mathbf{E}_{j,1} & \mathbf{0} \\ C_j \mathbf{E}_{j,1} & C_j \mathbf{E}_{j,2} \end{bmatrix}, \tag{13}$$

the rank of which, being block lower triangular, is such that

$$\begin{aligned} \text{rank}(\mathbf{C}_j \mathbf{E}_j) &\geq \text{rank}(\mathbf{E}_{j,1}) + \text{rank}(C_j \mathbf{E}_{j,2}) \\ &= \text{rank}(\mathbf{E}_{j,1}) + \text{rank}(\mathbf{E}_{j,2}) = \text{rank}(\mathbf{E}_j) \end{aligned}. \tag{14}$$

Hence, noting that $\text{rank}(\mathbf{C}_j \mathbf{E}_j) \leq \min(\text{rank}(\mathbf{C}_j), \text{rank}(\mathbf{E}_j))$, it follows that $\text{rank}(\mathbf{C}_j \mathbf{E}_j) = \text{rank}(\mathbf{E}_j)$, thus satisfying (C1).

To show that (C2) is satisfied for the augmented system matrices, first note that a block-diagonal matrix $\mathbf{S}_j$ composed of blocks $S_{j,1}$ and $S_{j,2}$ can be found such that $\mathbf{S}_j \mathbf{E}_j = \mathbf{0}$. This is due to existence of solutions to $S_{j,1}\mathbf{E}_{j,1} = \mathbf{0}$ and $S_{j,2}\mathbf{E}_{j,2} = \mathbf{0}$, from $\text{rank}(\mathbf{E}_{j,1}) < n_i$, $\text{rank}(\mathbf{E}_{j,2}) < q_i \leq n_i$ by construction. Therefore $\mathbf{S}_j \mathbf{A}_{jj} = \begin{bmatrix} S_{j,1}A^{art}_{jj} & \mathbf{0} \\ \star & S_{j,2}A_{jj} \end{bmatrix}$, where $\star$ represents the additional term. Hence, the pair $(\mathbf{C}_j, \mathbf{S}_j \mathbf{A}_{jj})$ is detectable:

$$\begin{aligned} \text{rank} \begin{bmatrix} s\mathbf{I} - \mathbf{S}_j \mathbf{A}_{jj} \\ \mathbf{C}_j \end{bmatrix} &= \text{rank} \begin{bmatrix} s\mathbf{I} - S_{j,1}A^{art}_{jj} & \mathbf{0} \\ \star & s\mathbf{I} - S_{j,2}A_{jj} \\ \mathbf{I} & \mathbf{0} \\ C_j & C_j \end{bmatrix} \\ &= n_j + \text{rank} \begin{bmatrix} s\mathbf{I} - S_{j,2}A_{jj} \\ C_j \end{bmatrix} \end{aligned}$$

which, given detectability of the pair $(C_j, S_j A_{jj})$ by hypothesis, is equal to $2n_j, \forall s \in \mathbb{C}^+$, with $S_{j,2} = S_j$.

The second part of the proposition holds, as $\text{rank}(\mathbf{C}_j) = n_j + \text{rank}(C_j) > \text{rank}(\mathbf{E}_{j,1}) + \text{rank}(\mathbf{E}_{j,2})$, given $\text{rank}(\mathbf{E}_{j,1}) + \text{rank}(\mathbf{E}_{j,2}) = \text{rank}(\bar{E}_j)$ by construction and $\text{rank}(C_j) = \text{rank}(\bar{E}_j)$ by hypothesis. ■

Remark 6. In the case of DCmGs, as can be seen from the definition of the system matrices in Appendix A, $\text{rank}(\bar{E}_j) = \text{rank}(C_j) = n_j$. As such it is not possible to design a UIO capable of detecting attacks and it is necessary to introduce the augmented state described above. Moreover, a good choice for the artificial state would be $x^{art}_{[j]} \triangleq x^{avg}_{[j]}$ (i.e. the state $x_{[i]}$ in (1) obtained by setting $u_{[i]} = u^{avg}_{[i]} \triangleq [V^{avg}_{ti}, \Delta V_i]$). ◁

In the sequel, we will consider that the observers in $\mathcal{O}^{UIO}_{i,j}$ are defined as in (8)-(9), with system matrices taken from augmented dynamics in (11). Furthermore, to stress the use of the augmented measurements $\mathbf{y}^c_{[j,i]}$, bold symbols $\mathbf{z}_{[j,i]}$, $\hat{\mathbf{x}}_{[j,i]}$ and $\hat{\mathbf{y}}_{[j,i]}$ will be used to denote the observer's state and the augmented state and output estimates.

Lemma 3. *If matrix* $A \in \mathbb{R}^{n \times n}$ *is Hurwitz stable, there exists a positive scalar* $\lambda > 0$, *and a matrix* $\Lambda \geq \mathbf{I}$ *such that:*

$$\left|e^{At}\right| \leq e^{-\lambda t}\Lambda \tag{15}$$

holds for all $t \geq 0$. □

Proof. The proof can be found in Appendix C. ■

Given the appropriate design of filter matrices (9), the estimation error $\epsilon_{[j,i]} \triangleq \mathbf{x}_{[j]} - \hat{\mathbf{x}}_{[j,i]}$ is stable, and it is therefore possible to design a time-varying threshold $\bar{r}_{[j,i]}$ capable of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

bounding the UIO's residual error defined as output estimation error $r_{[j,i]} \triangleq \mathbf{y}^c_{[j,i]} - \hat{\mathbf{y}}_{[j,i]}$:

$$\bar{r}_{[j,i]}(t) \triangleq \mathbf{C}_j e^{-\sigma_j t}\Sigma_j \left[\bar{\epsilon}_{[j,i]}(0) + |H_j|\,\bar{\boldsymbol{\rho}}_{[j]}\right] + |Z_j|\,\bar{\boldsymbol{\rho}}_{[j]} + \\ + \mathbf{C}_j \int_0^t e^{-\sigma_j (t-\tau)}\Sigma_j \left[|S_j|\,\bar{\tilde{\mathbf{w}}}_{[j]} + \left|\widehat{K}_j\right|\bar{\boldsymbol{\rho}}_{[j]}\right] d\tau. \tag{16}$$

where $Z_j \triangleq (\mathbf{I} - \mathbf{C}_j H_j)$ and $\mathbf{C}_j \geq \mathbf{0}$ is supposed without loss of generality; as F_j Hurwitz stable, scalar $\sigma_j > 0$ and matrix $\Sigma_j \geq \mathbf{I}$ can be found as in Lemma 3. The following proposition guarantees that $\bar{r}_{[j,i]}$ in (16) is indeed an upper bound to the corresponding residual.

Proposition 1. *In the absence of an attack, given F_j Hurwitz stable by design and Assumption 2, $\bar{r}_{[j,i]}(t)$ in* (16) *is such that the inequality:*

$$|r_{[j,i]}(t)| \leq \bar{r}_{[j,i]}(t) \tag{17}$$

holds for all $t < T_a^{j,i}, \forall j \in \mathcal{N}_i$. □

Proof. Given the definition of the UIO matrices, it is possible to derive the dynamics of the estimation error $\epsilon_{[j,i]}(t)$ as:

$$\begin{aligned}\dot{\epsilon}_{[j,i]}(t) &= \dot{\mathbf{x}}_{[j]}(t) - \dot{\hat{\mathbf{x}}}_{[j,i]}(t) \\ &= F_j \epsilon_{[j,i]}(t) + S_j \tilde{\mathbf{w}}_{[j]}(t) - H_j \dot{\boldsymbol{\rho}}_{[j]}(t) - \widetilde{K}_j \boldsymbol{\rho}_{[j]}(t),\end{aligned} \tag{18}$$

the solution of which, exploiting integration by parts, is:

$$\epsilon_{[j,i]}(t) = e^{F_j t}\left[\epsilon_{[j,i]}(0) + H_j \boldsymbol{\rho}_{[j]}(0)\right] - H_j \boldsymbol{\rho}_{[j]}(t) + \\ + \int_0^t e^{F_j (t-\tau)}\left[S_j \tilde{\mathbf{w}}_{[j]}(\tau) - \widehat{K}_j \boldsymbol{\rho}_{[j]}(\tau)\right] d\tau. \tag{19}$$

Given that $r_{[j,i]}(t) = \mathbf{C}_j \epsilon_{[j,i]}(t) + \boldsymbol{\rho}_{[j]}(t)$ in nominal conditions, the solution of residual $r_{[j,i]}(t)$ is:

$$r_{[j,i]}(t) = \mathbf{C}_j e^{F_j t}\left[\epsilon_{[j,i]}(0) + H_j \boldsymbol{\rho}_{[j]}(0)\right] + Z_j \boldsymbol{\rho}_{[j]}(t) + \\ + \mathbf{C}_j \int_0^t e^{F_j (t-\tau)}\left[S_j \tilde{\mathbf{w}}_{[j]}(\tau) - \widehat{K}_j \boldsymbol{\rho}_{[j]}(\tau)\right] d\tau. \tag{20}$$

By use of triangle inequality, bounds in Assumption 2, and Lemma 3, it is possible to bound the estimation error with:

$$\bar{\epsilon}_{[j,i]}(t) \triangleq e^{-\sigma_j t}\Sigma_j \left[\bar{\epsilon}_{[j,i]}(0) + |H_j|\,\bar{\boldsymbol{\rho}}_{[j]}\right] + |H_j|\,\bar{\boldsymbol{\rho}}_{[j]} + \\ + \int_0^t e^{-\sigma_j (t-\tau)}\Sigma_j \left[|S_j|\,\bar{\tilde{\mathbf{w}}}_{[j]} + \left|\widehat{K}_j\right|\bar{\boldsymbol{\rho}}_{[j]}\right] d\tau, \tag{21}$$

which will converge to a constant for $t \to \infty$, as F_j is Hurtwitz stable. Similarly, the threshold $\bar{r}_{[j,i]}(t)$ in (16) is such that inequality (17) is guaranteed to hold when the communication link between DGU j and i is not under attack, i.e. $t < T_a^{j,i}$, thus proving the Proposition. ■

Whenever inequality (5a) is violated, the monitoring module $\mathcal{D}_i$ detects the presence of an attack on the communication link between $\mathcal{S}_j$ and $\mathcal{S}_i$, thus isolating it. In order to perform detection using the UIO-based layer, $\mathcal{D}_i$ requires information *offline* to design the bank of UIOs, and information *online* to perform the updates to the estimate and to compute the residual. These requirements are found in Table I.

B. Detectability Properties of $\mathcal{O}^{UIO}_{j,i}$

We define a *detectable attack* as an attack function that is guaranteed to trigger the monitor $\mathcal{D}_i$ by some finite time $T_d \geq \check{T}_a^i$. In this and the following subsections, we will analyze the properties of the UIO-based detection module of $\mathcal{D}_i$ while under attack, i.e. for $t \geq \check{T}_a^i$. Note that, given that each UIO evaluates the security of a single communication line, we consider a single attack starting at $T_a^{j,i}$. Hence, let us define $T_a \triangleq T_a^{j,i}$ for clarity of exposition.

Once an attack is active on a communication link, i.e. for $t \geq T_a$, the residual error of $\mathcal{O}^{UIO}_{j,i}$ can be expressed as:

$$r_{[j,i]}(t) = r^h_{[j,i]}(t) + r^a_{[j,i]}(t) \tag{22}$$

where $r^h_{[j,i]}(t)$ is the same as the residual in nominal conditions defined in (20), and $r^a_{[j,i]}(t) \triangleq \mathbf{C}_j \epsilon^a_{[j,i]}(t) + \phi_{j,i}(t)$, with:

$$\epsilon^a_{[j,i]}(t) \triangleq -H_j \boldsymbol{\phi}_{j,i}(t) + e^{F_j (t-T_a)} H_j \boldsymbol{\phi}_{j,i}(T_a) + \\ - \int_{T_a}^t e^{F_j (t-\tau)} \widehat{K}_j \boldsymbol{\phi}_{j,i}(\tau) d\tau. \tag{23}$$

The class of attacks that are guaranteed to be detected can therefore be expressed in the following proposition:

Proposition 2. *If attack function $\boldsymbol{\phi}_{j,i}(t)$ is such that at any time $t \geq T_a$*

$$\left|r^a_{[j,i]}(t)\right| > 2\bar{r}_{[j,i]}(t) \tag{24}$$

holds for any component, then detector $\mathcal{D}_i$ operating in accordance with Algorithm 1 will detect the attack, thanks to the UIO observer $\mathcal{O}^{UIO}_{j,i}$. □

Proof. By using the triangle inequality, Proposition 1, and exploiting the decomposition in (22), one has

$$|r_{[j,i]}(t)| \geq \left|r^a_{[j,i]}(t)\right| - \left|r^h_{[j,i]}(t)\right| \geq \left|r^a_{[j,i]}(t)\right| - \bar{r}_{[j,i]}(t) \tag{25}$$

where we used the fact that $\left|r^h_{[j,i]}(t)\right|$ in (20) is upper bounded by $\bar{r}_{[j,i]}(t)$. For guaranteeing detection through violation of (5a), it is sufficient that the attack $\boldsymbol{\phi}_{j,i}(t)$ is such that

$$\left|r^a_{[j,i]}(t)\right| - \bar{r}_{[j,i]}(t) > \bar{r}_{[j,i]}(t) \tag{26}$$

is satisfied for some time $t > \check{T}_a^i$. As (24) is a sufficient condition for (26), this concludes the proof. ■

C. Classes of Attacks Stealthy to $\mathcal{O}^{UIO}_{j,i}$

Having evaluated the class of attacks which are guaranteed to be detected by $\mathcal{O}^{UIO}_i$ in $\mathcal{D}_i$, we now analyze the UIO-based module's *weakness*, i.e. those attacks which are *stealthy* to it.

Definition 1 (Stealthy Attacks)**.** *An attack is stealthy to $\mathcal{D}_i$ if it is guaranteed not to be detected at any time $t \geq \check{T}_a^i$.* ◇

It is worth recalling that, as described in Remark 2, the attack only influences the output communicated between controllers, while not attacking any subsystem's dynamics directly. Hence, the stealthiness properties differ with respect to those available in literature [10], [24]. Again we exploit the decomposition of $r_{[j,i]}(t)$ into healthy and attacked components to analyze stealthiness. In order to give a complete overview of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

the stealthy attacks for this module, we will separately treat three classes of attacks defined in [10]: false data injection attack; replay attack; covert attack.

a) False Data Injection Attacks: this class of attacks does not require any *disclosure* capabilities (i.e. the malicious agent does not need to eavesdrop the information sent through the communication link). By injecting an attack of this type, it is possible for the attacker to alter the equilibrium of the network as a whole. The influence of this type of attack on the residual $r_{[j,i]}(t)$ can be characterized as in (22).

Proposition 3. *If attacks $\boldsymbol{\phi}_{j,i}(t)$ are such that for all $t \geq T_a$:*

$$\left| r^a_{[j,i]}(t) \right| = 0, \tag{27}$$

then they will be stealthy to the UIO-based module in $\mathcal{D}_i$. □

Proof. Given that $\left| r^h_{[j,i]}(t) \right|$ is bounded by $\bar{r}_{[j,i]}(t)$ by construction, and exploiting the triangle inequality, it holds that:

$$\left| r_{[j,i]}(t) \right| = \left| r^h_{[j,i]}(t) + r^a_{[j,i]}(t) \right| \leq \bar{r}_{[j,i]}(t) + \left| r^a_{[j,i]}(t) \right|. \tag{28}$$

Given that, for the attack to be undetected, inequality (5a) must always hold, it is sufficient that $\boldsymbol{\phi}_{j,i}(t)$ is designed to satisfy $\left| r^a_{[j,i]}(t) \right| = 0, \forall t \geq \check{T}^i_a$ for it to be stealthy. ■

Remark 7. Recalling that $r^a_{[j,i]} = \mathbf{C}_j \epsilon^a_{[j,i]} + \phi_{j,i}$, it is sufficient for attacks to be such that

$$\boldsymbol{\phi}_{j,i}(T_a) = \mathbf{0}, \quad \boldsymbol{\phi}_{j,i}(t) \in ker \left(\begin{bmatrix} \widehat{K}_j \\ Z_j \end{bmatrix} \right), t > T_a \tag{29}$$

for condition (27) to be satisfied for all $t \geq T_a$. ◁

b) Replay Attacks: With an attacker capable of violating the integrity of the communication network (and thus to eavesdrop on the transmitted measurements) from some time $t = T_0$, a replay attack requires no knowledge of the system's model. Instead, it modifies the transmitted information by *replaying* stored old data, substituting it for the current data[2]. Hence communicated information (12) will be

$$\mathbf{y}^c_{[j,i]}(t) = \mathbf{y}_{[j]}(t - nT).$$

It has been shown that replay attacks may be undetectable to attack monitoring schemes [25], as the replayed data has both the same statistical properties of the non-attacked data, and it evolves following correct dynamics.

Note that, although a replay attack does not require any knowledge of the subsystem's dynamics, it is possible for the attacker to disguise any changes to the operating conditions of a unit from its neighbors, thus altering or disrupting the consensus equilibrium.

Specifically, in our scenario, the following condition can be given:

Lemma 4. *If a replay attack is such that:*

$$\Sigma_j \left| \epsilon^r_{[j,i]}(T_a) + H_j \boldsymbol{\rho}_{[j]}(T_a - T) \right| \leq \bar{\epsilon}_{[j,i]}(T_a) - |H_j| \, \bar{\boldsymbol{\rho}}_{[j]}, \tag{30}$$

then detection test (5a) *will hold for all $t \in [T_a, T_a + T)$, where $\epsilon^r_{[j,i]}(T_a) \triangleq \mathbf{x}_{[j]}(T_a - T) - \hat{\mathbf{x}}_{[j,i]}(T_a)$.* □

Proof. Given T_a and T, for time $t \in [T_a, T_a + T)$, the UIO estimation error residual takes the form:

$$\begin{aligned} r_{[j,i]}(t) &= \mathbf{y}^c_{[j,i]}(t) - \hat{\mathbf{y}}_{[j,i]}(t) \\ &= \mathbf{C}_j \epsilon^r_{[j,i]}(t) + \boldsymbol{\rho}_{[j]}(t - T). \end{aligned}$$

The dynamics of state estimation error under attack $\epsilon^r_{[j,i]}(t)$ can be derived from equations (1) and (8):

$$\begin{aligned} \dot{\epsilon}^r_{[j,i]}(t) =& F_j \epsilon^r_{[j,i]}(t) + S_j \tilde{\mathbf{w}}_{[j]}(t - T) + \\ &- \widetilde{K}_j \boldsymbol{\rho}_{[j]}(t - T) - H_j \dot{\boldsymbol{\rho}}_{[j]}(t - T) \end{aligned} \tag{31}$$

the solution of which is:

$$\begin{aligned} \epsilon^r_{[j,i]}(t) = e^{F_j(t - T_a)} &\left(\epsilon^r_{[j,i]}(T_a) + H_j \boldsymbol{\rho}_{[j]}(T_a - T) \right) + \\ -H_j \boldsymbol{\rho}_{[j]}(t - T) + \int_{T_a}^{t} e^{F_j(t-\tau)} &\left[S_j \tilde{\mathbf{w}}_{[j]}(\tau - T) + \right. \\ &\left. - \widehat{K}_j \boldsymbol{\rho}_{[j]}(\tau - T) \right] d\tau. \end{aligned} \tag{32}$$

Estimation error bound $\bar{\epsilon}_{[j,i]}(t)$ defined in (21) for time $t > T_a$ can be rewritten as:

$$\begin{aligned} \bar{\epsilon}_{[j,i]}(t) = e^{-\sigma_j(t - T_a)} &\left[\bar{\epsilon}_{[j,i]}(T_a) - |H_j| \, \bar{\boldsymbol{\rho}}_{[j]} \right] + |H_j| \, \bar{\boldsymbol{\rho}}_{[j]} + \\ &+ \int_{T_a}^{t} e^{-\sigma_j(t-\tau)} \Sigma_j \left[|S_j| \, \bar{\tilde{\mathbf{w}}}_{[j]} + \left| \widehat{K}_j \right| \bar{\boldsymbol{\rho}}_{[j]} \right] d\tau. \end{aligned} \tag{33}$$

In order to guarantee that $|r_{[j,i]}(t)| \leq \bar{r}_{[j,i]}(t)$, $\forall t \in [T_a, T_a + T)$, implying stealthiness, it is sufficient that $|\epsilon^r_{[j,i]}(t)| \leq \bar{\epsilon}_{[j,i]}(t)$. By comparison, all terms except

$$e^{F_j(t - T_a)} \left(\epsilon^r_{[j,i]}(T_a) + H_j \boldsymbol{\rho}_{[j]}(T_a - T) \right)$$

are guaranteed to be bounded by their corresponding terms in (33), given the definition of the disturbance bounds in (2). Thus, as the following inequality holds:

$$\begin{aligned} \left| e^{F_j(t - T_a)} \left(\epsilon^r_{[j,i]}(T_a) + H_j \boldsymbol{\rho}_{[j]}(T_a - T) \right) \right| \leq \\ \leq e^{-\sigma_j(t - T_a)} \Sigma_j \left| \epsilon^r_{[j,i]}(T_a) + H_j \boldsymbol{\rho}_{[j]}(T_a - T) \right|, \end{aligned}$$

it is sufficient for condition (30) to hold for stealthiness to be achieved, which proves the Lemma holds for $t \in (T_a, T_a + T)$.

To prove sufficiency of (30) for $|\epsilon^r_{[j,i]}(T_a)| \leq \bar{\epsilon}_{[j,i]}(T_a)$, we use the property of $\Sigma_j \geq \mathbf{I}$ and the inverse triangle inequality:

$$\begin{aligned} \Sigma_j |\epsilon^r_{[j,i]}(T_a) + H_j \boldsymbol{\rho}_{[j]}(T_a - T)| \geq \\ \geq |\epsilon^r_{[j,i]}(T_a) + H_j \boldsymbol{\rho}_{[j]}(T_a - T)| \geq \\ \geq \left| \epsilon^r_{[j,i]}(T_a) \right| - |H_j| \bar{\boldsymbol{\rho}}_{[j]}. \end{aligned} \tag{34}$$

Hence, if (30) is satisfied, the following holds:

$$\left| \epsilon^r_{[j,i]}(T_a) \right| - |H_j| \bar{\boldsymbol{\rho}}_{[j]} \leq \bar{\epsilon}_{[j,i]}(T_a) - |H_j| \, \bar{\boldsymbol{\rho}}_{[j]}, \tag{35}$$

and therefore detection will not occur at time $t = T_a$.

Note finally that, given definition of $\bar{\epsilon}_{[j,i]}(t)$ in (21), the right hand side of (30) is guaranteed to be greater than zero. Hence, (30) is well defined. This completes the proof. ■

[2]The analysis of the stealthiness of replay attacks in $\mathcal{O}^{UIO}_{j,i}$ was presented preliminarily in [44].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

Proposition 4. *If a replay attack is such that:*

$$\Sigma_j \left|\epsilon^r_{[j,i]}(T_a) + M_{j,i} + H_j\boldsymbol{\rho}_{[j]}(T_a - T)\right| \leq \bar{\epsilon}_{[j,i]}(T_a) - \Delta\bar{\epsilon}_{[j,i]}(T_a) - |H_j|\bar{\boldsymbol{\rho}}_{[j]} \tag{36}$$

holds, with $M_{j,i}, \Delta\bar{\epsilon}_{[j,i]}(T_a) \geq \mathbf{0}$ *appropriately defined vectors, then detection test* (5a) *will hold for all* $t \geq T_a$. □

Proof. The proof can be found in Appendix D. ■

Remark 8. Conditions in Lemma 4 and Proposition 4 depend on quantities unknown to the attacker, so it is not guaranteed that the attack will be able to satisfy them. However, as long as the attacker chooses T_a and T appropriately (i.e. such that $\mathbf{y}_{[j]}(T_a) \approx \mathbf{y}_{[j]}(T_a - T)$), it is *likely* (although not guaranteed) that (5a) will hold for all $t \geq T_a$. ◁

c) Covert Attacks: To perform a *covert attack*, the malicious agent must not only be able to disrupt the communication network, and be able to eavesdrop the information being transmitted, but must also have knowledge of the dynamics of $\mathcal{S}_j$. It is therefore capable of *simulating* the behavior of the subsystem and feeding this information to the control and monitoring architecture of $\mathcal{S}_i$. Specifically, a covert attack can be modeled as follows:

$$\boldsymbol{\phi}_{j,i}(t) = -\mathbf{y}_{[j]}(t) + \mathbf{y}^a_{[j]}(t), \tag{37}$$

where $\mathbf{y}^a_{[j]}(t)$ is the output of a simulated system with the following dynamics and initial condition:

$$\begin{aligned} \dot{\mathbf{x}}^a_{[j]}(t) &= \mathbf{A}_{jj}\mathbf{x}^a_{[j]}(t) + \mathbf{E}_j\mathbf{d}^a_{[j]}(t) \\ \mathbf{y}^a_{[j]}(t) &= \mathbf{C}_j\mathbf{x}^a_{[j]} \\ \mathbf{x}^a_{[j]}(T_a) &= \mathbf{C}_j^\dagger\mathbf{y}_{[j]}(T_a) \end{aligned}, \tag{38}$$

where $\mathbf{d}^a_{[j]}(t)$ is freely chosen by the attacker to substitute $\mathbf{d}_{[i]}$ in (11). Under this scenario, $\mathbf{y}^c_{[j,i]}(t) = \mathbf{y}^a_{[j]}(t)$.

Remark 9. Note that, differently to the *covert attack* described in [10], [26], we do not consider the case in which the attacker may alter the control input signals of $\mathcal{S}_j$, but only the information transmitted to $\mathcal{S}_i$, consistently with Remark 2. While limiting the scope of the attacker, through modification of the unknown input vector $\mathbf{d}^a_{[j]}(t) \neq \mathbf{d}_{[j]}(t)$, it is possible for it to change the operating condition of $\mathcal{S}_j$ as seen by $\mathcal{S}_i$, thus modifying the behavior of the LSS as a whole. ◁

Remark 10. For $\text{rank}(\mathbf{C}_j) < 2n_j$, the attacker may introduce initial state error $\Delta\mathbf{x}^a_{[j]}(T_a) \triangleq \mathbf{x}_{[j]}(T_a) - \mathbf{x}^a_{[j]}(T_a) \in \ker(\mathbf{C}_j)$. This, given observability of $(\mathbf{C}_j, \mathbf{A}_{jj})$, will be nonetheless observable from $\mathbf{y}^a_{[j]}(t), t > T_a$, and it may thus be possible for $\mathcal{D}_i$ to detect the attack. In Proposition 5 we have considered the worst case scenario in which $\Delta\mathbf{x}^a_{[j]}(T_a) = \mathbf{0}$. ◁

Proposition 5. *If an attack as in* (37) *is carried out, in which* $\mathbf{x}^a_{[j]}(t)$ *is the state of LTI system* (38)*, and if* $\boldsymbol{\rho}_{[j]}$ *is such that*

$$\mathbf{C}_j\Sigma_j|(\mathbf{C}_j^\dagger - H_j)\boldsymbol{\rho}_{[j]}(T_a)| \leq |Z_j|\bar{\boldsymbol{\rho}}_{[j]}, \tag{39}$$

inequality (5a) *will hold for all* $t \geq T_a$*, and the attack will be stealthy.* □

Proof. Start by noticing that for time $t = T_a$, the residual is:

$$r_{[j,i]}(T_a) = \mathbf{C}_j\mathbf{x}^a_{[j]}(T_a) - \hat{\mathbf{y}}_{[j,i]}(T_a) = \mathbf{y}_{[j]}(T_a) - \hat{\mathbf{y}}_{[j,i]}(T_a), \tag{40}$$

and therefore condition (5a) will hold, given Proposition 1. For $t > T_a$, $\varepsilon^a_{[j,i]}(t) \triangleq \mathbf{x}^a_{[j]}(t) - \hat{\mathbf{x}}_{[j,i]}(t)$. The dynamics of the residual error can therefore be written as:

$$\dot{\varepsilon}^a_{[j,i]}(t) = F_j\varepsilon^a_{[j,i]}(t) + S_j\mathbf{E}_j\mathbf{d}^a_{[j]}(t) = F_j\varepsilon^a_{[j,i]}(t),$$

as $S_j\mathbf{E}_j = \mathbf{0}$ by design (9a). Hence:

$$\begin{aligned} r_{[j,i]}(t) &= \mathbf{C}_j e^{F_j(t-T_a)}\varepsilon^a_{[j,i]}(T_a) = \\ &= \mathbf{C}_j e^{F_j t}\left[\epsilon_{[j,i]}(0) + H_j\boldsymbol{\rho}_{[j]}(0)\right] + \\ &+ \mathbf{C}_j e^{F_j(t-T_a)}\left(\mathbf{C}_j^\dagger - H_j\right)\boldsymbol{\rho}_{[j]}(T_a) + \\ &+ \mathbf{C}_j\int_0^{T_a} e^{F_j(t-\tau)}\left[S_j\tilde{\mathbf{w}}_{[j]}(\tau) - \widehat{K}_j\boldsymbol{\rho}_{[j]}(\tau)\right]d\tau. \end{aligned} \tag{41}$$

Comparing (41) to the definition of the residual in healthy conditions (20), we see that the only term not guaranteed to be bounded by the corresponding terms in (16) is $\mathbf{C}_j e^{F_j(t-T_a)}\left(\mathbf{C}_j^\dagger - H_j\right)\boldsymbol{\rho}_{[j]}(T_a)$. Hence, to guarantee that (5a) holds, we must demonstrate that

$$\begin{aligned} \left|\mathbf{C}_j e^{F_j(t-T_a)}\left(\mathbf{C}_j^\dagger - H_j\right)\boldsymbol{\rho}_{[j]}(T_a)\right| &\leq |Z_j|\,\bar{\boldsymbol{\rho}}_{[j]} + \\ &+ \mathbf{C}_j\int_{T_a}^{t} e^{-\sigma_j(t-\tau)}\Sigma_j\left[|S_j|\,\bar{\bar{\mathbf{w}}}_{[j]} + \left|\widehat{K}_j\right|\bar{\boldsymbol{\rho}}_{[j]}\right]. \end{aligned} \tag{42}$$

Recalling that $|e^{F_j t}| \leq \Sigma_j e^{\sigma_j t}$, it is sufficient for condition (39) to hold for (42) to be satisfied, and therefore detection condition (5a) will hold for all $t \geq T_a$. ■

In this Section we have presented $\mathcal{O}^{UIO}_{j,i}$, as well as its detectability properties. It is worth noting that this detection module does not rely on the physical interconnections between subsystems, but only on the communicated values received from its neighbors $\mathcal{S}_j$.

V. Distributed estimation of local states

A. Design of the detection module

The second module of the attack detection monitor $\mathcal{D}_i$ is based on a distributed Luenberger observer $\mathcal{O}^{Luen}_i$. The following assumption is made in this section, motivated by the application to microgrids (see Section II):

Assumption 5. *Matrix* C_i *is invertible for all* $\mathcal{S}_i$. ◁

We will give some indications as how this assumption could be removed in Remark 12. Note that, from Assumption 5, it follows that $\mathbf{C}_j$ is also non-singular. The dynamics of $\mathcal{O}^{Luen}_i$ can therefore be formulated as:

$$\begin{aligned} \dot{\hat{x}}_{[i]} &= A_{ii}\hat{x}_{[i]} + \hat{\xi}_{[i]} + B_i u_{[i]} + M_i d_{[i]} - L_i\left(y_{[i]} - \hat{y}_{[i]}\right), \\ \hat{y}_{[i]} &= C_i\hat{x}_{[i]} \end{aligned} \tag{43}$$

where L_i is designed such that $A_{Li} = (A_{ii} + L_iC_i)$ is Hurwitz stable, guaranteeing estimation error stability, and the effect of the physical interconnection with neighbors in $\mathcal{N}_i$

$$\hat{\xi}_{[i]} \triangleq \sum_{j\in\mathcal{N}_i} A_{ij}\hat{x}_{[j,i]} = \sum_{j\in\mathcal{N}_i} A_{ij}\Gamma\mathbf{C}_j^{-1}\mathbf{y}^c_{[j,i]},$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

where $\Gamma \mathbf{C}_j^{-1}\mathbf{y}^c_{[j,i]}$ is used as an estimate of $x_{[j]}$, with $\Gamma \triangleq \begin{bmatrix} \mathbf{I} & \mathbf{I} \end{bmatrix}$, recalling (10).

To verify whether hypothesis $\mathcal{H}_i^0(t)$ in Problem 1 is valid or not, $\mathcal{D}_i$ computes the residual error

$$r_{[i]}(t) \triangleq y_{[i]}(t) - \hat{y}_{[i]}(t), \tag{44}$$

and compares it with an appropriately defined time-varying threshold $\bar{r}_{[i]}(t)$, given by:

$$\bar{r}_{[i]}(t) \triangleq C_i e^{-\lambda_i t}\Lambda_i \bar{\epsilon}_{[i]}(0) + C_i \int_0^t e^{-\lambda_i(t-\tau)}\Lambda_i \bar{\eta}_{[i]} d\tau + \bar{\rho}_{[i]} \tag{45}$$

where $\lambda_i > 0$ and $\Lambda_i \geq \mathbf{I}$ are such that $|e^{A_{Li}t}| \leq e^{-\lambda_i t}\Lambda_i$ holds, thanks to Lemma 3; $\bar{\epsilon}_{[i]}(0)$ is an appropriately defined initial condition of the bound on the estimation error $\epsilon_{[i]}(t) \triangleq x_{[i]}(t) - \hat{x}_{[i]}(t)$; and $\bar{\eta}_{[i]} \triangleq \bar{w}_{[i]} + |L_i|\bar{\rho}_{[i]} + \sum_{j\in\mathcal{N}_i} |A_{ij}|\Gamma\mathbf{C}_j^{-1}\bar{\boldsymbol{\rho}}_{[j]}$. The following proposition holds:

Proposition 6. *Given Assumption 2 and that A_{Li} is Hurwitz stable by design, the inequality:*

$$|r_{[i]}(t)| \leq \bar{r}_{[i]}(t) \tag{46}$$

is guaranteed to be satisfied for all $t < \check{T}_a^i$, for residual $r_{[i]}$ in (44) *and threshold $\bar{r}_{[i]}$ computed by $\mathcal{D}_i$ as in* (45). □

Proof. The residual error can be rewritten as $r_{[i]} = C_i\epsilon_{[i]} + \rho_{[i]}$. The dynamics of $\epsilon_{[i]}(t)$ can be derived from (1) and (43):

$$\dot{\epsilon}_{[i]}(t) = A_{Li}\epsilon_{[i]}(t) + \eta_{[i]}(t) \tag{47}$$

where $\eta_{[i]} = -\sum_{j\in\mathcal{N}_i} A_{ij}\Gamma\mathbf{C}_j^{-1}\boldsymbol{\rho}_{[j]} + w_{[i]} - L_i\rho_{[i]}$. The following explicit solution can be found:

$$\epsilon_{[i]}(t) = e^{A_{Li}t}\epsilon_{[i]}(0) + \int_0^t e^{A_{Li}(t-\tau)}\eta_{[i]}(\tau)d\tau. \tag{48}$$

Since A_{Li} is Hurwitz stable by design of L_i for all $i \in \mathcal{N}$, estimation error $\epsilon_{[i]}(t)$ is BIBO stable, and, given Assumption 2, it can be bounded by a time-varying quantity $\bar{\epsilon}_{[i]}(t)$. Using the triangle inequality and bounds defined in (2) as well as Lemma 3, a bound on the estimation error can be computed:

$$\bar{\epsilon}_{[i]}(t) \triangleq e^{-\lambda_i t}\Lambda_i\bar{\epsilon}_{[i]}(0) + \int_0^t e^{-\lambda_i(t-\tau)}\Lambda_i\bar{\eta}_{[i]}d\tau, \tag{49}$$

where $\lambda_i > 0$ and Λ_i are found following Lemma 3, and $\bar{\epsilon}_{[i]}(0)$ is appropriately defined. The threshold in (45) on the residual can similarly be computed by using the triangle inequality. ■

The information required by $\mathcal{D}_i$ to compute the estimate $\hat{x}_{[i]}(t)$ and threshold $\bar{r}_{[i]}(t)$ is provided in Table I.

B. Detectability Properties of $\mathcal{O}_i^{Luen}$

In this and the following subsections, we will analyze the properties of the Luenberger-observer-based detection module of $\mathcal{D}_i$ while under attack, i.e. for $t \geq \check{T}_a^i$. Once an attack is active on a communication link, it will affect both the computation of the networked control $u_{[i]}(t)$ and of the variable $\hat{\xi}_{[i]}(t)$ in (43), which will become:

$$\begin{aligned}\hat{\xi}_{[i]}(t) = &\sum_{j\in\mathcal{N}_i} A_{ij}\Gamma\mathbf{C}_j^{-1}\left(\mathbf{C}_j\mathbf{x}_{[j]}(t) + \boldsymbol{\rho}_{[j]}(t)\right) \\ &+ \sum_{j\in\widehat{\mathcal{N}}_i(t)} A_{ij}\Gamma\mathbf{C}_j^{-1}\boldsymbol{\phi}_{j,i}(t)\,,\ \forall t \geq \check{T}_a^i\end{aligned} \tag{50}$$

where $\widehat{\mathcal{N}}_i(t) \triangleq \left\{j \in \mathcal{N}_i : t \geq T_a^{j,i}\right\} \subseteq \mathcal{N}_i$ is the set of neighbors whose transmissions to $\mathcal{S}_i$ have been attacked at time t. As the attack is additive with respect to the dynamics (43), it is possible to write the residual as:

$$r_{[i]}(t) = r^h_{[i]}(t) + r^a_{[i]}(t), \tag{51}$$

where:

$$r^h_{[i]}(t) \triangleq C_i e^{A_{Li}t}\epsilon_{[i]}(0) + C_i\int_0^t e^{A_{Li}(t-\tau)}\eta_{[i]}(\tau)d\tau + \rho_{[i]}(t) \tag{52}$$

is the *healthy* part of the residual, and is independent of $\check{T}_a^i$. Hence $\left|r^h_{[i]}(t)\right| \leq \bar{r}_{[i]}(t)$ will hold for all $t \geq 0$. Moreover,

$$r^a_{[i]}(t) \triangleq C_i\int_{\check{T}_a^i}^t e^{A_{Li}(t-\tau)}\sum_{j\in\widehat{\mathcal{N}}_i(t)} A_{ij}\Gamma\mathbf{C}_j^{-1}\boldsymbol{\phi}_{j,i}(\tau)d\tau, \tag{53}$$

for all $t \geq \check{T}_a^i$, is the part of the residual affected by the attack.

Proposition 7. *If attack function $\boldsymbol{\phi}_{j,i}(t) \in \mathbb{R}^{n_j+p_j}$ is such that at any time $t > \check{T}_a^i$*

$$\left|r^a_{[i]}(t)\right| > 2\bar{r}_{[i]}(t) \tag{54}$$

holds for any of its components, then detector $\mathcal{D}_i$ operating in accordance with Algorithm 1 will detect the attack at some finite time $T_d > \check{T}_a^i$ thanks to $\mathcal{O}_i^{Luen}$. □

Proof. The proof follows that of Proposition 2. ■

Having evaluated the class of attacks which are guaranteed to be detected by $\mathcal{O}_i^{Luen}$ in $\mathcal{D}_i$, we now analyze the Luenberger-observer-based module's *weakness*, i.e. the class of attacks which are *stealthy* to it.

We again exploit the decomposition of the residual $r_{[i]}(t)$ into healthy and attacked components to analyze stealthiness.

Proposition 8. *If attacks $\boldsymbol{\phi}_{j,i}(t)$ are such that for all $t \geq \check{T}_a^i$*

$$\left|r^a_{[i]}(t)\right| = 0, \tag{55}$$

holds, then they will be stealthy to the Luenberger-observer-based module in $\mathcal{D}_i$. □

Proof. The proof follows that of Proposition 3. ■

Remark 11. For Proposition 8 to hold for all $t \geq \check{T}_a^i$, $\boldsymbol{\Phi}_{[i]}(t) \triangleq \mathrm{col}(\boldsymbol{\phi}_{j,i}(t)), \forall j \in \widehat{\mathcal{N}}_i(t)$ must satisfy:

$$\begin{aligned}\boldsymbol{\Phi}_{[i]}(t) &\in \ker(\mathbb{A}_{ij}(t)), \\ \mathbb{A}_{ij}(t) &\triangleq \left[A_{ij_1}\Gamma\mathbf{C}_{j_1}^{-1}, \ldots, A_{ij_{\widehat{N}_i(t)}}\Gamma\mathbf{C}_{j_{\widehat{N}_i(t)}}^{-1}\right]\end{aligned} \tag{56}$$

where $\mathbb{A}_{ij}(t)$ *collects* physical coupling matrices of the neighbors whose communication has been attacked, and as such may be time-varying, with $\widehat{N}_i(t) \triangleq |\widehat{\mathcal{N}}_i(t)|$. This is revealing, as it shows the dependency of the detectability of $\mathcal{O}_i^{Luen}$ on the physical interconnections of $\mathcal{S}_i$ and its neighbors. Specifically, (56) implies that to design an attack stealthy to $\mathcal{O}_i^{Luen}$ an attacker could either leverage knowledge of the structure of the interconnection between subsystems, and therefore of a subset of the state $x_{[j]}$ that does not influence (1), or compensate its effect on the residual through multiple channels, depending on whether matrices $A_{ij}, j \in \mathcal{N}_i$ are singular. ◁

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

Remark 12. As previously mentioned, the analysis in this section was performed considering an invertible C_j. In the case when it is singular, it is possible to exploit the estimation of the neighbors' states $\hat{x}_{[j,i]} \triangleq \Gamma\hat{\mathbf{x}}_{[j,i]}$. Propositions 6-8 can then be showed to hold by making appropriate changes to $\bar{r}_{[i]}$, $\epsilon_{[i]}$, and $r^a_{[i]}$ in (45), (47), and (53), respectively. Specifically, while recalling that the estimation error of $\mathcal{O}^{UIO}_{j,i}$ can be decomposed in its *healthy* and *attacked* components, we change $\mathbf{C}_j^{-1}\bar{\rho}_{[j]}$ in definition of $\bar{\eta}_{[i]}$ to $\bar{\epsilon}_{[j,i]}$, $\mathbf{C}_j^{-1}\rho_{[j]}$ in $\eta_{[i]}$ to $\epsilon^h_{[j,i]}$, and $\mathbf{C}_j^{-1}\boldsymbol{\phi}_{j,i}$ to $\epsilon^a_{[j,i]}$. Hence, proofs of Propositions 6-8 follow.

A significant difference implied by this alteration of $\mathcal{O}^{Luen}_{j,i}$ is that the two modules in $\mathcal{D}_i$ are directly *coupled*, and that attack vector $\boldsymbol{\phi}_{[j,i]}$ no longer directly affects (53), but rather affects it through $\epsilon^a_{[j,i]}$.

In such a scenario, the Luenberger-observer-based detector will require from the UIO-based module, at all times $t \geq 0$, the state estimate $\hat{\mathbf{x}}_{[j,i]}(t)$ and the bound on its estimation error $\bar{\epsilon}_{[j,i]}(t)$. Thus, for Proposition 8 to hold it is sufficient for the attack vector to satisfy a condition similar to that in (29) with Z_j replaced by H_j. Furthermore, an attack would satisfy (55) also if it were such that $\epsilon^a_{[j,i]}(t)$ lie within $\ker(\mathbb{A}_{ij})$ with $\mathbf{C}^{-1}_{j_k} = \mathbf{I}$ for all $t \geq T_a$. Both these conditions rely on knowledge of parameters of $\mathcal{O}^{UIO}_{[j,i]}$. ◁

VI. Detectability Analysis of $\mathcal{D}_i$

We will show that the combined use of the two modules in $\mathcal{D}_i$ has advantages in terms of detectability. In fact, it is sufficient for either conditions in Proposition 2 or 7 to be satisfied for an attack to be guaranteed to be detected. In this section, we will therefore focus on two specific cases:

i. the class of bias injection attacks stealthy to $\mathcal{D}_i$;
ii. the detectability of a replay or covert attack.

For the first of the two cases, it is clear to see that for invertible C_j, to be stealthy to $\mathcal{D}_i$, it is sufficient that:

$$\boldsymbol{\Phi}_{[i]}(t) \in \ker\left(\begin{bmatrix} \mathbb{Z}_j(t) \\ \widehat{\mathbb{K}}_j(t) \end{bmatrix}\right) \cap \ker\left(\mathbb{A}_{ij}(t)\right), \tag{57}$$

while also satisfying $\boldsymbol{\phi}_{j,i}(T^a_{[j,i]}) = \mathbf{0}$, where $\widehat{\mathbb{K}}_j \triangleq \text{diag}(\widehat{K}_{j_1}, \dots, \widehat{K}_{j_{\widehat{N}_i(t)}})$ and $\mathbb{Z}_j \triangleq \text{diag}\left(Z_{j_1}, \dots, Z_{j_{\widehat{N}_i(t)}}\right)$. In fact, if (57) holds, then conditions for both Propositions 3 and 8 will hold. This, in turn, implies that neither of the modules of $\mathcal{D}_i$ will detect the attack, which will therefore be stealthy.

Remark 13. For the case of singular C_j, we refer to Remark 12 for derivation of equivalent conditions. ◁

In the second case, while replay and covert attacks are stealthy to the UIO-based module of $\mathcal{D}_i$, they may be detected by the Luenberger-based one. In order to simplify the analysis of this scenario, let us note that both replay and covert attacks can be interpreted as attack function:

$$\boldsymbol{\phi}_{j,i}(t) = -\mathbf{y}_{[j]}(t) + \mathbf{y}^a_{[j]}(t), \tag{58}$$

where $\mathbf{y}^a_{[j]}(t)$ is the output of the following LTI system:

$$\begin{aligned} \dot{\mathbf{x}}^a_{[j]}(t) &= \mathbf{A}_{jj}\mathbf{x}^a_{[j]}(t) + \mathbf{E}_j\mathbf{d}^a_{[j]}(t) + \tilde{\mathbf{w}}^a_{[j]}(t) \\ \mathbf{y}^a_{[j]}(t) &= \mathbf{C}_j\mathbf{x}^a_{[j]}(t) + \boldsymbol{\rho}^a_{[j]}(t), \end{aligned} \tag{59}$$

TABLE II: Values for interpretation of replay and covert attacks

	Replay Attacks	**Covert Attacks**
$\mathbf{x}^a_{[j]}(T_a)$	$\mathbf{x}_{[j]}(T_a - T)$	$\mathbf{C}_j^{\dagger}\mathbf{y}_{[j]}(T_a)$
$\mathbf{d}^a_{[j]}(t)$	$\mathbf{d}_{[j]}(t - nT)$	$\mathbf{d}^a_{[j]}(t)$
$\tilde{\mathbf{w}}^a_{[j]}(t)$	$\tilde{\mathbf{w}}_{[j]}(t - nT)$	$\mathbf{0}$
$\boldsymbol{\rho}^a_{[j]}(t)$	$\boldsymbol{\rho}_{[j]}(t - nT)$	$\mathbf{0}$

and the values of $\mathbf{d}^a_{[j]}(t), \tilde{\mathbf{w}}^a_{[j]}(t), \boldsymbol{\rho}^a_{[j]}(t)$ and initial condition $\mathbf{x}^a_{[j]}(T_a)$ can be defined as in Table II. Note furthermore that, for replay attacks, $\mathbf{x}^a_{[j]}(t)$ is periodic, and may be discontinuous in time for $t \in \mathcal{T} \triangleq \{t \in \mathbb{R} | t = T_a + nT, \forall n \in \mathbb{N}^0\}$, as $\mathbf{x}^a_{[j]}(T_a + nT) = \mathbf{x}^a_{[j]}(T_a), \forall n \in \mathbb{N}^0$. In this case we abuse notation by using (59), as it holds for $t \geq T_a, t \notin \mathcal{T}$.

For both covert and replay attacks it is possible to rewrite $\mathbf{d}^a_{[j]}(t) \triangleq \mathbf{d}_{[j]}(t) + \Delta\mathbf{d}_{[j]}(t)$, $\boldsymbol{\rho}^a_{[j]}(t) \triangleq \boldsymbol{\rho}_{[j]}(t) + \Delta\boldsymbol{\rho}_{[j]}(t)$, and $\tilde{\mathbf{w}}^a_{[j]}(t) \triangleq \tilde{\mathbf{w}}_{[j]}(t) + \Delta\tilde{\mathbf{w}}_{[j]}(t)$ as a nominal term, and a deviation term specific to the attack, derived from definitions in Table II. Note that bounds $\bar{\tilde{\mathbf{w}}}_{[j]}$ and $\bar{\boldsymbol{\rho}}_{[j]}$ are always satisfied.

Note that it is possible to redefine the state of (59) as $\mathbf{x}^a_{[j,i]}(t) \triangleq \mathbf{x}_{[j]}(t) + \Delta\mathbf{x}_{[j,i]}(t)$, where $\Delta\mathbf{x}_{[j,i]}(t)$ includes the effect of the attack on the state. The solution of (59) can therefore be computed for both covert and replay attacks as:

$$\begin{aligned} \mathbf{x}^a_{[j,i]}(t) = \mathbf{x}_{[j]}(t) + e^{\mathbf{A}_{jj}(t - T_a - nT)}\Delta\mathbf{x}_{[j,i]}(T_a + nT) + \\ + \int_{T_a + nT}^{t} e^{\mathbf{A}_{jj}(t-\tau)}\left[\mathbf{E}_j\Delta\mathbf{d}_{[j]}(\tau) + \Delta\tilde{\mathbf{w}}_{[j]}(\tau)\right]d\tau, \end{aligned} \tag{60}$$

where, in the case of covert attacks, $nT \triangleq 0$. From this, a (possibly discontinuous for replay attacks) solution of $\Delta\mathbf{x}_{[j,i]}(t)$ can be derived. The following holds for nonsingular $\mathbf{C}_j$:

Theorem 1. *If a replay or covert attack as in* (58)*, with dynamics as in* (59)*, and stealthy to the UIO-based detector in $\mathcal{D}_i$, is such that:*

$$\left| C_i \int_{T_a}^{t} e^{A_{Li}(t-\tau)} \left[\sum_{j \in \widehat{\mathcal{N}}_i} A_{ij}\Gamma\Delta\mathbf{x}_{[j,i]}(\tau)\right] d\tau \right| > 2\bar{r}_{[i]}(t) \tag{61}$$

is satisfied for some $t \geq T_a$, then the attack will be detected by the Luenberger-observer-based detector in $\mathcal{D}_i$. □

Proof. In order to prove that detection occurs, we must verify that either (5a) or (5b) must be violated, for some $t \geq T_a$. As it is assumed that attack function $\boldsymbol{\phi}_{j,i}(t)$ is defined as in (58), and is stealthy to $\mathcal{O}^{UIO}_{j,i}$ in $\mathcal{D}_i$, (5b) must not hold.

First, exploiting the formulation of the attack dynamics in (59), and the definition of $\Delta\mathbf{x}_{[j,i]}(t)$, one can see that $\mathbf{y}^c_{[j,i]}(t) = \mathbf{C}_j\mathbf{x}_{[j]}(t) + \boldsymbol{\rho}^a_{[j]}(t) + \mathbf{C}_j\Delta\mathbf{x}_{[j,i]}(t)$ and to detect the attack, $|r_{[i]}(t)| > \bar{r}_{[i]}(t)$ must be satisfied. Noting that, as seen from Table II, $|\rho^a_{[j]}| \leq \bar{\rho}_{[j]}$ is always satisfied, it is possible to divide the residual in healthy and attacked parts, as in Section V, with

$$r^a_{[j]}(t) = -C_i \int_{T_a}^{t} e^{A_{Li}(t-\tau)} \left[\sum_{j \in \mathcal{N}_i} A_{ij}\Gamma\Delta\mathbf{x}_{[j,i]}(\tau)\right] d\tau.$$

The rest of the proof follows that of Proposition 6, through the use of the triangle inequality. ■

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

Remark 14. In the case of singular C_j matrix, sufficient condition (61) changes to

$$\left| C_i \int_{T_a}^{t} e^{A_{Li}(t-\tau)} \left[\sum_{j \in \widehat{\mathcal{N}}_i} A_{ij} \Gamma \epsilon^a_{[j,i]}(\tau) \right] d\tau \right| > 2\bar{r}_{[i]}(t),$$

where $\epsilon^a_{[j,i]}$ is as in (23) with $\boldsymbol{\phi}_{[j,i]}(t) = \mathbf{C}_j \Delta x_{[j,i]}(t)$, i.e. the effect on $\mathcal{O}_i^{Luen}$ of the deviation provoked by $\Delta x_{[j,i]}$ on the UIO state estimate. ◁

Note that Theorem 1 provides bounds for *how much* an attacker implementing a covert or a replay attack may alter the behavior of the LSS, by establishing the maximum deviation of $\mathbf{x}^a_{[j,i]}$ from $\mathbf{x}_j$ before $\mathcal{O}_i^{Luen}$ is guaranteed to detect it. For replay attacks, this implies that if the operating condition of $\mathcal{S}_j$ changes significantly over time, then it will be detected by $\mathcal{O}_i^{Luen}$. On the other hand, for covert attacks, if the attacker's input $\mathbf{d}^a_{[j]}(t)$ deviates significantly from the true $\mathbf{d}_{[j]}(t)$, it will be detected, limiting the malicious agent's impact on the LSS overall.

VII. Simulation Results

A. Simulation Setup

The proposed scheme is verified through realistic simulations in Simulink, using the *Specialized Power Systems Toolbox* [45]. The considered microgrid topology is that in Figure 1, having source voltages $V_{si} = 60V, \forall i \in \mathcal{N}$, and employing bidirectional Buck converters realized as non-ideal IGBT switches, operating at 10 kHz, with snubbers to suppress large transients and protect the equipment. Although power lines are considered to be purely resistive in the development of the results, RL power lines are employed for the physical connection of DGUs in the simulations.

The parameters of the electrical components and primary controllers are taken from [46]. Voltages are measured in $[V]$ and currents are measured in $[A]$, whereas the unit of the integrator state is $[V \cdot s]$. The effect of model mismatch is modeled as bounded process noise $w_{[i]}$, $\forall i \in \mathcal{N}$. The process and measurement noises satisfy Assumption 2, with $\bar{w}_i = [0.05, 0.05, 0.01]^\top$ and $\bar{\rho}_i = [0.01, 0.01, 0]^\top$, $\forall i \in \mathcal{N}$.

The Luenberger observer gains L_i are calculated to assign the eigenvalues of A_{Li} to $\{-50, -100, -500\}$ for each DGU i. The UIO matrices $S_j = \mathbf{I} - H_j \mathbf{C}_j$ are selected to ensure $S_j \mathbf{E}_j = \mathbf{0}$. Matrices K_j are calculated to assign the eigenvalues of F_j to $\{-1, -1.5, -2, -2.5, -3, -3.5\}$. All other UIO matrices are computed as in (9). Two attack scenarios will be discussed in the following subsections. In the first, attacks on $\mathbf{y}^c_{[2,4]}$ and $\mathbf{y}^c_{[3,4]}$ will be designed to be stealthy to the Luenberger-like observer as per condition in Proposition 8. In the second scenario, a covert attack will be implemented on $\mathbf{y}^c_{[2,4]}$. These two scenarios have been specifically designed to demonstrate the interplay between the two modules of $\mathcal{D}_i$.

For both scenarios, the simulation proceeds as follows. At time $t = 0s$, all DGUs are started disconnected from each other, i.e., DGUs are running separately; therefore, power lines and communication links in Figure 1 are not in place. Consequently, at this phase of simulations, the secondary controllers of DGUs are not active and primary controllers track a constant voltage reference of $V_{ref} = 48V$. At time $t = 2s$, the DGUs are connected to each other through both RL power lines and communication links, and secondary controllers are activated. At this phase of the simulations, the communications are *healthy*, i.e., no attacks are active, and therefore, the secondary controllers will achieve current sharing and voltage balancing. Finally, at $t = 8s$, the attack is launched on the corresponding communication channels.

B. Scenario I – False data injection stealthy to $\mathcal{O}_i^{Luen}$

In the first scenario, constant bias injection attacks are directed to communications $\mathbf{y}^c_{[2,4]}(t)$ and $\mathbf{y}^c_{[3,4]}(t)$, where the elements of the attack vector $\boldsymbol{\phi}^{bi}_{2,4}$ are selected randomly from a uniform distribution in the interval $[-0.02, 0.02]$. The fourth element of the attack vector $\boldsymbol{\phi}^{bi}_{3,4}$ is selected as $\boldsymbol{\phi}^{bi}_{3,4_4} = -\frac{R_{34}}{R_{24}} \boldsymbol{\phi}^{bi}_{2,4_4}$ making them stealthy to $\mathcal{O}_i^{Luen}$, as per condition (55). Remaining elements of the attack vector $\boldsymbol{\phi}^{bi}_{3,4}$ are again drawn from a uniform distribution in the interval specified above. Specifically, the constant attack vectors are:

$$\boldsymbol{\phi}^{bi}_{2,4} = [-0.0139, 0.0149, -0.0031, -0.0014, -0.0095, -0.0011]^\top$$
$$\boldsymbol{\phi}^{bi}_{3,4} = [0.0037, 0.0185, 0.0174, 0.0021, 0.0178, 0.0180]^\top.$$

Figure 2 displays the residuals and corresponding thresholds for the Luenberger-observer-based module for DGU 4, and UIO-based modules for communication $\mathbf{y}^c_{[j,4]}(t), j \in \{2, 3\}$, in Figures 2a, 2b, and 2c respectively. Moreover, in these figures, the vertical dashed lines in black indicate the time of the start of the attacks, i.e., $T_a^{2,4} = T_a^{3,4} = 8s$, whereas those in green indicate the time of detection for the corresponding module.

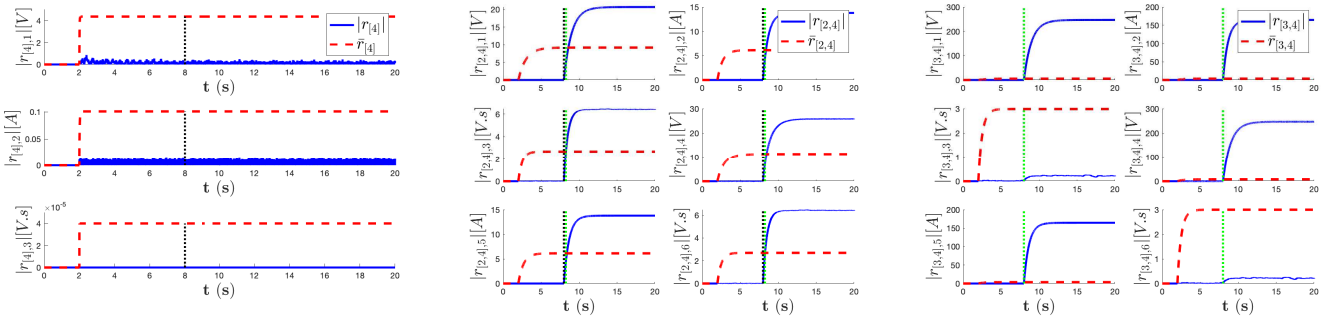
One can see that, through the proper selection of the attack vectors $\boldsymbol{\phi}^{bi}_{2,4}$ and $\boldsymbol{\phi}^{bi}_{3,4}$, the attacker is able to achieve stealthiness condition (55) for $\mathcal{O}_i^{Luen}$. Hence, the residual of this module is unaffected by the attack, preventing detection, as is shown in Figure 2a. Nevertheless, the residuals of $\mathcal{O}_{j,i}^{UIO}$monitoring the two communication links are affected by the attack, leading to the violation of (17) for $(j, i) = (2, 4)$ and $(j, i) = (3, 4)$ in turn triggering detection in both modules. The attacks are detected at times $T_d^{2,4} = 8.270$s and $T_d^{3,4} = 8.015$s, shortly after activation.

C. Scenario II – Covert attack

In the second scenario, a covert attack $\boldsymbol{\phi}^c_{2,4}$ is launched on the communication $\mathbf{y}^c_{[2,4]}(t)$, with dynamics (38). The inputs $\mathbf{d}^a_{[2]}$ are such that the state dynamics of the attacked system act as if DGU 2 were disconnected from the rest of the microgrid, i.e. dynamics of $\mathbf{x}_{[2]}$ not influenced by its neighboring states nor by secondary consensus input $\alpha_{[2]}$. Furthermore, the attacker also specifies a difference in load current I^a_{L2} in $\mathbf{d}^a_{[2]}$, selected such that $\Delta I_{L2} = 2A$, to alter the operation point of $\mathbf{x}^a_{[2]}$ compared to $\mathbf{x}_{[2]}$.

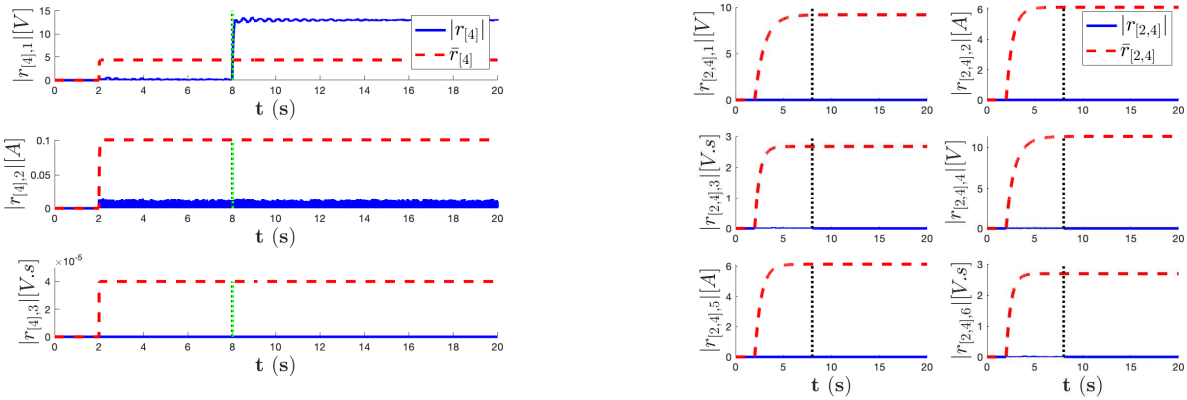
Figures 3a-3b show the residuals and corresponding thresholds for the second attack scenario, for the Luenberger-observer-based module for DGU 4 and UIO-based module for communication $\mathbf{y}^c_{[2,4]}(t)$, respectively. Since this covert attack complies with the dynamics in (38), it is stealthy to the UIO-based detection module as proven in Proposition 5. Indeed,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control



(a) Residuals and thresholds – $\mathcal{O}_4^{Luen}$ (b) Residuals and thresholds – $\mathcal{O}_{2,4}^{UIO}$ (c) Residuals and thresholds – $\mathcal{O}_{3,4}^{UIO}$

Fig. 2: Residual and detection thresholds of the different modules in $\mathcal{D}_4$ under Scenario I. The false data injection attacks $\phi_{2,4}^{bi}$ and $\phi_{3,4}^{bi}$ are detected by the UIO modules $\mathcal{O}_{2,4}^{UIO}$ and $\mathcal{O}_{3,4}^{UIO}$, whilst not detected by $\mathcal{O}_4^{Luen}$, as Proposition 8 is satisfied.



(a) Residuals and thresholds – $\mathcal{O}_4^{Luen}$ (b) Residuals and thresholds – $\mathcal{O}_{2,4}^{UIO}$

Fig. 3: Residual and detection thresholds of the different modules in $\mathcal{D}_4$ under Scenario II. The covert attack $\phi_{2,4}^{c}$ is stealthy to the UIO modules, but is detected by $\mathcal{O}_4^{Luen}$.

one can see from Figure 3b that the residual of the UIO-based module is unchanged by the onset of the attack and this module fails to detect the attack. On the other hand, residual of the Luenberger-observer-based module reflects the effect of the attack, and the covert attack is quickly detected at time $T_d^4 = 8.001s$.

VIII. Concluding Remarks

In this paper, we have presented a novel *distributed* attack detection technique for LSS inspired by and applied to low-voltage DC microgrids. We have discussed the architecture and the properties of a two-module local detection unit $\mathcal{D}_i$, composed of a Luenberger-like observer and a bank of suitably designed unknown-input observers estimating local and neighboring states, respectively. Details on the information necessary for the design of each module are given explicitly, requiring knowledge of dynamics of the local subsystem and of its neighbors. Thorough analysis has been provided as well as extensive simulation results on a realistic model of a DC microgrid showing the methodology's effectiveness.

The detection architecture relies on the assumption that model uncertainties are unstructured but bounded, as well as on an assumption of *ideality* of the communication network. The relaxation of the latter, through the introduction of, e.g., delays, will be the focus of future research, further improving the technique's real-world applicability. We also wish to explore the possibility of exploiting the Plug-and-Play capabilities of the controllers to develop an automatic reconfiguration strategy after detection and isolation, focusing on the scalability of the proposed method.

Finally, we intend to analyze the relationship between the properties of $\mathcal{D}_i$ and the information used in its design and operation. Specifically, future improvements to the detector's properties may be found if further measurements, containing information content different to that used in this work, were available. Further to this, we will consider potential alternatives to reduce the information required in the design of the two proposed modules.

Appendix

A. DGU Dynamics

Matrices A_{ii}, B_i, M_i, A_{ij}, K_i, and C_i are defined as [8]:

$$A_{ii} = \begin{bmatrix} -\sum_{j\in\mathcal{N}_i} \frac{1}{R_{ij}C_{ti}} & \frac{1}{C_{ti}} & 0 \\ -\frac{1}{L_{ti}} & -\frac{R_{ti}}{L_{ti}} & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2982577, IEEE Transactions on Automatic Control

$$B_i = \begin{bmatrix} 0 & 0 \\ \frac{1}{L_{ti}} & 0 \\ 0 & 1 \end{bmatrix}, K_i = \begin{bmatrix} k_{i,1} & k_{i,2} & k_{i,3} \end{bmatrix},$$

$$M_i = \begin{bmatrix} -\frac{1}{C_{ti}} & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, A_{ij} = \begin{bmatrix} \frac{1}{R_{ij}C_{ti}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where $R_{ti}, L_{ti}, C_{ti}, R_{ij}$ are electrical parameters of the DGU as seen in Figure 1. For the design of the UIOs, the DGU dynamics are rearranged as in (11), with $\bar{E}_j = \mathbf{I}$ and $\hat{E}_j$ defined accordingly.

B. Proof of Lemma 1

We provide a sketch of the proof. To simplify notation, without risk of ambiguity, we remove all subscripts from variables in (6) and (9), and replace $\bar{E}$ with E:

$$\dot{x}(t) = Ax(t) + Ed(t) + w(t)$$
$$y^c(t) = Cx(t) + \rho(t) + \beta(t - T_a)\phi(t)$$

Exploiting [47, Lemma 1], condition (C1) implies that there are nonsingular matrices P and Q such that

$$P^{-1}E = \begin{bmatrix} E_1 \\ \mathbf{0} \end{bmatrix} \quad Q^{-1}CP = \begin{bmatrix} C_1 & \mathbf{0} \end{bmatrix}, \tag{62}$$

where E_1 and C_1 have the same dimension and are both invertible. It is possible to construct a UIO for the transformed dynamics for state $\bar{x} = P^{-1}x$ and output $\bar{y}^c = Q^{-1}y^c$, noting that conditions (C1) and (C2) hold for the transformed dynamics, and defining $\bar{A} \triangleq P^{-1}AP$, $\bar{E} \triangleq P^{-1}E$, and $\bar{C} \triangleq Q^{-1}CP$, with $\bar{A}$ a 2-by-2 block matrix with entries $\bar{A}_{lk}, l, k \in \{1, 2\}$. From (9) one derives the following[3]:

$$\bar{H} = \begin{bmatrix} C_1^{-1} \\ \mathbf{0} \end{bmatrix}, \quad \bar{S} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-q} \end{bmatrix}, \quad \bar{\tilde{K}} = \begin{bmatrix} \bar{\tilde{K}}_1 \\ \bar{\tilde{K}}_2 \end{bmatrix}$$

$$\bar{F} = \begin{bmatrix} -\bar{\tilde{K}}_1 C_1 & \mathbf{0} \\ \bar{A}_{21} - \bar{\tilde{K}}_2 C_1 & \bar{A}_{22} \end{bmatrix}, \quad \bar{\hat{K}} = \begin{bmatrix} \mathbf{0} \\ \bar{A}_{21} C_1^{-1} \end{bmatrix}$$

where $\bar{F}$ is Hurwitz stable by design. Note that the pair $(\bar{F}, \bar{C})$ is not observable and thus the state $\bar{z}$ can be written as $\bar{z} = \left[\bar{z}_1^\top, \bar{z}_2^\top\right]^\top$, where $\bar{z}_1$ and $\bar{z}_2$ are respectively the observable and unobservable portions of the state. Furthermore, given the structure of transformed matrices above, it is evident that $\bar{y}^c$ does not influence the observable part of the state, $\bar{z}_1$, and $\hat{\bar{y}} = C_1\bar{z}_1 + \bar{y}^c$. Therefore, the residual defined as $r = Q\left(\bar{y}^c - \hat{\bar{y}}\right) = QC_1\bar{z}_1$ is independent of $\bar{y}^c(t)$ and $\phi(t), \forall t \geq T_a$. ■

C. Proof of Lemma 3

For any matrix A, it is possible to find its Jordan normal form $J = P^{-1}AP$. This implies that the equivalence $e^{At} = Pe^{Jt}P^{-1}$ holds. Note that e^{Jt} also has the same block-diagonal structure of J, where each block $e^{J_k t} \in \mathbb{R}^{n_k \times n_k}$ is upper-triangular. Following this, we define a block-diagonal matrix $\mathbf{P}$ such that each block $\mathbf{P}_k \in \mathbb{R}^{n_k \times n_k}$ is upper-triangular with all entries 1, and thus has the same non-zero structure as $e^{J_k t}$. We exploit the property that for any matrix M with element m_{ij} at i-th row and j-th column, it holds that $\max|m_{ij}| \leq \|M\|$ to show that $\left|Pe^{Jt}P^{-1}\right| \leq |P|\left|e^{Jt}\right||P^{-1}| \leq |P|\left\|e^{Jt}\right\|\mathbf{P}|P^{-1}|$ is satisfied elementwise. Hence, noting that if A is Hurwitz stable then J is also Hurwitz stable, it is possible to find scalars $\lambda > 0$ and $\mu \geq 1$ such that $\|e^{Jt}\| \leq \mu e^{-\lambda t}$, and to define $\Lambda \triangleq \mu|P|\mathbf{P}|P^{-1}|$. Finally, note that matrix Λ is such that $\Lambda \geq \mathbf{I}$ holds, as the following relationships can be derived: $|P|\mathbf{P}|P^{-1}| = |P|(\mathbf{I}+\Xi)|P^{-1}| = |P||P^{-1}| + |P|\Xi|P^{-1}| \geq \mathbf{I} + \mathbf{0}$, given that $\mathbf{P} = \mathbf{I} + \Xi$ with $\Xi \geq \mathbf{0}$, and that $|P||P^{-1}| \geq |PP^{-1}| = \mathbf{I}$, $|P|\Xi|P^{-1}| \geq \mathbf{0}$. ■

[3]A bar $\bar{\cdot}$ has been added to the matrices to highlight their dependence on the transformed system.

D. Proof of Proposition 4

We generalize Lemma 4 to find a condition on $\epsilon^r_{[j,i]}(T_a + kT)$, for any $k \in \mathbb{N}$ such that (5a) will hold for all $t \in [T_a + kT, T_a + (k+1)T)$. Following the Proof of Lemma 4, if

$$\Sigma_j \left|\epsilon^r_{[j,i]}(T_a + kT) + H_j\boldsymbol{\rho}_{[j]}(T_a - T)\right| \leq \leq \bar{\epsilon}_{[j,i]}(T_a + kT) - |H_j|\bar{\boldsymbol{\rho}}_{[j]}, \tag{63}$$

then the replay attack will not be detected for $t \in [T_a + kT, T_a + (k+1)T)$. We therefore must characterize the solution of estimation error $\epsilon_{[j,i]}(T_a + kT)$ as $k \to \infty$. To do so, we note that the solution to the state estimate under replay attack for time $t \in \mathcal{T} \triangleq \{t | t = T_a + kT, \forall k \in \mathbb{N}\}$ is:

$$\begin{aligned}\hat{\mathbf{x}}_{[j,i]}(T_a + kT) &= e^{F_j kT}\mathbf{z}_{[j,i]}(T_a) + H_j\mathbf{y}^c_{[j,i]}(T_a + kT) \\ &+ \int_{T_a}^{T_a+kT} e^{F_j(T_a+kT-\tau)}\widehat{K}_j\mathbf{y}^c_{[j,i]}(\tau)d\tau \\ &= e^{F_j kT}\mathbf{z}_{[j,i]}(T_a) + H_j\mathbf{y}_{[j]}(T_a - T) + \\ &+ \sum_{s=0}^{k-1} e^{sTF_j}\int_{T_a-T}^{T_a} e^{F_j(T_a-\tau)}\widehat{K}_j\mathbf{y}_{[j]}(\tau)d\tau.\end{aligned} \tag{64}$$

Given that F_j is Hurwitz by design, the series $\sum_{s=0}^{k-1} e^{sTF_j}\int_{T_a-T}^{T_a} e^{F_j(T_a-\tau)}\widehat{K}_j\mathbf{y}_{[j]}(\tau)d\tau$ converges. Hence $\epsilon^r_{[j,i]}(T_a + kT)$ can be expressed as:

$$\begin{aligned}\epsilon^r_{[j,i]}(T_a + kT) &= \epsilon^r_{[j,i]}(T_a) + (\mathbf{I} - e^{F_j kT})\mathbf{z}_{[j,i]}(T_a) + \\ &- \sum_{s=0}^{k-1} e^{sTF_j}\int_{T_a-T}^{T_a} e^{F_j(T_a-\tau)}\widehat{K}_j\mathbf{y}_{[j]}(\tau)d\tau \\ &= \epsilon^r_{[j,i]}(T_a) + \Delta\epsilon_{[j,i]}(k).\end{aligned}$$

Given the convergence of the series, it is possible to bound the estimation error under attack by:

$$\left|\epsilon^r_{[j,i]}(T_a + kT)\right| \leq \left|\epsilon^r_{[j,i]}(T_a) + M_{j,i}\right| \tag{65}$$

where $M_{j,i} \geq \mathbf{0}$ is such that:

$$\left|\Delta\epsilon_{[j,i]}(k)\right| \leq M_{j,i}\,, \forall k \in \mathbb{N}^0.$$

Hence, $M_{j,i}$ can be defined as $M_{j,i} \triangleq \sup_{k\in\mathbb{N}^0}\left|\Delta\epsilon_{[j,i]}(k)\right|$. Given the monotonicity of the LHS of the previous, $M_{j,i}$ is an upper bound on $\left|\Delta\epsilon_{[j,i]}(k)\right|$ for all $k \in \mathbb{N}$, implying

$$|\epsilon^r_{[j,i]}(T_a + kT) + H_j\boldsymbol{\rho}_{[j]}(T_a - T)| \leq \left|\epsilon^r_{[j,i]}(T_a) + M_{j,i} + H_j\boldsymbol{\rho}_{[j]}(T_a - T)\right|. \tag{66}$$

Finally, to complete the proof, note that $\bar{\epsilon}_{[j,i]}(t)$ is monotonic. We define a variable

$$\Delta\bar{\epsilon}_{[j,i]}(T_a) \triangleq \max\left(\mathbf{0}, \lim_{t\to\infty} \bar{\epsilon}_{[j,i]}(T_a) - \bar{\epsilon}_{[j,i]}(t)\right)$$

which is $\mathbf{0}$ if $\bar{\epsilon}_{[j,i]}(t)$ monotonically increasing, and greater than $\mathbf{0}$ otherwise. This definition allows us to state that $\bar{\epsilon}_{[j,i]}(T_a + kT) \geq \bar{\epsilon}_{[j,i]}(T_a) - \Delta\bar{\epsilon}_{[j,i]}(T_a), \forall k \in \mathbb{N}_0$.

Therefore, recalling $\Sigma_j \geq \mathbf{I}$, if (36) is satisfied, (63) will also hold, as:

$$\begin{aligned}\left|\epsilon^r_{[j,i]}(T_a + kT)\right| - |H_j|\bar{\boldsymbol{\rho}}_{[j]} \leq \\ \leq \Sigma_j \left|\epsilon^r_{[j,i]}(T_a + kT) + H_j\boldsymbol{\rho}_{[j]}(T_a - T)\right| \leq \\ \leq \bar{\epsilon}_{[j,i]}(T_a) - \Delta\bar{\epsilon}_{[j,i]}(T_a) - |H_j|\bar{\boldsymbol{\rho}}_{[j]} \leq \\ \leq \bar{\epsilon}_{[j,i]}(T_a + kT) - |H_j|\bar{\boldsymbol{\rho}}_{[j]}, \ \forall k \in \mathbb{N}_0.\end{aligned}$$

Consequently, $|\epsilon^r_{[j,i]}(t)| \leq \bar{\epsilon}_{[j,i]}(t)$ will hold for all $t \geq T_a$. ■

REFERENCES

[1] L. Meng, Q. Shafiee, G. Ferrari-Trecate, H. Karimi, D. Fulwani, X. Lu, and J. M. Guerrero, "Review on control of DC microgrids and multiple microgrid clusters," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 5, no. 3, pp. 928–948, 2017.

[2] T. Samad and T. Parisini, "Systems of systems," *The Impact of Control Technology*, pp. 175–183, 2011.

[3] A. Mitra and S. Sundaram, "Secure distributed observers for a class of linear time invariant systems in the presence of byzantine adversaries," in *55th IEEE Conference on Decision and Control (CDC)*, pp. 2709–2714, 2016.

[4] H. Nishino and H. Ishii, "Distributed detection of cyber attacks and faults for power systems," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 11 932–11 937, 2014.

[5] S. Dibaji, M. Pirani, A. Annaswamy, K. Johansson, and A. Chakrabortty, "Secure control of wide-area power systems: Confidentiality and integrity threats," in *57th IEEE Conference on Decision and Control (CDC)*, pp. 7269–7274, 2018.

[6] F. Boem, A. J. Gallo, G. Ferrari-Trecate, and T. Parisini, "A distributed attack detection method for multi-agent systems governed by consensus-based control," in *56th IEEE Conference on Decision and Control (CDC)*, pp. 5961–5966, 2017.

[7] A. J. Gallo, M. S. Turan, P. Nahata, F. Boem, T. Parisini, and G. Ferrari-Trecate, "Distributed cyber-attack detection in the secondary control of DC microgrids," in *European Control Conference (ECC)*, pp. 344–349, 2018.

[8] M. Tucci, S. Riverso, and G. Ferrari-Trecate, "Line-independent plug-and-play controllers for voltage stabilization in DC microgrids," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 3, pp. 1115–1123, 2018.

[9] Y. Mo, T. H.-J. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli, "Cyber–physical security of a smart grid infrastructure," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 195–209, 2012.

[10] A. Teixeira, D. Pérez, H. Sandberg, and K. H. Johansson, "Attack models and scenarios for networked control systems," in *1st International Conference on High Confidence Networked Systems*. ACM, pp. 55–64, 2012.

[11] P. Cheng, L. Shi, and B. Sinopoli, "Guest editorial special issue on secure control of cyber-physical systems," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 1–3, 2017.

[12] H. Sandberg, S. Amin, and K. H. Johansson, "Cyberphysical security in networked control systems: An introduction to the issue," *IEEE Control Systems*, vol. 35, no. 1, pp. 20–23, 2015.

[13] D. I. Urbina, J. Giraldo, A. A. Cardenas, J. Valente, M. Faisal, N. O. Tippenhauer, J. Ruths, R. Candell, and H. Sandberg, *Survey and new Directions for Physics-Based Attack Detection in Control Systems*. US Department of Commerce, National Institute of Standards and Technology, 2016.

[14] A. A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *28th International Conference on Distributed Computing Systems Workshops*. IEEE, pp. 495–500, 2008.

[15] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security & Privacy*, vol. 7, no. 3, pp. 75–77, 2009.

[16] H. Khurana, M. Hadley, N. Lu, and D. A. Frincke, "Smart-grid security issues," *IEEE Security & Privacy*, vol. 8, no. 1, pp. 81–85, 2010.

[17] A. R. Metke and R. L. Ekl, "Security technology for smart grid networks," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 99–107, 2010.

[18] X. Zhong, L. Yu, R. Brooks, and G. K. Venayagamoorthy, "Cyber security in smart dc microgrid operations," in *1st IEEE International Conference on DC Microgrids (ICDCM)*, pp. 86–91, 2015.

[19] G. Fiore, A. Iovine, E. De Santis, and M. D. Di Benedetto, "Secure state estimation for DC microgrids control," in *13th IEEE Conference on Automation Science and Engineering (CASE)*, pp. 1610–1615, 2017.

[20] O. A. Beg, L. V. Nguyen, T. T. Johnson, and A. Davoudi, "Signal temporal logic-based attack detection in DC microgrids," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3585–3595, 2019.

[21] S. Sahoo, S. Mishra, J. C.-H. Peng, and T. Dragicevic, "A stealth cyber attack detection strategy for DC microgrids," *IEEE Transactions on Power Electronics*, vol. 34, no. 8, pp. 8162–8174, 2018.

[22] K. Paridari, N. O'Mahony, A. E.-D. Mady, R. Chabukswar, M. Boubekeur, and H. Sandberg, "A framework for attack-resilient industrial control systems: Attack detection and controller reconfiguration," *Proceedings of the IEEE*, vol. 106, no. 1, pp. 113–128, 2018.

[23] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

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

[25] Y. Mo, S. Weerakkody, and B. Sinopoli, "Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs," *IEEE Control Systems*, vol. 35, no. 1, pp. 93–109, 2015.

[26] R. S. Smith, "Covert misappropriation of networked control systems: Presenting a feedback structure," *IEEE Control Systems*, vol. 35, no. 1, pp. 82–92, 2015.

[27] R. Anguluri, V. Katewa, and F. Pasqualetti, "Attack detection in stochastic interconnected systems: Centralized vs decentralized detectors," in *57th IEEE Conference on Decision and Control (CDC)*, pp. 4541–4546, 2018.

[28] F. Pasqualetti, F. Dörfler, and F. Bullo, "A divide-and-conquer approach to distributed attack identification," in *54th IEEE Conference on Decision and Control*, pp. 5801–5807, 2015.

[29] A. Barboni, H. Rezaee, F. Boem, and T. Parisini, "Distributed detection of covert attacks for interconnected systems," in *European Control Conference (ECC)*, pp. 2240–2245, 2019.

[30] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, "Distributed fault diagnosis and fault-tolerant control," in *Diagnosis and Fault-Tolerant Control*. Springer, pp. 467–518, 2016.

[31] I. Shames, A. M. Teixeira, H. Sandberg, and K. H. Johansson, "Distributed fault detection for interconnected second-order systems," *Automatica*, vol. 47, no. 12, pp. 2757–2764, 2011.

[32] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "Distributed fault detection and isolation resilient to network model uncertainties," *IEEE Transactions on Cybernetics*, vol. 44, no. 11, pp. 2024–2037, 2014.

[33] M. Davoodi, N. Meskin, and K. Khorasani, "Simultaneous fault detection and consensus control design for a network of multi-agent systems," *Automatica*, vol. 66, pp. 185–194, 2016.

[34] F. Boem, R. M. G. Ferrari, C. Keliris, T. Parisini, and M. M. Polycarpou, "A distributed networked approach for fault detection of large-scale systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 18–33, 2017.

[35] S. Riverso, F. Boem, G. Ferrari-Trecate, and T. Parisini, "Plug-and-play fault detection and control-reconfiguration for a class of nonlinear large-scale constrained systems," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 3963–3978, 2016.

[36] F. Pasqualetti, A. Bicchi, and F. Bullo, "Consensus computation in unreliable networks: A system theoretic approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 90–104, 2012.

[37] R. D. Middlebrook and S. Cuk, "A general unified approach to modelling switching-converter power stages," in *IEEE Power Electronics Specialists Conference*, pp. 18–34, 1976.

[38] J. Zhao and F. Dörfler, "Distributed control and optimization in DC microgrids," *Automatica*, vol. 61, pp. 18–26, 2015.

[39] M. Tucci, L. Meng, J. M. Guerrero, and G. Ferrari-Trecate, "Stable current sharing and voltage balancing in DC microgrids: A consensus-based secondary control layer," *Automatica*, vol. 95, pp. 1–13, 2018.

[40] C. De Persis, E. Weitenberg, and F. Dörfler, "A power consensus algorithm for DC microgrids," *Automatica*, vol. 89, pp. 364–375, 2018.

[41] G. Cavraro, S. Bolognani, R. Carli, and S. Zampieri, "The value of communication in the voltage regulation problem," in *55th IEEE Conference on Decision and Control (CDC)*, pp. 5781–5786, 2016.

[42] Z. Wang, B. Chen, J. Wang, and J. Kim, "Decentralized energy management system for networked microgrids in grid-connected and islanded modes," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 1097–1105, 2016.

[43] 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.

[44] A. J. Gallo, M. S. Turan, F. Boem, G. Ferrari-Trecate, and T. Parisini, "Distributed watermarking for secure control of microgrids under replay attacks," in *7th IFAC Workshop on Distributed Estimation and Control in Networked Systems (NecSys'18)*, pp. 182–187, 2018.

[45] Simscape electrical: user's guide (specialized power systems). [Online]. Available: https://mathworks.com/help/physmod/sps/specialized-power-systems.html

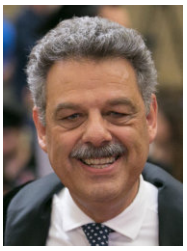
[46] M. Tucci, L. Meng, J. M. Guerrero, and G. Ferrari-Trecate, "Consensus algorithms and plug-and-play control for current sharing in DC microgrids," *arXiv preprint arXiv:1603.03624*, 2016.

[47] M. Corless and J. Tu, "State and input estimation for a class of uncertain systems," *Automatica*, vol. 34, no. 6, pp. 757–764, 1998.

Alexander J. Gallo (S'16) Alexander J. Gallo received the MEng in Electrical and Electronic Engineering from Imperial College, London, UK, in 2016. He is currently pursuing a PhD at the Electrical and Electronic Engineering Department, Imperial College, London, UK, with the Control and Power Research Group. His research interests include distributed fault diagnosis for large-scale systems, as well as distributed methods for secure control of cyber physical systems.

Mustafa S Turan Mustafa Sahin Turan received the B.S. degree in mechatronics engineering from Sabanci University, Istanbul, Turkey and the M.S. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2015 and 2017, respectively. He is currently pursuing the Ph.D. degree with École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, under the supervision of Prof. Giancarlo Ferrari Trecate. His current research interests include networked control and cyber-security of microgrids.

Francesca Boem (M'18) received the MSc degree (cum laude) in Management Engineering in 2009 and the PhD degree in Information Engineering in 2013, both from the University of Trieste, Italy. She was Post-Doc at the University of Trieste with the Machine Learning Group from 2013 to 2014. From 2014 to 2018, she was Research Associate at the Department of Electrical and Electronic Engineering, Imperial College London, with the Control and Power Research Group. Since 2015 she has been part of the team at Imperial College working on the flagship EU H2020-WIDESPREAD-TEAMING project for the development of the EU KIOS Research and Innovation Centre of Excellence, a strategic partnership between University of Cyprus and Imperial College London. In 2018 Dr. Boem has been appointed as a Lecturer in the Department of Electronic and Electrical Engineering at University College London. Her current research interests include distributed fault diagnosis and fault-tolerant control methods for large-scale networked systems and distributed estimation methods for sensor networks. Dr. Boem is member of the IFAC Technical Committee 6.4 ("Fault Detection, Supervision & Safety of Technical Processes - SAFEPROCESS") and Associate Editor for the IEEE Systems Journal, for the IEEE Control System Society Conference Editorial Board and for the EUCA Conference Editorial Board.



Thomas Parisini (F'11) received the Ph.D. degree in Electronic Engineering and Computer Science in 1993 from the University of Genoa. He was with Politecnico di Milano and since 2010 he holds the Chair of Industrial Control and is Director of Research at Imperial College London. He is a Deputy Director of the KIOS Research and Innovation Centre of Excellence, University of Cyprus. Since 2001 he is also Danieli Endowed Chair of Automation Engineering with University of Trieste. In 2009-2012 he was Deputy Rector of University of Trieste. In 2018 he received an *Honorary Doctorate* from University of Aalborg, Denmark. He authored or co-authored more than 320 research papers in archival journals, book chapters, and international conference proceedings. His research interests include neural-network approximations for optimal control problems, distributed methods for cyber-attack detection and cyber-secure control of large-scale systems, fault diagnosis for nonlinear and distributed systems, nonlinear model predictive control systems and nonlinear estimation. He is a co-recipient of the IFAC Best Application Paper Prize of the Journal of Process Control, Elsevier, for the three-year period 2011-2013 and of the 2004 Outstanding Paper Award of the IEEE Trans. on Neural Networks. He is also a recipient of the 2007 IEEE Distinguished Member Award. In 2016, he was awarded as Principal Investigator at Imperial of the H2020 European Union flagship Teaming Project KIOS Research and Innovation Centre of Excellence led by University of Cyprus. In 2012, he was awarded an ABB Research Grant dealing with energy-autonomous sensor networks for self-monitoring industrial environments. Thomas Parisini currently serves as 2020 President-Elect of the IEEE Control Systems Society and has served as Vice-President for Publications Activities. During 2009-2016 he was the Editor-in-Chief of the IEEE Trans. on Control Systems Technology. Since 2017, he is Editor for Control Applications of Automatica and since 2018 he is the Editor in Chief of the European Journal of Control. He is also the Chair of the IFAC Technical Committee on Fault Detection, Supervision & Safety of Technical Processes - SAFEPROCESS. Among other activities, he was the Program Chair of the 2008 IEEE Conference on Decision and Control and General Co-Chair of the 2013 IEEE Conference on Decision and Control. Prof. Parisini is a Fellow of the IEEE and of the IFAC.

Giancarlo Ferrari-Trecate (SM'12) Giancarlo Ferrari-Trecate received the Ph.D. degree in Electronic and Computer Engineering from the Università degli Studi di Pavia in 1999. Since September 2016 he is Professor at EPFL, Lausanne, Switzerland. In spring 1998, he was a Visiting Researcher at the Neural Computing Research Group, University of Birmingham, UK. In fall 1998, he joined as a Postdoctoral Fellow the Automatic Control Laboratory, ETH, Zurich, Switzerland. He was appointed Oberassistent at ETH, in 2000. In 2002, he joined INRIA, Rocquencourt, France, as a Research Fellow. From March to October 2005, he was researcher at the Politecnico di Milano, Italy. From 2005 to August 2016, he was Associate Professor at the Dipartimento di Ingegneria Industriale e dell'Informazione of the Università degli Studi di Pavia.

His research interests include scalable control, microgrids, networked control systems, hybrid systems and machine learning.

Giancarlo Ferrari-Trecate was the recipient of the Researcher Mobility Grant from the Italian Ministry of Education, University and Research in 2005. He is currently a member of the IFAC Technical Committees on Control Design and Optimal Control, and the Technical Committee on Systems Biology of the IEEE SMC society. He has been Editor at large for the 2019 ACC and has been serving on the editorial board of Automatica for three terms and of Nonlinear Analysis: Hybrid Systems.