Distributed Cyber-Attack Detection in the Secondary Control of DC Microgrids

A. J. Gallo, M. S. Turan, P. Nahata, F. Boem, T. Parisini, G. Ferrari-Trecate

***Abstract*— The paper considers the problem of detecting cyber-attacks occurring in communication networks typically used in the secondary control layer of DC microgrids. The proposed distributed methodology allows for scalable monitoring of a microgrid and is able to detect the presence of data injection attacks in the communications among Distributed Generation Units (DGUs) - governed by consensus-based control - and isolate the communication link over which the attack is injected. Each local attack detector requires limited knowledge regarding the dynamics of its neighbors. Detectability properties of the method are analyzed, as well as a class of undetectable attacks. Some results from numerical simulation are presented to demonstrate the effectiveness of the proposed approach.**

I. Introduction

Microgrids offer several advantages over conventional electrical power networks. Due to localized generation and compatibility with various renewable sources, they ensure clean high quality power with minimal transmission losses. Microgrids, both AC and DC, are composed of distributed generation units (DGUs), storage, and loads. DC microgrids (DCmGs) hold tremendous promise, as a large portion of loads are inherently DC, and have attracted significant research attention [1].

To ensure stable and efficient operation of microgrids, a hierarchical control architecture is adopted [2]. The primary control performs the decentralized control of local power, voltage, and current [3], [4], [5]. The secondary and tertiary layers deal with power quality regulation, load sharing, DGU coordination, microgrid synchronization and optimization, and enforcing system regulations [1], [5], [6], [7]. As shown in [1], [8], certain secondary and tertiary objectives require communication between the control layers. However, the information flowing through communication channels is vulnerable to malicious attacks and can be compromised [9], [10], [11]. These attacks are aimed at hindering normal microgrid operation and can have detrimental consequences like voltage instability, damage of critical loads, blackouts, etc [12]. Therefore, detection and isolation of attacks are necessary in order to undertake remedial actions.

The field of attack detection and isolation, which plays a central role in secure control systems, has attracted growing interest in recent years [13], [14]. This is due to the fact that an ever growing number of control systems are integrating communication tools to regulate the processes. This in turn has exposed them to attacks. Some approaches in security of cyber-physical systems stem from prior research in the field of fault detection and isolation (FDI), a well established area of research which focuses on identifying if the behaviour of the underlying process is *healthy* or whether it is subject to a fault. The available literature on attack detection and isolation, differently to that on cyber-security in the computer science sense, attempts to exploit knowledge of the system dynamics in order to verify whether communicated information is corrupted or not [15]. Most work in the secure control literature is based on centralized architectures [9], which in many cases may not be appropriate given the complexity of most cyber-physical systems. Hence, it is necessary to develop distributed methodologies. In the literature, few works, e.g. [10], [16], [17] point in this direction. In [10], [16] the authors develop a distributed methodology to detect attacks on the physical processes of subsystems, but it is assumed that the communication between detectors is secure. Furthermore, for design of the method in [17] knowledge of the model of the entire system is required. On the other hand, in the literature several contributions proposing distributed FDI techniques have been presented [18], [19], [20], [21], [22], [23]. In this work, we propose an attack detection and isolation methodology which is able to identify attacks present in the communication network supporting a consensus-based secondary control for current sharing in DCmGs. A novel estimator has been developed, based on Unknown Input Observers (UIOs) [24]. The proposed estimator is able to identify, with limited information of the overall DCmG's model, whether output measurements received from its neighbors are corrupted by a data-injection attack or not. Preliminary results have been presented in [25]. With respect to [25], local UIOs are exploited to estimate the state of neighboring DGUs. Detection thresholds are developed which ensure the absence of false alarms. To the best of the authors' knowledge, this is the first work designing an attack detection architecture for DC microgrids.

This work has been partially supported by European Union's Horizon 2020 research and innovation programme under grant agreement No 739551 (KIOS CoE). This work has also been conducted as part of: i) the research project *Stability and Control of Power Networks with Energy Storage* (STABLE-NET) which is funded by the RCUK Energy Programme (contract no: EP/L014343/1); ii) 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, P. Nahata, G. Ferrari-Trecate are with the Automatic Control Laboratory, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. Email: {mustafa.turan, pulkit.nahata, giancarlo.ferraritrecate}@epfl.ch

F. Boem is with the Department of Electronic and Electrical Engineering, University College London, UK. Email: francesca.l.boem@gmail.com

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

The rest of the paper is structured as follows. In Section II the models of the DC microgrid and of the attack are presented and the attack-detection problem is formulated. In Section III, we briefly recall the distributed estimation technique and detection strategy presented in [25], and its limitations for DC microgrids are higlighted. In Section IV, the monitoring architecture is sketched, unknown input observers are designed to estimate the state of the neighbors, hence a detection threshold is designed and attack detectability is analyzed. In order to validate the proposed technique, in Section V, simulation results are provided showing the effectiveness of the proposed technique.

Notation: In the paper, the operator $|\cdot|$ applied to a set determines its cardinality, while used with matrices or vectors it defines their component-by-component absolute value. The operator $\|\cdot\|$ is used to define the matrix norm. In general, in this paper inequalities are considered component-by-component.

II. Problem Formulation

A. Model of the DC Microgrid

Consider a DC microgrid made of N distributed generation units (DGUs), which are modeled as in [4] and are interconnected through power lines. In particular, [4] exploits the Quasi Stationary Line (QSL) approximation where interconnection power lines are supposed to be purely resistive. The entire DC microgrid modeled in this way may be seen as an undirected graph where the nodes represent the DGUs, the edges represent the power lines connecting the DGUs, and the weights of the edges are the resistances of the power lines. The voltage and current dynamics for DGU i are characterised as follows:

$$\begin{aligned} \frac{dV_i}{dt} &= \frac{1}{C_{ti}} I_{ti} + \sum_{j \in \mathcal{N}_i} \frac{1}{C_{ti} R_{ij}} (V_j - V_i) - \frac{1}{C_{ti}} I_{Li} \\ \frac{dI_{ti}}{dt} &= \frac{1}{L_{ti}} V_{ti} - \frac{R_{ti}}{L_{ti}} I_{ti} - \frac{1}{L_{ti}} V_i \end{aligned} , \quad (1)$$

where (V_{ti}, I_{Li}) are inputs to the DGU, (V_i, I_{ti}) are the states, $V_j \in \mathcal{N}_i$ are the interconnection terms between the DGUs, and $\mathcal{N}_i \subset \mathcal{N} \equiv \{1, \dots, N\}$ is the set containing neighboring DGUs, i.e. DGUs connected to DGU i through power lines. R_{ti}, C_{ti}, L_{ti} are electrical parameters of the RLC filter of DGU i. R_{ij} is the resistance of the power line connecting DGUs i and j.

We now briefly summarize the control methodology in [26], where a primary controller is designed to stabilize the voltage of the DGUs, and a secondary consensus-based controller is designed to achieve current sharing. The state-space model of DGU i, considering also state disturbances and measurement noise, is given by

$$\begin{aligned} \dot{x}_{[i]}(t) &= A_{ii} x_{[i]}(t) + B_i u_{[i]}(t) + G_i \alpha_{[i]}(t) \\ &\quad + M_i d_{[i]}(t) + \xi_{[i]}(t) + w_{[i]}(t) \\ y_{[i]}(t) &= C_i x_{[i]}(t) + \rho_{[i]}(t) \end{aligned} , \quad (2)$$

where $x_{[i]} = [V_i, I_{ti}, v_{[i]}]^\top$ is the local state. Note that $v_{[i]}$ is added to the state to enable the integrator action required by the primary control, and its dynamics are $\dot{v}_{[i]} = V_{ref,i} - V_i$, $V_{ref,i}$ being the reference for voltage V_i. Variables $u_{[i]}$ and $\alpha_{[i]}$ are the primary and secondary inputs of the DGU. The term $d_{[i]} = [I_{Li}, V_{ref,i}]^\top$ is the exogenous input, and $\xi_{[i]} = \sum_{j \in \mathcal{N}_i} A_{ij} x_{[j]}$ is a vector modeling the physical influence of neighboring DGUs. The vectors $w_{[i]}(t)$ and $\rho_{[i]}(t)$ model the unknown state disturbances and measurement noises, respectively. The following assumption is needed:

Assumption 1: Process noise $w_{[i]}(t)$ and measurement noise $\rho_{[i]}(t)$ are unknown vectors which are bounded for all t by some known bounds, i.e.

$$|w_{[i]}(t)| \leq \bar{w}_{[i]}, \quad |\rho_{[i]}(t)| \leq \bar{\rho}_{[i]}, \quad \forall t \geq 0\,. \quad (3)$$

The primary control input is $u_{[i]} = [V_{ti}] = K_i y_{[i]}$, where matrix K_i is designed following the methodology in [4]. The secondary consensus-based controller input $\alpha_{[i]}(t)$ is defined according to the following consensus protocol:

$$\dot{\alpha}_{[i]}(t) = -\sum_{j \in \mathcal{N}_i} [0\; k_I\; 0] \left(\frac{y_{[i]}(t)}{I^s_{ti}} - \frac{y^c_{[j,i]}(t)}{I^s_{tj}} \right), \quad (4)$$

where $I^s_{ti} > 0, \forall i \in \mathcal{N}$ are scaling factors appropriately defined, and k_I is the consensus weight common to all DGUs. The term $y^c_{[j,i]}(t)$ represents the output measurement of DGU j which is communicated to DGU i, and is defined in the next subsection. Note that the secondary consensus-based control in [26] requires that there be a communication network connecting the controllers of DGUs. In this paper, without loss of generality, we suppose that the topology of the graph representing the physical interconnections of the DGUs and that of the communication network are the same.

Matrices A_{ii}, B_i, M_i, A_{ij}, K_i, and C_i are defined as [3].

B. Attack Model

We now explain how the attack is modeled. We define the attack as a data injection attack on the communicated data between neighboring DGUs. In the considered scenario, it is assumed that an attacker is able to inject falsified data in the communication network linking neighboring DGUs. We model these attacks by defining the output measurement vector $y^c_{[i,j]}(t)$, communicated by DGU i to DGU j, as:

$$y^c_{[i,j]}(t) = y_{[i]}(t) + \beta(t - T_{a_{[i,j]}}) \phi_{i,j}(t), \quad (5)$$

where $y_{[i]}(t)$ is the unattacked measurement of state of the i-th DGU, as given in (2), and $\beta(t)$ is an activation function whose value is 0 for $t < 0$ and 1 for $t \geq 0$. $T_{a_{[i,j]}}$ denotes the time of occurrence of an attack on the communication line connecting node i to j. Note that neighboring DGUs share the full output measurement vector.

We suppose that the attack can be active on a subset $\widehat{\mathcal{N}}_i \subseteq \mathcal{N}_i$ of the communication lines between DGU i and its neighbors, whilst we assume that, for each DGU, its own measurements of its state are secure.

In the following, for notation simplicity, and without loss of generality, we assume that the attack is concurrent among all the attacked communication lines, i.e. $T_{a_{[j,i]}} = T_a$. The unknown attack function $\phi_{i,j}(t)$ is designed by the attacker and models the data which is injected. In the sequel, a detection technique to detect $\phi_{i,j}(t)$ is illustrated.

III. Distributed Estimation of Local State Variables

We design a distributed state estimator - similar to the one given in [25] - based on a full-order Luenberger-like observer and used by each DGU to estimate its local state. This requires local model information, as well as output measurements that are communicated from its neighbors and are subject to possible attacks.

The dynamics of the local state estimator is:

$$\begin{aligned}\dot{\hat{x}}_{[i]}(t) &= A_{ii}\hat{x}_{[i]}(t) + B_i u_{[i]}(t) + G_i\alpha_{[i]}(t) + M_i d_{[i]}(t) \\ &\quad + \hat{\xi}_{[i]}(t) + L_i\left(y_{[i]}(t) - \hat{y}_{[i]}(t)\right) \\ \hat{y}_{[i]}(t) &= C_i\hat{x}_{[i]}(t)\end{aligned}, \tag{6}$$

where $\hat{x}_{[i]}(t), \hat{y}_{[i]}(t)$ are the state and output estimates, and $\hat{\xi}_{[i]}(t) = \sum_{j\in\mathcal{N}_i} A_{ij}y^c_{[j,i]}(t)$ is the locally computed interconnection vector from communicated measurements. Matrix L_i is designed such that $A_{Li} = (A_{ii} - L_iC_i)$ is Hurwitz stable. We assume that the estimator has perfect knowledge of the exogenous input $d_{[i]}(t) = [I_{Li}, V_{ref,i}]^\top$. Indeed, I_{Li} can be measured, and $V_{ref,i}$ is a design parameter.

In order to determine whether the above distributed estimator allows for attack detection, we analyze the residual

$$r_{[i]}(t) = y_{[i]}(t) - \hat{y}_{[i]}(t)\,.$$

Using (2) and (6) and owing to the definition of C_i, for time $t < T_a$, i.e. before the onset of an attack, it is possible to rewrite the residual as $r_{[i]}(t) = \epsilon_{[i]}(t) + \rho_{[i]}(t)$, where $\epsilon_{[i]}(t) = x_{[i]}(t) - \hat{x}_{[i]}(t)$ is the state estimation error obeying the following dynamics:

$$\dot{\epsilon}_{[i]}(t) = A_{L_i}\epsilon_{[i]}(t) - \sum_{j\in\mathcal{N}_i} A_{ij}\rho_{[j]}(t) + w_{[i]}(t) - L_i\rho_{[i]}(t)\,. \tag{7}$$

Since A_{Li} are Hurwitz stable for all $i \in \mathcal{N}$, the dynamics in (7) are BIBO stable and independent from the primary, secondary, or exogenous inputs. Indeed, Assumption 1 implies the boundedness of $\epsilon_{[i]}(t)$. More specifically, let us compute the solution of (7):

$$\epsilon_{[i]}(t) = e^{A_{Li}t}\epsilon_{[i]}(0) + \int_0^t e^{A_{Li}(t-\tau)}\eta_{[i]}(\tau)d\tau, \tag{8}$$

where $\eta_{[i]}(t) = -\sum_{j\in\mathcal{N}_i} A_{ij}\rho_{[j]}(t) + w_{[i]}(t) - L_i\rho_{[i]}(t)$. As A_{Li} is Hurwitz stable by design, exist constants $\nu, \lambda > 0$ such that

$$\|e^{A_{Li}t}\| \leq \nu e^{-\lambda t}\,.$$

Because of Assumption 1 and the BIBO stability of (7), it is possible to use the triangle inequality to design a bound $\bar{\epsilon}_{[i]}(t)$ for the estimation error

$$\bar{\epsilon}_{[i]}(t) = \nu e^{-\lambda t}\bar{\epsilon}_{[i]}(0) + \int_0^t \nu e^{-\lambda(t-\tau)}\bar{\eta}_{[i]}d\tau, \tag{9}$$

where $\bar{\eta}_{[i]} = \sum_{j\in\mathcal{N}_i} |A_{ij}|\bar{\rho}_{[j]} + \bar{w}_{[i]} + |L_i|\bar{\rho}_{[i]}$. The use of the triangle inequality for the design of the bound implies that, in the absence of an attack, the inequality

$$|\epsilon_{[i]}(t)| \leq \bar{\epsilon}_{[i]}(t)$$

holds for suitable choice of $\bar{\epsilon}_{[i]}(0) \geq |\epsilon_{[i]}(0)|$. It is then possible to design an *attack-detection threshold* as follows:

$$\bar{r}_{[i]}(t) = \bar{\epsilon}_{[i]}(t) + \bar{\rho}_{[i]}, \tag{10}$$

such that

$$|r_{[i]}(t)| \leq \bar{r}_{[i]}(t) \tag{11}$$

holds in healthy conditions. Hence it is sufficient that a component of the residual crosses the corresponding threshold to state that an attack is present. A thorough analysis of the detectability properties of this method can be found in [25].

Limitations on attack detection

In the following we will analyze the limitations of this method when applied to attack-detection for DC microgrids.

Proposition 1: If the attack functions $\phi_{j,i}(t)$ in (5) take on the form

$$\phi_{j,i}(t) = \left[0\ \gamma_{j,i}(t)\ \theta_{j,i}(t)\right]^\top, \forall j \in \widehat{\mathcal{N}}_i, \forall t \geq T_a \tag{12}$$

where $\gamma_{j,i}(t)$ and $\theta_{j,i}(t)$ are arbitrary functions, then residual (10) is not affected by the attack.

Proof: The proofs of the propositions in this paper are omitted due to space constraints. ■

Remark 1: Note that undetectable attacks characterized in Proposition 1 turn out to be problematic. As $\gamma_{j,i}(t)$ in (12) influences the secondary control input in (4), an attacker seeking to alter the consensus-based control input may do so with limited knowledge of the DGU model, and without violating detection test (11).

Given the above remark, it is necessary to enhance the attack-detection scheme, as defined in the next section.

IV. Estimation of Neighboring DGU States

To overcome the significant issue pointed out in Remark 1, local validation of the transmitted information is necessary. To this end, let us define the structure of the proposed attack detection architecture used for each DGU. We design an unknown input observer for the estimation by DGU i of the state of each of its neighbors. Hence, a bank of $|\mathcal{N}_i|$ UIOs is designed for detection, one for each of the neighbors of DGU i. Once estimators are defined, bounds on the estimation errors are derived, which lead to the design of $|\mathcal{N}_i|$ detection thresholds, through which each of the residuals are tested. Following design, detectability is analyzed, and a sufficient condition for an attack to be stealthy is derived.

A. Unknown Input Observers

UIOs allow for state estimation even in the presence of unknown inputs (see, for instance, [24]). This is a particularly valuable feature in the case of DCmGs, as it allows DGU i to estimate the states of DGUs j, $j \in \mathcal{N}_i$, without requiring knowledge of their secondary input, exogenous inputs and of the states of their neighbors.

In order to design a UIO, we first recast the dynamics of DGU j in (2) to be consistent with the typical UIO structure [24]. Rewriting system dynamics in (2), one has:

$$\begin{aligned}\dot{x}_{[j]}(t) &= A_{Kj}x_{[j]}(t) + \bar{E}_j\bar{d}_{[j]}(t) + w_{[j]}(t) + B_jK_j\rho_{[j]}(t) \\ y_{[j]}(t) &= C_jx_{[j]}(t) + \rho_{[j]}(t)\end{aligned}, \tag{13}$$

where $A_{Kj} = A_{jj} + B_j K_j$, matrix $\bar{E}_j$ and vector $\bar{d}_{[j]}(t)$ contain the inputs to DGU j which are unknown to the UIO in DGU i. Specifically, they are defined as

$$\bar{E}_j = \begin{bmatrix} \frac{1}{C_{tj}} & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \bar{d}_{[j]} = \widehat{E}_j \hat{d}_{[j]}(t),$$

where

$$\widehat{E}_j = \begin{bmatrix} -1 & 0 & 0 & \frac{1}{R_{jk_1}} & 0 & 0 & \dots & \frac{1}{R_{jk_{|\mathcal{N}_j|}}} & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix},$$

$$\hat{d}_{[j]} = \left[d_{[j]}^\top(t), \alpha_{[j]}(t), x_{[k_1]}^\top(t), \dots, x_{[k_{|\mathcal{N}_j|}]}^\top(t) \right]^\top,$$

where $\{k_1, k_2, \dots, k_{|\mathcal{N}_j|}\}$ are all the elements of $\mathcal{N}_j$. Note that for UIO design [24], it is necessary that $\bar{E}_j$ be full column rank. The full order observer for estimating the state of DGU j in DGU i is:

$$\begin{aligned} \dot{z}_{[j,i]}(t) &= F_j z_{[j,i]}(t) + T_j B \bar{u}_{[j]}(t) + \widehat{K}_j y^c_{[j,i]}(t) \\ \tilde{x}_{[j,i]}(t) &= z_{[j,i]}(t) + H_j y^c_{[j,i]}(t) \end{aligned} \tag{14}$$

where $\bar{u}_{[j]}(t) = 0, \forall t$. The filter matrices are designed following [24] as:

$$(H_j C_j - I)\bar{E}_j = 0 \tag{15}$$

$$T_j = I - H_j C_j \tag{16}$$

$$F_j = T_j A_{Kj} - \widetilde{K}_j C_j \tag{17}$$

$$\bar{K}_j = F_j H_j \tag{18}$$

$$\widehat{K}_j = \widetilde{K}_j + \bar{K}_j \tag{19}$$

Proposition 2: If the UIO filter parameters are designed following (15)-(19), in the absence of attacks the estimation error $\widetilde{\epsilon}_{[j,i]}(t) = x_{[j]}(t) - \tilde{x}_{[j,i]}(t)$ behaves according to the following dynamics:

$$\dot{\widetilde{\epsilon}}_{[j,i]}(t) = F_j \widetilde{\epsilon}_{[j,i]}(t) + T_j \tilde{w}_{[j]}(t) - H_j \dot{\rho}_{[j]}(t) - \widetilde{K}_j \rho_{[j]}(t) \tag{20}$$

where $\tilde{w}_{[j]}(t) = w_{[j]}(t) + (B_j K_j)\rho_{[j]}(t)$.

In order to design a UIO which converges to the actual value of the state, it is necessary that the system dynamics satisfy the following conditions:

i. $rank(C_j \bar{E}_j) = rank(\bar{E}_j)$;

ii. the pair $(C_j, T_j A_{Kj})$ is detectable

Then, the following result can be proved.

Proposition 3: Given dynamics (13) and matrices defined as in Section II-A, conditions (i) and (ii) hold.

In order to design the UIO it is necessary to define H_j such that (15) holds, and $\widetilde{K}_j$ such that F_j in (17) is stable. Then the remaining matrices in (15)-(19) can be derived. Matrix H_j satisfies (15) as long as it is in the following form:

$$H_j = \begin{bmatrix} 1 & H_{j,1} & 0 \\ 0 & H_{j,2} & 0 \\ 0 & H_{j,3} & 1 \end{bmatrix} \tag{21}$$

in which $H_{j,1}, H_{j,2}, H_{j,3}$ can be arbitrarily assigned. Hence we compute matrices T_j, $\widetilde{K}_j$, $\bar{K}_j$, and $\widehat{K}_j$. Given the structure of $T_j = I - H_j$, the term $T_j A_{Kj}$ in (17) is

$$T_j A_{Kj} = \begin{bmatrix} T_{j,1}A_{Kj,21} & T_{j,1}A_{Kj,22} & T_{j,1}A_{Kj,23} \\ T_{j,2}A_{Kj,21} & T_{j,2}A_{Kj,22} & T_{j,2}A_{Kj,23} \\ T_{j,3}A_{Kj,21} & T_{j,3}A_{Kj,22} & T_{j,3}A_{Kj,23} \end{bmatrix}$$

where $T_{j,1} = -H_{j,1}$, $T_{j,2} = 1 - H_{j,2}$, and $T_{j,3} = -H_{j,3}$, and $A_{Kj,ab}$ is the (a,b)-th term of A_{Kj}. We now highlight the information flow which must occur between DGUs j and i such that the latter can estimate the state of the first:

1) At design time, DGU i requires that its neighbors communicate the second row of their closed-loop matrix
$$A_{Kj,2} = [A_{Kj,21}, A_{Kj,22}, A_{Kj,23}],$$
as well as the value of their bounds $\bar{w}_{[j]}$ and $\bar{\rho}_{[j]}$ in (3), needed for local computation of (25);
2) During online operations, DGU i requires transmission of measurement output $y^c_{[j,i]}(t)$.

We have designed the UIOs in DGU i to estimate the states of its neighbors. In the next subsection the residual $\widetilde{r}_{[j,i]}(t) = y^c_{[j,i]}(t) - C_j \tilde{x}_{[j,i]}(t)$ will be analyzed and a detection threshold will be defined.

B. Detection Thresholds

As in the case for the estimator in Section III, note that the residual in the absence of attacks, i.e. at time $t < T_a$ is:

$$\widetilde{r}_{[j,i]}(t) = C_j \widetilde{\epsilon}_{[j,i]}(t) + \rho_{[j]}(t). \tag{22}$$

Hence, by designing a bound on error $\widetilde{\epsilon}_{[j,i]}$, it is possible to exploit the triangle inequality to design a detection threshold. The solution to (20) is:

$$\begin{aligned} \widetilde{\epsilon}_{[j,i]}(t) &= e^{F_j t}\widetilde{\epsilon}_{[j,i]}(0) + \\ &+ \int_0^t e^{F_j(t-\tau)} \left[T_j \tilde{w}_{[j]}(\tau) - \widetilde{K}_j \rho_{[j]}(\tau) - H_j \dot{\rho}_{[j]}(\tau) \right] d\tau. \end{aligned} \tag{23}$$

Using integration by parts, to remove dependence from $\dot{\rho}_{[j]}$:

$$\begin{aligned} \widetilde{\epsilon}_{[j,i]}(t) &= e^{F_j t}\left[\widetilde{\epsilon}_{[j,i]}(0) + H_j \rho_{[j]}(0)\right] - H_j \rho_{[j]}(t) + \\ &+ \int_0^t e^{F_j(t-\tau)} \left[T_j w_{[j]}(\tau) + \left(T_j B_j K_j - \widehat{K}_j \right) \rho_{[j]}(\tau) \right] d\tau. \end{aligned} \tag{24}$$

Note that, because F_j is Hurwitz stable by design, it is possible to define constants $\kappa, \mu > 0$ such that

$$\|e^{F_j t}\| \leq \kappa e^{-\mu t}, \forall t \geq 0.$$

Therefore a time-varying bound $\bar{\widetilde{\epsilon}}_{[j,i]}(t)$ is designed:

$$\begin{aligned} \bar{\widetilde{\epsilon}}_{[j,i]}(t) &= \kappa e^{-\mu t}\left[\bar{\widetilde{\epsilon}}_{[j,i]}(0) + |H_j|\bar{\rho}_{[j]}\right] + |H_j|\bar{\rho}_{[j]} \\ &+ \int_0^t \kappa e^{-\mu(t-\tau)} \left[|T_j|\, \bar{w}_{[j]} + \left| T_j B_j K_j - \widehat{K}_j \right| \bar{\rho}_{[j]} \right] d\tau, \end{aligned} \tag{25}$$

which, for suitably defined $\bar{\widetilde{\epsilon}}_{[j,i]}(0) \geq |\widetilde{\epsilon}_{[j,i]}(0)|$, guarantees that, in the absence of an attack,

$$|\widetilde{\epsilon}_{[j,i]}(t)| \leq \bar{\widetilde{\epsilon}}_{[j,i]}(t) \tag{26}$$

holds. Finally, it is possible to design the detection threshold:

$$\bar{\widetilde{r}}_{[j,i]}(t) = \bar{\widetilde{\epsilon}}_{[j,i]}(t) + \bar{\rho}_{[j]} \tag{27}$$

for which in the absence of an attack the following is satisfied

$$|\widetilde{r}_{[j,i]}(t)| \leq \bar{\widetilde{r}}_{[j,i]}(t). \tag{28}$$

This inequality is then used to detect the presence of attacks. In fact, it is sufficient for it to be violated for at least one component of residual $\widetilde{r}_{[j,i]}(t)$ at some time T_d for an attack to be detected and the compromised communication link isolated.

C. Detectability Analysis

We now analyze the detectability properties of the proposed method. We start by analyzing the effect of the attack modeled as in (5) on the estimation error dynamics in (20). In this case $\widetilde{\epsilon}_{[j,i]}(t)$ is given by

$$\begin{aligned}\widetilde{\epsilon}_{[j,i]}(t) &= e^{F_j(t-T_a)}\left[\widetilde{\epsilon}_{[j,i]}(T_a) + H_j\left(\rho_{[j]}(T_a) + \phi_{j,i}(T_a)\right)\right] \\ &- H_j\left(\rho_{[j]}(t) + \phi_{j,i}(t)\right) + \int_{T_a}^{t} e^{F_j(t-\tau)} T_j w_{[j]}(\tau) d\tau + \\ &+ \int_{T_a}^{t} e^{F_j(t-\tau)}\left[\left(T_j B_j K_j - \widehat{K}_j\right)\rho_{[j]}(\tau) - \widehat{K}_j \phi_{j,i}(\tau)\right] d\tau,\end{aligned} \tag{29}$$

where, similarly to (24) we used integration by parts. The residual is $\widetilde{r}_{[j,i]}(t) = \widetilde{\epsilon}_{[j,i]}(t) + \rho_{[j]}(t) + \phi_{j,i}(t)$. We would like to find a condition for which at time T_d the inequality $|\widetilde{r}_{[j,i]}(T_d)| > \widetilde{\bar{r}}_{[j,i]}(T_d)$ is guaranteed to be satisfied for at least one component of the residual.

Proposition 4: If exists time $t = T_d$ such that

$$\begin{aligned}\Big| e^{F_j(t-T_a)} H_j \phi_{j,i}(T_a) + T_j \phi_{j,i}(T_d) + \\ - \int_{T_a}^{T_d} e^{F_j(t-\tau)} \widehat{K}_j \phi_{j,i}(\tau) d\tau \Big| > 2\widetilde{\bar{r}}_{[j,i]}(T_d)\end{aligned} \tag{30}$$

is fulfilled for at least one component of $\widetilde{\bar{r}}_{[j,i]}(t)$, then attack detection is guaranteed, for any value of $\rho_{[j]}(t)$ and $w_{[j]}(t)$.

D. Stealthy Attacks

We now define a condition on $\phi_{j,i}(t)$ causing an attack to be undetectable by the UIO-based detection strategy.

Proposition 5: If the attack is designed in such a way that

$$\begin{aligned}\Big| e^{F_j(t-T_a)} \phi_{j,i}(T_a) + T_j \phi_{j,i}(t) + \\ - \int_{T_a}^{t} e^{F_j(t-\tau)} \widehat{K}_j \phi_{j,i}(\tau) d\tau \Big| = 0\end{aligned} \tag{31}$$

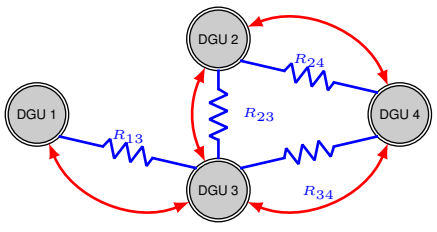

Fig. 1: Graph of DCmG considered in simulation. The blue lines represent the power lines connecting the DGUs, and the red arrows represent the communication graph. Note that communication links are bidirectional, although links may be attacked in one direction without the other being affected.

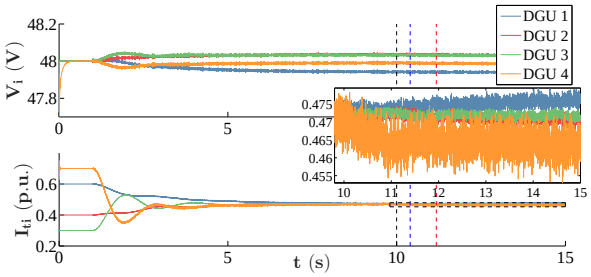

Fig. 2: State evolution of voltages and currents of the DGUs. Vertical lines show initial time of attack, in black, as well as detection and isolation of attack on communications from DGUs 2 and 3, in blue and red, respectively. In the highlighted window, a detail of trajectories is presented, showing that consensus is not achieved in presence of attack.

is satisfied for all $t \geq T_a$, then its detection by the UIO-based detection scheme is not possible.

Attacks satisfying Proposition 5 are part of a class of attacks which are undetectable by the proposed methodology, termed *stealthy*. Note that, to generate a stealthy attack as in (31) the attacker must have a lot of information regarding the filters of the UIO. The complete characterization of the class of stealthy attacks is out of scope of this preliminary paper, and will be the subject of further research.

V. Simulation Results

The proposed attack detection architecture is evaluated by conducting simulations on MATLAB software. In Figure 1 we present the topology of the considered DC microgrid. The same parameters as in [26] are used in the simulation to define DGU matrices and controllers. In the following, unless otherwise stated, currents are measured in A and volts in V. Process and measurement noises are modeled as random processes verifying Assumption 1, for all $t \geq 0$, where: $\bar{w}_{[i]} = [0.1, 0.1, 0.01]^\top$, and $\bar{\rho}_{[i]} = [0.01, 0.01, 0]^\top$, and where each scalar process is uncorrelated, and generated from a uniform distribution on the corresponding interval. We consider that, before time $t = 1.5s$, the DGUs are disconnected. Then, after time $T_a = 10s$, the following attack function affects transmitted measurements $y^c_{[2,4]}(t)$ and $y^c_{[3,4]}(t)$:

$$\phi_{j,4}(t) = \begin{bmatrix} 0 & 0.1 & 0 \end{bmatrix}^T, \forall j \in \mathcal{N}_4 = \{2,3\}, \forall t \geq T_a. \tag{32}$$

Function (32) models an attack on the measurement of the current communicated to compute secondary control of DGU 4, which is equipped with a detector with two UIOs computing estimates of the state of DGUs neighboring DGU 4. The filter matrices for both the unknown-input observers are designed as follows. For $j \in \{2,3\}$, the second column of H_j in (21) is $[0.1, 0.9, 0.1]^\top$, matrices T_j satisfy (17), $\widetilde{K}_j$ are such that eigenvalues of $F_j = \{-1, -1.5, -2\}$, and $\bar{K}_j$, $\widehat{K}_j$ are computed as in (18)-(19).

In Figure 2 we show the state evolution of V_i and I_{ti} for all the DGUs of the DCmG. The scaling factors in (4) are taken to be the current ratings of each DGU, i.e. $I^s_{ti} = 10A, i \in \{1,2,3\}$, and $I^s_{t4} = 5A$, therefore currents are shown in $p.u.$ Furthermore, this implies that consensus

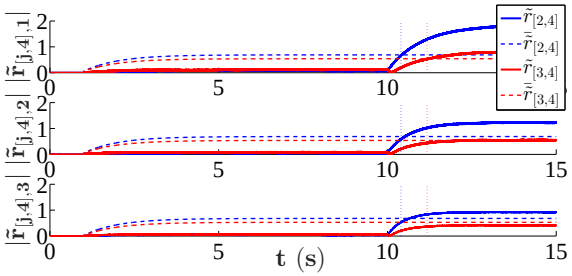

Fig. 3: Component-by-component comparison of residuals $\widetilde{r}_{[j,4]}(t)$ (22) with their corresponding thresholds $\bar{\widetilde{r}}_{[2,4]}(t)$ (27), in continuous and dotted lines, respectively. Those corresponding to estimate of DGU 2 are in blue, whilst those for estimate of DGU 3 are in red. Detection occurs when the residual of the first component of the state crosses the corresponding threshold (vertical lines).

is achieved when currents are the same in Figure 2. Note that, with the defined ratings, the attack in (32) corresponds to $[0, 0.02, 0]^\top\ p.u.$ Together with the state evolution of the systems, in Figure 2 we present the time instance when attack begins, $t = T_a$ (black dashed line), as well as the times at which the residuals $\widetilde{r}_{[2,4]}(t)$ and $\widetilde{r}_{[3,4]}(t)$ violate their respective thresholds (in blue and red, respectively). Additionally, in Figure 2 an enlarged portion of the evolution of the current is also shown, to emphasize that consensus is not achieved under attack.

Finally, we show the residuals from the UIOs which estimate the neighboring states, in Figure 3. We see that before time T_a, the estimates of the UIOs lie within the bounds, whilst after the onset of the attack, the residuals increase until they cross the threshold, thus detection is achieved at times $T_{d,2} = 10.4s$ and $T_{d,3} = 11.18s$. Note that detection time is also isolation time, at which the detector of DGU 4 is able to identify the communication line over which the information is compromised.

VI. Concluding Remarks

In this paper, we have introduced a distributed attack detection and isolation methodology for DC microgrids. Each DGU locally computes estimates of its neighbors' states through UIOs. In the paper, local detection thresholds are designed, and detectability analysis is performed. Furthermore, a class of stealthy attacks is identified. As future work, we aim to conduct further analysis in the design of the matrices of the UIOs, in order to minimize their effect on the size of the detection threshold. Moreover, we aim to study the possibility of developing a reconfiguration strategy to restore consensus in the event of an attack.

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, Sept 2017.

[2] J. M. Guerrero, J. C. Vásquez, and R. Teodorescu, "Hierarchical control of droop-controlled DC and AC microgrids: a general approach towards standardization," in *2009 35th Annual Conference of IEEE Industrial Electronics*, 2009, pp. 4305–4310.

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

[4] M. Tucci, S. Riverso, and G. Ferrari-Trecate, "Line-independent plug-and-play controllers for voltage stabilization inDCmicrogrids," *IEEE Transactions on Control Systems Technology*, 2017, to appear.

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

[6] C. De Persis, E. Weitenberg, and F. Dörfler, "A power consensus algorithm for DC microgrids," *Automatica*, 2016, submitted.

[7] M. Tucci, L. Meng, J. M. Guerrero, and G. Ferrari-Trecate, "Plug-and-play control and consensus algorithms for current sharing in DC microgrids," in *Preprints of the 20th IFAC world congress*. Toulouse France, 2017, pp. 12 951–12 956.

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

[9] 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.

[10] 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.

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

[12] P. Danzi, C. Stefanovic, L. Meng, J. M. Guerrero, and P. Popovski, "On the impact of wireless jamming on the distributed secondary microgrid control," in *2016 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2016, pp. 1–6.

[13] F. Pasqualetti, "Secure control systems: A control-theoretic approach to cyber-physical security," Ph.D. dissertation, University California, Santa Barbara, 2012.

[14] S. Weerakkody, X. Liu, S. H. Son, and B. Sinopoli, "A graph-theoretic characterization of perfect attackability for secure design of distributed control systems," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 60–70, 2017.

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

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

[17] A. Teixeira, H. Sandberg, and K. H. Johansson, "Networked control systems under cyber attacks with applications to power networks," in *American Control Conference, 2010*, 2010, pp. 3690–3696.

[18] F. Arrichiello, A. Marino, and F. Pierri, "Observer-based decentralized fault detection and isolation strategy for networked multirobot systems," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 4, pp. 1465–1476, 2015.

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

[20] F. Boem, L. Sabattini, and C. Secchi, "Decentralized fault diagnosis for heterogeneous multi-agent systems," in *2016 3rd Conference on Control and Fault-Tolerant Systems (SysTol)*, 2016, pp. 771–776.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

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

[26] 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.