Reinforcement Learning-Based Method to Exploit Vulnerabilities of False Data Injection Attack Detectors in Modular Multilevel Converters

Claudio Burgos-Mellado, *Member, IEEE*, Claudio Zuñiga-Bauerle, Diego Muñoz-Carpintero, *Member, IEEE*, Yeiner Arias-Esquivel, *Student, IEEE*, Roberto Cárdenas-Dobson, *Senior, IEEE*, Tomislav Dragičević, *Senior, IEEE*, Felipe Donoso, *Graduate Student Member, IEEE*, Alan Watson, *Senior, IEEE*

***Abstract*—Implementing control schemes for modular multilevel converters (M2Cs) involves both a cyber and a physical level, leading to a cyber-physical system (CPS). At the cyber level, a communication network enables the data exchange between sensors, control platforms, and monitoring systems. Meanwhile, at the physical level, the semiconductor devices that comprise the M2C are switched ON/OFF by the control system. In this context, almost all published works in this research area assume that the CPS always reports correct information. However, this may not be the case when the M2C is affected by cyber-attacks, such as the one named false data injection attack (FDIA), where the data seen by the control system is corrupted through illegitimate data intrusion into the CPS. To deal with this situation, FDIA detectors for the M2C are recently starting to be studied, where the goal is to detect and mitigate the attacks and the attacked sub-modules. This paper proposes a reinforcement learning (RL)-based method to uncover the deficiencies of existing FDIAs detectors used for M2C applications. The proposed method auto-generates complex attack sequences able to bypass FDIA detectors. Therefore, it points out the weaknesses of current detectors: This valuable information can be used later to improve the performance of the detectors, establishing more reliable cybersecurity solutions for M2Cs. The RL environment is developed in Matlab/Simulink augmented by PLECS/blockset, and it is made available to researchers on a website to motivate future research efforts in this area. Hardware-in-the-loop (HIL) studies verify the proposal's effectiveness.**

***Index Terms*—Modular Multilevel Converter, Distributed Control, False Data Injection Attack, Reinforcement Learning.**

This work was supported in part by "Agencia Nacional Investigacion y Desarrollo" (ANID) under grants: ANID/FONDECYT de Iniciación/11220989 and ANID/FONDECYT de Iniciación/11221230. The support of the Basal project FB0008 and Fondecyt 1221392 is also acknowledged.

C. Burgos-Mellado and Diego Muñoz-Carpintero are with the Institute of Engineering Sciences, Universidad de O'Higgins, Libertador Bernardo O'Higgins 611, Rancagua 2841959, Chile (emails: claudio.burgos@uoh.cl; diego.munoz@uoh.cl)

C. Zuñiga and R. Cárdenas-Dobson are with the Department of Electrical Engineering, Faculty of Physical and Mathematical Sciences, University of Chile, Santiago 8370451, Chile (emails: claudio.zuniga.b@ug.uchile.cl, rcd@ieee.org)

Y. Arias-Esquivel is with the Department of Mechatronic Engineering, Instituto Tecnológico de Costa Rica, Cartago 30101, Costa Rica (email: yarias@tec.ac.cr).

T. Dragičević is with the Electrical Engineering Department of Technical University of Denmark, Copenhagen, Denmark (email: tomdr@elektro.dtu.dk)

F. Donoso, is with the Research and Development Department, Siemens Drives Technology Unit, Congleton CW12 1PH, U.K (email: felipe.donoso-merlet@siemens.com)

A. Watson is with the Power Electronics, Machines and Control Group, University of Nottingham, Nottingham NG72RD, U.K. (email: alan.watson@nottingham.ac.uk)

I. Introduction

MODULAR Multilevel Converters (M2Cs) are a prominent solution for high-voltage direct current (HVDC) transmission systems that enable electric power transfer over a long distance [1]. Such converters have been used in several commercial projects worldwide (e.g Trans Bay Cable [2], Dolwin2, etc.), and one of its main advantages is that M2Cs can withstand high voltages by connecting hundred of submodules (SMs) in series. Traditionally, this power converter has been controlled using a centralised control architecture, where a central controller is in charge of processing all the information required for implementing the whole control system, running the control algorithm, and sending back the control actions to the semiconductor devices that compose the M2C submodules (SMs). This control architecture limits the modularity, flexibility and expandability of the M2C in terms of software development, especially for the case where the converter has a considerable number of SMs, i.e., for HVDC applications. Additionally, the centralised control approach makes the system prone to a single point of failure. To overcome some of these issues, distributed control schemes have been proposed to control the M2C (see [3]–[9]). In this approach, local controllers (LCs), physically implemented in each M2C SM, perform low-level control tasks, whereas high-level control tasks are undertaken by a central controller (CC). Thereby, the computational burden on the CC is reduced, and a more reliable and modular system with fewer signal wires is obtained.

Both centralised and distributed control architectures are implemented in a cyber-physical system (*CPS*), composed of sensors, communication links, etc., where interactions between the physical elements and computational processes occur [10]. The *CPS* is vulnerable to malicious cyber-attacks, which could potentially degrade the converter operation, lead to sub-optimal or even unstable operation, and induce protection system tripping in severe cases. Examples of common types of cyber-attacks are [11]: (i) false data injection attack (FDIA), (ii) replay attack, and (iii) denial of service (DoS) attack. Note that cyber-attack issues are intensively investigated in other electrical systems such as microgrids [11]–[13], smart grids [14], [15], modern power systems [16], [17], and electric ve-

hicle charging infrastructure [18], [19]. However, for modular multilevel converters, this area has been scarcely explored.

In the field of M2C applications, so far only [10], [20] have addressed cybersecurity issues in M2Cs; in both cases, false data injection attacks (FDIAs) were considered. In [20], it is shown, via simulations, that an FDIA in the M2C sensors can affect the stability of the centralised control system. On the other hand, in [10], an FDIA detector, based on the Kalman filter, is proposed with the M2C being controlled using a distributed control scheme: the proposal [10] is experimentally validated. In references [10], [20] a simple FDIA is considered: it is assumed that the attack sequence $V_i^a(k)$ in (1) is performed by introducing a false step variation in the voltage of the ith module seen by the M2C control system (known hereon as a step FDIA). However, their protection performance against more sophisticated FDIAs is not guaranteed; for instance, for the case where the attack sequence $V_i^a(k)$ in (1), is another, more elaborated time-variant attack sequence.

Based on the above discussion, it is essential to have a methodology that can explore the vulnerabilities of existing detection methods (as well as those developed in the future), identifying their weaknesses to improve the robustness of these detection algorithms. In this context, the vulnerabilities of FDIA detectors must be understood as those FDIAs able to deceive the detection algorithm, i.e. those which are not pinpointed by the detection algorithms. With this information, the performance of existing detectors can be improved by incorporating proper modifications to detect FDIAs not perceived initially. To the best of the author's knowledge, this has not been previously investigated for applications related to M2Cs.

To fulfil the identified research gap, this paper proposes a general reinforcement learning (RL)-based method to obtain an FDIA attacker (a neural network called the agent) able to discover the vulnerabilities of FDIA detectors and provide valuable information about attack sequences able to deceive the aforementioned FDIA detectors: Fig. 1 shows the general scheme of the RL-based method presented in this work. Considering the large number of floating capacitors used by the converter topology, this paper focuses on FDIAs targeting the capacitor voltage measurements of M2C. However, the proposed methods could be modified to study other cyber-attack targets, for instance, the M2C arm currents. The Twin Delayed Deep Deterministic Policy Grandient (TD3) variant of RL is considered in this work for training the agent (FDIA attacker) which explores FDIAs of continuous values, thus allowing for precise identification of the potential vulnerable spots. In this paper, the vulnerable spots correspond to the FDIAs that deceive the studied FDIA detector and, therefore, are not detected. In particular, the proposal performance is evaluated considering the detector recently proposed in [10]. To this end, the proposed RL-scheme is used to generate an agent able to generate attack sequences [see $V_i^a(k)$ in (1))] which are able to fool the detector of [10], causing detrimental effects to the operation of an M2C. The proposal's performance is validated using hardware in the loop (HIL) studies. The main contributions of this paper are:

1) This paper proposes a novel RL-based method for automatically discovering the vulnerabilities of existing FDIA detectors utilised in applications related to modular multilevel converters. More precisely, RL is used to train an agent that, in each discrete time, outputs an FDIA value that can bypass the analyzed FDIA detector when applied (see Fig. 1). Notice that the ability of the FDIA attacker to achieve its goal will depend on the given training (namely FDIA detectors, operation regimes, etc.). Thus, the proposal provides valuable information for effectively revealing the vulnerability of the studied FDIA detector, thereby potentially contributing to achieving a more robust and efficient FDIA detector for M2C applications.
2) The proposed vulnerability discovery and exploitation scheme is implemented using TD3. This enables a continuous space action for a more fine exploration of the FDIA to decide on the studied FDIA detector. Notice that, as discussed in Section IV-A, TD3 is considered the state of the art on RL methods with continuous action space.
The proposed agent takes as inputs the information that may be obtained by the attacker in the ith M2C SM and outputs corrupted voltage measurement of that SM to be used by LCs. Note that the proposed RL environment facilitates effective integration into existing FDIA detectors used for M2Cs.
3) This paper provides the foundation for establishing more reliable cybersecurity solutions for M2Cs since, so far, there are few works on this topic, i.e. only [10], [20] have addressed cybersecurity issues in modular multilevel topologies. The proposed RL method source code is available in [21] to promote future research in this area.

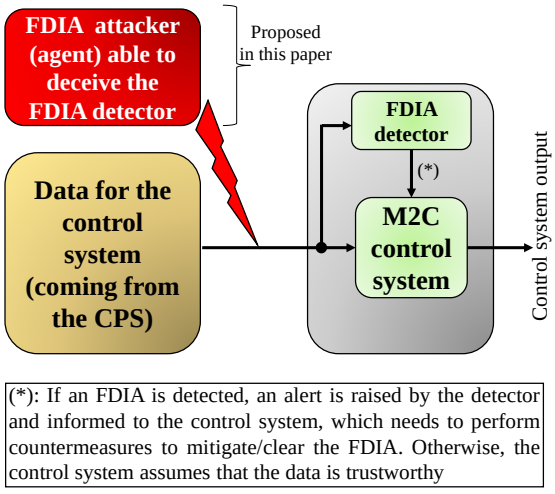


Fig. 1. Graphical representation of the proposal presented in this paper.

The rest of this paper is organised as follows: Section II presents the definition of the FDIA and its potential impacts on M2Cs. Section III introduces and discusses the control system used for managing the M2C. The proposed RL-based method for generating an FDIA attacker able to deceive FDIA detectors for M2C is presented in Section IV. In Section V,

the HIL validation of the proposal is provided. Finally, the conclusions are presented in Section VI.

II. Definitions and Potential Impacts of FDIA in the M2C

This paper considers that the voltage measurements in the SM capacitors of the M2C are prone to FDIAs. In this scenario, an FDIA on the voltage measurement targeting the ith LC is modelled as (1). The term $V_i^f(k)$ is the attacked voltage seen by the M2C control system, $V_i(k)$ is the real voltage, and $V_i^a(k)$ is the attack sequence. In (1), κ=1 denotes the presence of an attack sequence $V_i^a(k)$ in $V_i(k)$, otherwise κ=0. It must be pointed out that in (1), it is considered that the FDIA is affecting the ith SM of the M2C. (For the rest of the SMs, the model is analogous)

$$Sensor \quad attack: \quad V_i^f(k) = V_i(k) + \kappa V_i^a(k) \tag{1}$$

As stated in [10], the FDIA (1) has an enormous impact on individual voltage balancing control. In this case, for the M2C control system, all looks normal; however, the individual SM capacitor voltages are not adequately regulated, producing high and/or low voltages [depending on the sign of the attack sequence $V_i^a(k)$ in (1)] in the attacked SMs, and eventually, tripping of the M2C protection systems. For this reason, methods for detecting FDIAs and countermeasures to deal with them need to be investigated further.

It must be pointed out that in this work, FDIAs targeting voltage sensor measurements are discussed. This is due to the fact that the capacitor voltage measurements are transmitted to the control system using the cyber-physical network.

Finally, it must be highlighted the development of FDIA detectors is an active research field in electrical systems such as microgrids, smart grids, and modern power systems. These methods are based mainly on the use of observer-based estimation methods (state-space-based observers) such as Kalman filter [14], Particle filter [22], Luenberger observer [23], and adaptive sliding-mode observer [24]. Also, recently the use of artificial intelligence (AI)-based observers [25] has increased attention from researchers. For instance, methods based on neural networks [12], support vector machine [26], and deep learning techniques [27] have been published.

A. Potential Impacts of FDIAs targeting the M2C

As stated in the introductory section, cybersecurity issues in M2C have been scarcely explored in the literature. So far, the papers [10], [20] are the only ones that have studied this topic. Both articles studied the effects of FDIAs on the M2C operation. It is worth to be aware that [20] studied the impact of the step FDIA on the M2C operation when a centralised control scheme controls this converter, whereas in [10], the same study was performed but considering that the M2C is controlled using a distributed control scheme based on the consensus theory [28]. From these references, the following potential impacts of step FDIAs on the operation of the M2C were found:

Instability issues: In [20] a small-signal model of a centralised-controlled M2C (modelled in the dq reference frame), is developed to quantify cyber-attack effects. From that work, it was concluded that cyber-attacks could move the eigenvalues of the modelled system to the right-half plane, leading the system to instability.

Power quality issues: In [10], it is shown that FDIAs could generate unbalanced currents on the M2C AC side when the converter is regulated by a consensus-based distributed control scheme.

Operability issues: From the two previous points discussed, it can be concluded that FDIAs eventually can lead to a shutdown of the M2C from the rest of the system due to the tripping of the protection system. This can be produced by either the FDIA leading the control system to instability or because the cyberattack discreetly penetrated the control system and artificially increased and/or decreased one or more measurements, producing the activation of the aforementioned protection system associated with those variables for overvalues and/or undervalues.

References [10], [20] showed that cyber-attacks affecting M2Cs could produce severe problems in the normal operation of this converter. Thus, they need to be investigated further. In this regard, it is paramount to mention that the effects of cyber-attacks, especially the one studied here (FDIA), will depend on the characteristics of the M2C control system. For instance, the impact of FDIA in a given control scheme might be less than the effects of the same FDIA targeting an M2C with a different control scheme. For this reason, more research is required to develop robust control schemes for the M2C against FDIAs.

B. Plausibility of FDIAs targeting M2Cs

As stated above, this paper considers the M2C in the context of HVDC systems. This topology has been used as a prominent solution worldwide to develop commercial HVDC projects: For instance, Siemens implemented the 400MW Transbay Cable project HVDC project using M2C technology [29]. The Southern Grid's Nan'ao link, the first multi-terminal HVDC project, was implemented using M2Cs [30]. The Zhoushan HVDC project is the first five terminal M2C-HVDC system in the world [31]. In these systems, the M2C is placed in a substation along with other electrical systems, as illustrated in Fig. 2. In addition, all the systems in the substation are monitored and coordinated by a supervisory control and data acquisition (SCADA) system. The SCADA system, among other tasks, shares information between the substation and the control centre. This information is shared through a communication network between the control centre and the substation (see Fig. 2). Thus, a cyber-attack on the SCADA system might allow the hacker to access all the substation components, particularly the M2C control system. In this situation, the attacker might look to maximising the effects of the cyber-attack. This could be measured by the period the system stops operating due to the cyber-attack. To achieve this aim, from the attacker's point of view, FDIAs targeting some of the hundreds of SMs that compose the M2C could be more effective (more challenging to localise and remediate) than an FDIA targeting other electrical systems of the HVDC station. This situation shows that cyber-attacks could be credible threats to M2Cs. Although this is a

recent topic, several actors, such as researchers, governments and universities, have realised the detrimental impact of cyber-attacks on M2C-based HVDC systems [20], [32]–[34]. By disabling HVDC transmission lines, it is possible to provoke a shutdown of the whole transmission system, producing a blackout. For instance, in 2015, the Ukrainian grid became unstable, and later it experimented a blackout after a false data injection attack compromised measurements from electricity grid sensors [35]. In general, several reasons support more research regarding cyber-attacks against M2C projects, such as the increment of M2C-based solutions for HVDC applications [29]–[31], the steady increase in cloud-based solutions in power electronics [6], [36], [37], and a rise of cyber-attacks against critical infrastructure [20], [32]–[35].

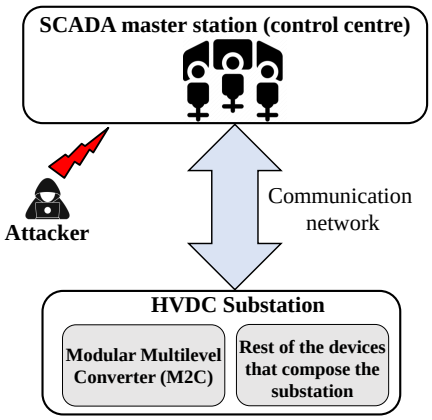


Fig. 2. Implementation of an M2C in a substation: All the devices of the substation are monitored by a SCADA system located in the control centre.

Additionally, commercial, real-life implementations of M2C have used some of the following communication protocols: Modbus [38], [39], Ethercat [40]–[44] and Profinet [9], [45], [46]. Note that these communication protocols are been widely used in SCADA systems [47]–[51]. In this context, it must be pointed out that these protocols have already shown vulnerabilities to cyber-attacks in SCADA systems, evidenced by numerous attacks targeting real-life applications [35], [52], [53].

III. Distributed control architecture for driving M2Cs

As discussed in the introduction section, the traditional control approach for driving M2C is the centralised control architecture where a central controller is in charge of implementing the whole control system. This approach is a good solution for M2C composed of a few SMs. However, it can have many issues for M2C in HVDC applications, where the number of SM is much larger. In this case, the main disadvantage of the centralised control system is that the central controller needs extensive processing capability, multiple analogue inputs, multiple digital outputs and communication channels for the switching signals, increasing the complexity and cost of the system.

A promising alternative to centralised control of M2C, is the utilisation of distributed control schemes for managing this power converter [3]–[9]. In this control architecture, a central controller performs high-level tasks (such as current control and total energy control). And local controllers (LCs) placed on the SMs, are in charge of low-level tasks such as the capacitor voltage balancing control and PWM generation. By doing this, the processing burden is shared among the LCs. Distributed control schemes have already been proposed for several modular multilevel cascaded converters, such as the cascaded H-bridge (CHB) multilevel converter [54]–[58], the M2C [3]–[9], and the modular multilevel matrix converter (M3C) [59], [60]. Note that distributed control schemes for modular multilevel cascaded converters have been studied recently and correspond to an active research field. The main distributed approaches reported in the literature, with a focus on the M2C, are discussed in the paragraphs below.

References [9], [61] proposed a hierarchical control system where a central controller performs the average, circulating, and grid current control; and the capacitor voltage balancing control and the modulation stages are performed in the local controllers of the SMs. This approach requires communication links among all the SMs that belong to the same arm. A similar approach is proposed in [3], [4], [43]. It reduces the computational burden (compared with [9], [61]) by regulating the capacitor voltage of groups of SMs in the same arm, and a controller regulates the voltage balancing among different groups. In [45], [62], [63] distributed control schemes based on a master/slave approach are proposed. For example, in [45], the master controller deals with high-level tasks such as current control, circulating current control and power controls. In contrast, slave controllers (placed on the SMs) perform low-level tasks such as capacitor voltage balance control and PWM generation. In [6], [64], [65], a hybrid control scheme for an M2C is proposed: a central controller regulates the output current, whereas local controllers regulate the SM capacitor voltages based on a distributed control scheme and are in charge of the PWM generation.

Due to its effectiveness in microgrid (MG) applications [66], distributed control schemes based on the consensus theory [28], have been recently proposed for controlling modular multilevel cascaded topologies, [5], [38], [39], [57], [58]. In particular, consensus-based distributed control schemes for balancing the capacitor voltages of the SMs of multilevel converters are proposed in [5], [57]; and for balancing the state of charge (SOC) of battery energy storage systems (BESS) based on modular multilevel cascaded topologies are proposed in [38], [39], [58], [67]. In those references, consensus-based distributed control schemes show promising results, and it demonstrated that they do not need communication with the central controller for running their control algorithms. Additionally, consensus-based distributed control schemes have excellent performance when events such as SM failures, time delays in the communication network, and failures in the communication links are produced. Based on these characteristics, distributed control schemes based on the consensus theory seem to be a promising control architecture to manage M2C with a high number of SMs, having recently published papers working with this approach [5], [10], [38], [39], [57], [58], [67].

A. Consensus theory for controlling M2Cs

Fig. 3 displays the control system considered in this work for managing the M2C. A central controller (CC) system performs high control tasks (the output current, the circulating currents, the arm balance, the total energy and the DC-port voltage), whereas local controllers work in a consensus-based distributed control architecture for individual voltage balancing. Also, in the LCs, the PWM generation is realised. The CC is implemented in the $\Sigma\Delta\alpha\beta 0$ reference frame discussed in [68], whereas the capacitor voltage balancing control is based on the consensus theory. As the latter approach is relatively new for M2C applications, it is discussed in more detail in this section.

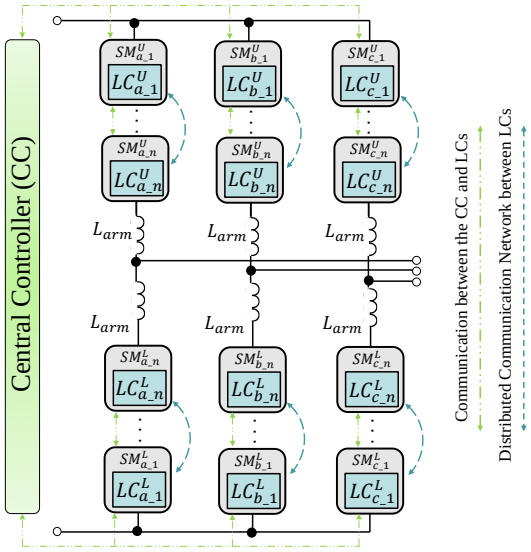


Fig. 3. Scheme of the control architecture used in this work for driving the M2C.

To discuss the use of the consensus theory for controlling the capacitors' voltage balancing, let us consider the upper arm of phase a in the M2C of Fig. 3, as shown in Fig. 4 (for the rest of the arms, the procedure is analogous). The distributed communication network shown in Fig. 4 corresponds to an undirected cyber graph $\mathbb{G} = (\mathfrak{N}, \xi, B)$ among the SMs $\mathfrak{N} = \{1, ..., n\}$, where ξ is the set of communication links and B is a non-negative $n \times n$ weighted adjacency matrix [28], and n is the number of SMs in the arm. The elements of B are $b_{ij} = b_{ji} \geq 0$, with $b_{ij} \geq 0$ if and only if $\{i, j\} \in \xi$. In other words, the terms b_{ij} of a non-negative B represent information flow among LCs, and they are determined as (2).

$$b_{ij} = \begin{cases} 1 & \text{Data from LC}_j \text{ arrives at LC}_i \\ 0 & \text{Data from LC}_j \text{ does not arrive at LC}_i \\ 0 & j = i \end{cases} \tag{2}$$

In the situation described in the previous paragraph and illustrated in Fig. 4, let us consider that each SM corresponds to a node of the graph $\mathbb{G}$ with a scalar first-order single-integrator dynamics, and the capacitor voltage in the ith and jth sub-modules are respectively V_i and V_j. Under this framework, it can be said that the capacitor voltages which belong to the cyber graph $\mathbb{G}$ (see Fig. 4) achieve consensus if $[V_i(k) - V_j(k)] \to 0$ as $k \to \infty$ [5], [57]. This is achieved

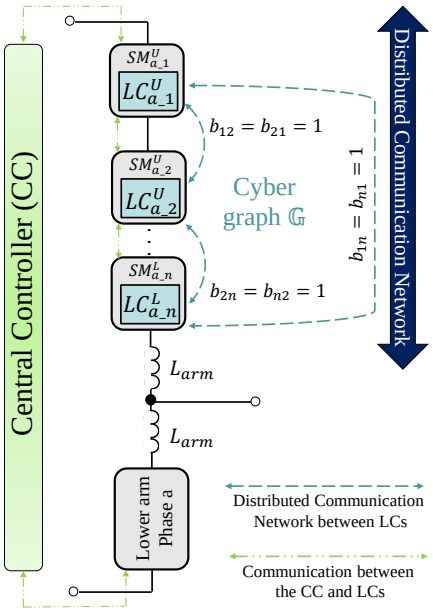


Fig. 4. Scheme of M2C upper arm operating in a consensus-based distributed control scheme. (For simplicity, LCs in the lower arm are not shown)

through a feedback loop by applying the protocol u_i given by (3) [28], which does not require communication between LCs and the CC for its implementation. The parameter h_i modifies the transient behaviour of the controller (3). Note that the controller gain h_i was tuned using the heuristic approach reported in [69], where a first approximation of the gains was obtained using the root locus method. Other methods for tuning the parameters of consensus algorithms are discussed in references [70], [71]. This paper uses the controller (3) to balance the voltage of the capacitors in the M2C. This control is distributed in that it only depends on the immediate neighbours $j \in \mathfrak{N}(i)$ of node i in the graph topology.

$$u_i(k) = -\frac{1}{h_i} \sum_{j \in \mathfrak{N}(i)} b_{ij} \cdot (V_i(k) - V_j(k)), \tag{3}$$

Fig. 5 shows the implementation of the consensus-based distributed control scheme for SM voltage balancing, considering the ith SM on the upper arm of phase a. As observed, the overall control action $U_i^{overall}$ sent to the modulation stage is composed of two control actions: U_i/n, which is generated by the CC (to control high-level tasks), and u_i, which is produced by the consensus-based distributed control scheme (3) for individual SM voltage balancing. The block labelled as FDIA detector is discussed in the coming section.

B. Consensus-based Distributed Control of M2C and Kalman Filter-based FDIA detector

Implementing (3) requires an exchange of voltage measurements between LCs, as shown in Fig. 5. As observed, the voltage measurements exchanged are realised by LCs through the distributed communication network. Thus, FDIAs targeting these measurements can highly affect this control architecture's performance [10]. Because of this, in [10], an FDIA detector based on a modified Kalman filter-based method was proposed, as shown in Fig. 5. The FDIA detector [10] is

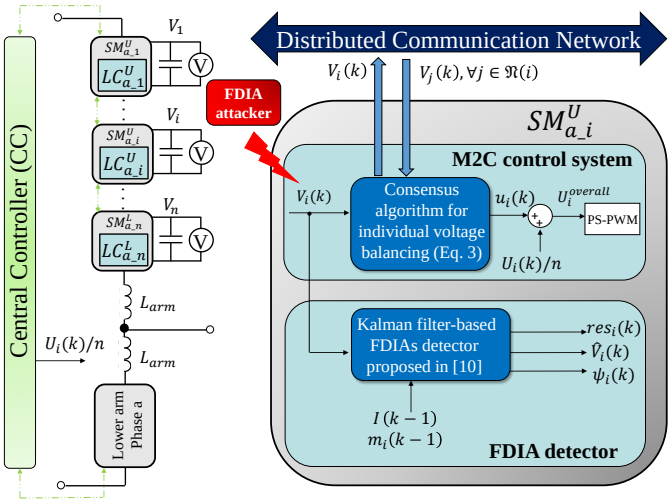


Fig. 5. Control system and FDIA detector [10] implemented on the LC of the $SM^U_{a\_i}$.

implemented in each LC of the M2C and aims to discriminate if the voltage measured locally by each LC and sent to the neighbour LCs via the distributed communication network is truthful. Fig. 5 shows its implementation considering the ith LC. As shown in this graphic, its input corresponds to the capacitor voltage $V_i(k)$ at the time instant k; the modulation index $m_i(k-1)$ and arm current $I(k-1)$. Based on that information, the detector proposed in [10], outputs at each time instant k, the residual index $res_i(k)$, the indicator function $\psi_i(k)$, and the estimated voltage $\hat{V}_i(k)$. If $res_i(k)$ is greater than a predefined threshold c, an FDIA is detected and $\psi_i(k) = 1$, otherwise $\psi_i(k) = -1$.

It must be recalled that the FDIA detector [10] only was validated considering step FDIAs, i.e., for cases where the attack sequence $V_i^a(k)$ in (1) is constant in steady state. However, its performance in detecting more sophisticated FDIAs is not guaranteed. In this sense, it is crucial to have a technique to analyse this detector and automatically determine all the FDIAs able to bypass the detector and, from this insightful information, develop strategies for its improvement. In this regard, this paper proposes a tool based on the artificial intelligent method named RL to achieve this objective. The proposed technique is validated using the detector [10]. However, it could be used to analyse the vulnerabilities of any FDIA detector in terms of FDIAs targeting the voltage measurements of the M2C SMs, proposed for M2Cs. The proposal is introduced in the next section.

IV. Proposed RL scheme to unveil vulnerabilities on FDIAs detectors

This paper proposes a reinforcement learning (RL) based method to obtain an FDIA attacker (a neural network) able to discover the vulnerabilities of existing false data injection attack (FDIA) detectors used for M2Cs (in terms of FDIAs targeting the voltage measurements of the M2C SMs). In this regard, the proposal's performance is validated using the FDIA detector [10]. The general scheme of the proposal is shown in Fig. 1 and Fig. 5.

Additionally, it must be highlighted that in the RL-based method proposed in this paper, it is necessary to define the following elements to use the RL technique for obtaining the FDIA attacker (see Fig. 1 and Fig. 5): (i) the inputs of the actor, which is the NN that will define the attack, and the critic, another NN that evaluates the cost; (ii) the output of the actor; (iii) the reward function that drives the training; and (iv) the experiment design. These elements are introduced and discussed in detail in the next section.

A. Reinforcement Learning and TD3

Reinforcement Learning [72] is a machine learning technique where an agent that computes decisions learns according to its interaction with the environment driven by a reward function $r(\cdot)$ [see Fig. 6 and (4)]. In this context, a reward function evaluates how well an agent behaves in the environment, and one desires that the resulting agent after training is the one that maximises the reward. RL is chosen for this problem for its capacity to learn according to the observed dynamics of the system, in this case, the dynamic behaviour of the Kalman Filter-based FDIA detector [10] in M2C environment (see Fig. 5).

In actor-critic RL algorithms, such as TD3 [73], the actor (or policy) $\pi(\cdot)$ is a function which receives as an input an observation (or state) s of the system and outputs the desired action a to be applied to the environment such that $a = \pi(s)$. The critic $Q^\pi(\cdot)$ or value function, evaluates the future performance of the system, or reward, as the Q-value $Q^\pi(s,a)$ for a given policy π, current observation s and current action a. In this proposal, the policy $\pi(\cdot)$ depends on the weights of the neural network θ, i.e., $\pi(\cdot) = \pi(\theta)$. Thus, the critic is given by $Q^{\pi(\theta)}(\cdot)$. To simplify the notation, from here onward, the latter expression (the critic) is considered as $Q^\theta(\cdot)$.

During learning, the objectives are to find: (i) the actor π_ϕ, given by a neural network (NN) with parameters ϕ, that yields (nearly) optimal performance; and (ii) the critic Q^θ, given by the NN with parameters θ, that evaluates (as precisely as possible) the future performance of the system such that π_ϕ is the policy. The conditions these NNs must satisfy in the optimum can be characterized by the Bellman Equation (4) [72].

$$\begin{aligned} Q^\theta\left(s(k),a(k)\right) =& r\left(s(k),a(k)\right) \\ &+\mathbb{E}_{s(k+1),a(k+1)}\left\{Q^\theta\left(s(k+1),a(k+1)\right)\right\} \end{aligned} \tag{4}$$

where $s(k), a(k)$ and $s(k+1), a(k+1)$ are the observation and action vectors at discrete times k and $k+1$ respectively, such that $a(k+1) = \pi^\phi\left(s(k+1)\right)$ and $s(k+1) = f\left(s(k),a(k)\right)$. In (4), $\mathbb{E}(\cdot)$ corresponds to the expected value of the function $Q^\theta(\cdot)$ given the observation and action vectors at time instant $k+1$. The learning process aims that the actor and critic satisfy these conditions by iteratively updating the critic and actor networks driven by the observed reward.

In this work, the TD3 variant of RL is chosen for its potential high performance and ability to handle continuous-space decision problems; it is considered the state of the art on RL methods with continuous actions space [73]. Besides the basic scheme described in the paragraphs above, TD3 uses a target actor and critic to obtain transitions and Q-value references. These target networks' evolution is lagging that of the original actor and critic, and these are used to avoid instability issues. Additionally, as TD3 aims to tackle a problem of overestimation bias in the Q-function produced by other RL methods [74], it incorporates three other features: (i) it learns two Q-functions instead of one (thus 'twin'), and the smaller Q-value is used to form the targets for the Bellman error loss functions (a measure of how far the system of satisfying the Bellman equation (4) [72]); (ii) the policy (and target networks) are updated less frequently than the Q-function; and (iii) noise is added to the target action to make it harder for the policy to exploit Q-function errors. Details concerning the operation of the TD3 algorithm are discussed in [73].

Note that s is referred to as the state or observation vector. RL theory is based on the use of s as the typical state space vector typically considered in dynamic systems. However, in practical RL, it may consist of more or less information, depending on the application, and thus for generality, s is referred to as an observation vector. The system's state may not always be measured, so one may need to use an observation vector s, which is smaller than the state of the system, to find the action. Alternatively, one may want to obtain the action as a function of other variables besides the state (such as external measurements, disturbances, etc.), which would make the observation vector s greater than the actual state vector. Obviously, one may also use part of the state and external variables for the observation s.

Finally, it is highlighted that the aim of the proposed RL scheme illustrated in Fig. 6 is to train an actor (an NN) able to generate FDIAs to deceive FDIAs used for M2Cs. Once the actor is trained via simulation work, its performance is validated using HIL work (see section V-B).

B. FDIA with RL

In this work, RL is used to train an actor for the ith LC, which will implement the decision of the FDIA [i.e. $V_i^a(k)$ in (1)] aiming to exploit the weaknesses of FDIA detectors used for M2Cs. It is worth remembering that in this work, weaknesses of FDIA detectors are understood as the set of attack sequences $V_i^a(k)$ not detected by the FDIA detector studied. The trained neural network (NN), which corresponds to the actor in the proposed RL environment, could be used later to improve the detection performance of the studied FDIA detector. This paper evaluates the proposal's effectiveness using the FDIA detector [10]

In Fig. 6, the output of the actor associated with ith LC is the attack signal $V_i^a(k)$ (see (1)) which aims to bypass detection. In order to guide the learning of the FDIA attacker, the reward signal is a function of the error between the real voltage and that estimated by the Kalman filter-based FDIA

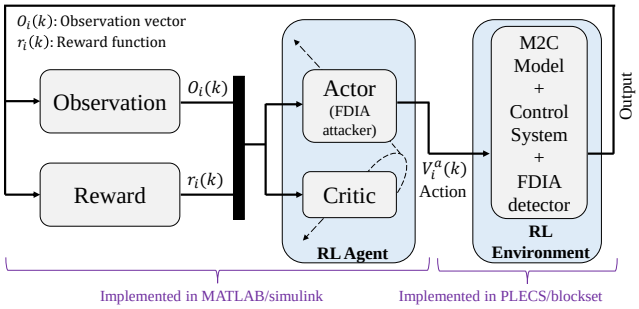


Fig. 6. RL scheme proposed in this work. In this proposal, the actor is trained via simulations, and its performance is validated through HIL work. (Note that the proposal only considers information of the ith SM in its formulation)

detector [10], $e_i = V_i - \hat{V}_i$. Indeed, if e_i is large, the attack is working as the detector is wrong about the voltage value, but only as long as the attack has not been discovered. If it is discovered, the system operator may take measures to stop or react to the attack. Therefore, we want a reward function that provides large values if the error e_i is large when there is an attack that is not discovered. This concept is used to construct the reward function r_i at sampling time k as shown in (5). In this case, r_i correspond to $r(\cdot)$ in (4).

$$r_i(k) = -\psi_i(k)(e_i^2(k) + [1 + |\dot{e}_i(k)|]^2) \tag{5}$$

where $\psi_i(k)$ indicates whether the attack has been discovered by the FDIA detector ($\psi_i(k) = 1$) or not ($\psi_i(k) = -1$). Note in (5) that $\dot{e}_i = [e(k) - e(k-1)]/T$. If the attack has not been discovered, then $r_i(k) = e_i^2(k) + (1 + |\dot{e}_i(k)|)^2 > 0$, which will assign greater rewards if e_i is large, and will make the agent to move toward this kind of behaviours. On the other hand, if the attack has been discovered, then $r_i(k) = -e_i^2(k) - (1 + |\dot{e}_i(k)|)^2 < 0$, which is a negative reward, and thus the agent will be dissuaded from this type of behaviours. Finally, note that the reward function defined by (5) is related to the voltage across capacitor of the ith SM in the M2C.

After these definitions, a pseudo-code showing the usual operation of TD3, but applied to this problem, is shown on Algorithm 1. The idea is basically that at each sampling time the transitions composed of the observation, the FDIA values obtained as the current output of the actor, the reward and the next observation (depending on the applied FDIA values) are saved in a buffer. Then, a subset of those transitions are sampled, a target output is computed for them and the critic is updated by minimising the error between the critic output and the target. Finally, every d sampling periods, the actor is updated by a policy gradient, and target networks are updated by a delay of the non-target networks. This procedure is repeated at each sampling time, and the new information is added to the buffer and used in the update of the networks. Notice how the update law of the actor aims to satisfy the Bellman equation. The algorithm stops when all simulation scenarios set for training have been used, or when the reward of the transitions in the buffer does not increase anymore. In Algorithm 1, T is a number of discrete instants considered

for training, and σ and $\tilde{\sigma}$ are the variances of the exploration noises which is limited by c for the targets.

Algorithm 1 TD3 for training of FDIA-agent on SM i

Initialize critic networks Q_{θ_1}, Q_{θ_2}, and actor network π_ϕ with random parameters θ_1, θ_2, ϕ Q_{θ_1}.
Initialize target networks $\theta_1' \leftarrow \theta_1$, $\theta_2' \leftarrow \theta_2$, $\phi' \leftarrow \phi$.
Initialize replay buffer $\mathcal{B}$.
for $k = 1 \dots T$ **do**
  Select FDIA action as a function of observation $O_i(k)$ with exploration noise. Inject FDIA attack to node i, and obtain reward $r_i(k)$ and new observation $O_i(k+1)$.
  Store transition tuple $(O_i(k), V_a^i(k), r_i(k), O_i(k+1))$ in $\mathcal{B}$.
  Sample mini-batch of N transitions (O_i, V_a^i, r_i, O_i^+) from $\mathcal{B}$, where O_i^+ is the next observation following O_i.
  for all sampled transitions **do**
    Obtain target outputs and then target Q-values as
    $\tilde{V_a^i} \leftarrow \pi_{\phi'}(O_i^+) + \epsilon$, $\epsilon \sim$ clip$(\mathcal{N}(0, \tilde{\sigma}), -c, c)$
    $y \leftarrow r + \gamma \min_{j=1,2} Q_{\theta_j'}(O_i^+, \tilde{V_a^i})$
  end for
  Update critics by minimizing error with respect to target reward of sampled transitions
  $\theta_j \leftarrow \arg\min_{\theta_j} \frac{\Sigma(y - Q_{\theta_j}(O_i^+, \tilde{V_a^i}))^2}{N}$.
  if $k \bmod d$ **then**
    Update ϕ by the deterministic policy gradient on the sampled transitions:
    $\nabla_\phi J(\phi) = \Sigma \nabla_{V_a^i} Q_{\theta_1}(O_i, V_a^i)|_{V_a^i = \pi_\phi(O_i)} \nabla_\phi \pi_\phi(O_i)/N$
    Update target networks:
    $\theta_i' \leftarrow \tau\theta_i + (1-\tau)\theta_i'$, $\phi' \leftarrow \tau\phi + (1-\tau)\phi'$
    If $O_i(k+1)$ is last state of the episode, reset the system for a new episode.
  end if
end for

As RL is a learning method, the ability of the agent to learn will firstly depend on the data it is trained with. Therefore, it is necessary to select the scenarios where the TD3 learning will happen. In this context, the concept of epistemic uncertainty refers to the operating conditions that the agent is not trained with and, thus, is not prepared to handle. Future research will explore how to best handle the situations when the current operating conditions have not been considered in training.

As shown in Table I, three possible settings for the agent associated with the ith LC are considered in this work. The different cases characterise what may be known by the FDIA attacker, and define what is included in the observation vector used in the actor and the critic. In the first case, it is assumed only voltage estimations, and thus error signals, are known; thus the observations at time k are $O_i(k) = \left\{e_i(k), \dot{e}_i(k), \int e_i(k)\right\}$. In the second case, it is assumed that the attacker knows the FDIA detector [10] (see Fig. 5); therefore the observations are given by $O_i(k) = \left\{e_i(k), \dot{e}_i(k), \int e_i(k), res_i(k)\right\}$. In the third case, the attacker knows more information of the system than just the estimated voltages, and the observations are given by $O_i(k) = \left\{e_i(k), \dot{e}_i(k), \int e_i(k), I_i(k), m_i(k)\right\}$, where $I_i(k)$ is the arm current and $m_i(k)$ is the modulation index (see Fig. 5). These cases were selected to evaluate the performance of the FDIA attacker when it it trained with different levels of information in the system. Note that the cases studied in this work (see Table I) consider electrical variables associated with the M2C control system, such as the capacitor voltage, the arm current and the modulation index.

TABLE I
OBSERVATION VECTORS USED IN THE PROPOSED RL-BASED SCHEME

	Information known by the FDIA attacker (Observation vector)
Case 1	$O_i(k) = \left\{e_i(k), \dot{e}_i(k), \int e_i(k)\right\}$
Case 2	$O_i(k) = \left\{e_i(k), \dot{e}_i(k), \int e_i(k), res_i(k)\right\}$
Case 3	$O_i(k) = \left\{e_i(k), \dot{e}_i(k), \int e_i(k), I_i(k), m_i(k)\right\}$

It should be noted that training of RL agents can be very time-consuming and computationally intensive. However, in this case, the FDIA agent only deals with a single SM of the M2C, considering that the behaviour of all the M2C SM is similar if the control system is working correctly. This keeps the system simple, and as will be seen in the results section, training times are reasonable, even when implemented on standard desktop computers.

We should notice that there are 3 fundamental paradigms in machine learning: RL, supervised learning (SL), and unsupervised learning (USL) [72]. In SL, the NN is taught to replicate input-output pairs. In the context of this application, we would need an already-designed FDIA attacker that the NN would learn to mimic. The USL training is used to find patterns and/or group data in different clusters; this cannot be used to train an attacker on its own. Therefore, RL is better suited for this problem, as the NN will learn directly through interaction with the system, as shown in Fig. 6.

Finally, note that the proposed method generates an FDIA attacker (a neural network) able to generate attack sequences that deceive the studied FDIA detector, considering as attack target the ith SM (see Fig. 6): This FDIA attacker can be used to attack any other M2C SM, as shown in Fig. 7. In that figure, i SMs are attacked by i FDIA attackers: The only difference among them is their observation vectors $O(k)$.

V. HARDWARE IN THE LOOP VALIDATION

It must be pointed out that if in the proposal displayed in Fig. 6, the reward and critic blocks are eliminated from the scheme, which corresponds to the case when the FDIA-agent is being used, after training, it is clear that the actor (FDIA attacker) works as a feedback controller aiming to avoid detection of the attack. This configuration is used and implemented in HIL, as shown in Fig. 8 to validate the performance of the FDIA attacker (see Fig. 6).

Based on the above discussion, it must be noted that the proposed RL scheme comprises two stages; the first one corresponds to the one where the FDIA attacker is trained to generate attack sequences to deceive the FDIA detector studied (see Fig. 6). And the second stage, where the performance of the FDIA attacker is validated, as shown in Fig. 8. Note that the first stage is performed via simulation work, whereas the second one is validated via HIL work.

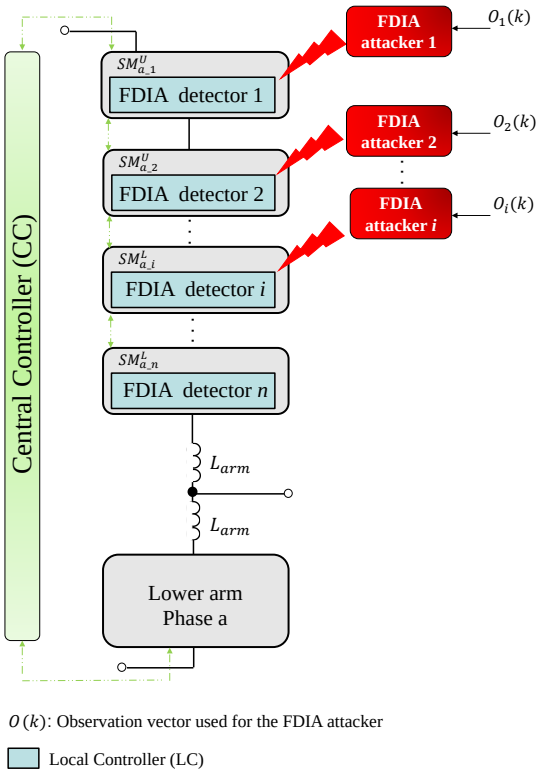


Fig. 7. Scheme that shows the use of the FDIA attacker generated by the proposal. (For simplicity, only the upper arm of the M2C is detailed)

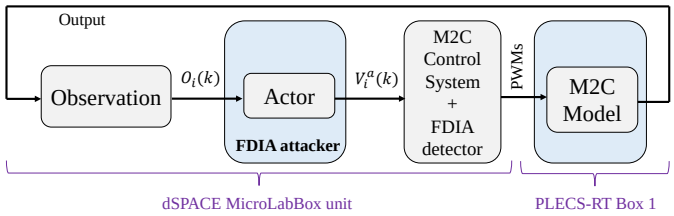


Fig. 8. Scheme used for the HIL validation.

It is worth remembering that the performance of the FDIA attacker for generating complex FDIAs able to deceive the FDIA detectors is evaluated considering the Kalman filter-based FDIA detector [10]. To this end, the M2C shown in Fig. 9 with the parameters listed in Table II is used in this work to train the RL scheme (see Fig. 6) and for HIL validation (see Fig. 8). From the training process, the actor (FDIA attacker) with the highest reward is selected to evaluate performance on the detection of [10] using the HIL-based experimental system shown in Fig. 10. In this setup, an M2C with the characteristics displayed in Table II is modeled by two PLECS-RT Box-1 HIL platforms, whereas the control system and the FDIA detector [10] are implemented on a dSPACE MicroLabBox unit.

A. Training process

As explained in section IV-A, the proposed RL scheme (see Fig. 6) is trained through simulation work. Then, the actor obtained (FDIA attacker) is used to validate, in the HIL

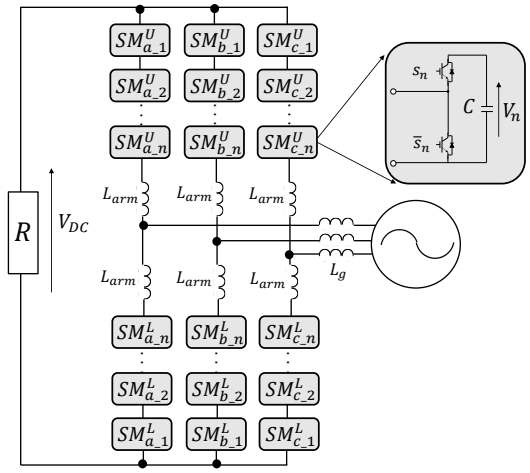


Fig. 9. M2C used for the HIL validation.

TABLE II
M2C PARAMETERS USED FOR THE HIL VALIDATION

Description	Experimental M2C
Grid inductance (L_g)	0.8mH
Arm inductance (L_{arm})	4.15mH
SM capacitance (C)	3.3mF
Number of SM per arm (n)	3 (18 SMs in total)
Carrier frequency (PS-PWM modulation)	8kHz
Grid frequency (f_g)	50Hz
Grid voltage (V_g)	$60V_{RMS}$
Power (P_n)	1.5kW
SM voltage reference (V_n^*)	70V
DC-link voltage (V_{DC})	182V
Resistive load (R)	22.5Ω
Consensus gain (h_i)	1
Threshold (c) for FDIA detection	$0.05V_n^*$

environment, the actor's effectiveness for generating FDIAs to deceive the studied detector.

In the proposed RL algorithm illustrated in Fig. 6, the actor and critic correspond to NNs with the following structure:

Actor Network: It is a multilayer perceptron where the input corresponds to the observation vector $O_i(k)$. The hidden layer is a fully-connected-layer composed by 10 neurons with sigmoid activation functions. Finally, the output layer is a single neuron with an activation function given by the identity, and corresponds to the attack signal $V_i^a(k)$.

Critic Network: It is also a multilayer perceptron but the inputs are the observation vector $O_i(k)$ and the attack signal $V_i^a(k)$. The activation function in every neuron is a ReLu [75]. Note that ReLus are popular because, unlike other activation functions, they do not activate all neurons at the same time [75]. The hidden layer is a fully-connected-layer with 32 neurons and the output layer has a single neuron which approximates the Q-value for the current observation vector and attack signal.

The training of the actor and the critic are performed with TD3, using the parameters listed in Table III. Finally, it must be remembered that the training of the actor (see

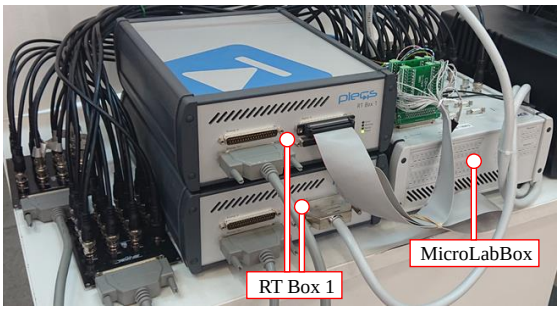


Fig. 10. Platform used for hardware-in-the-loop validation.

TABLE III
PARAMETERS OF BOTH ACTOR AND CRITIC IN THE TRAINING PROCESS.

Parameter	value
Max. episodes	400
Max. steps per episode	800
Score Averaging window length	5
Stop training criteria	Average reward
Stop training value	$5 \cdot 10^9$

Fig. 6) is made through simulation work implemented in Matlab/Simulink augmented by PLECS/blockset: This code is made available to researchers on a website (see [21]), to motivate future research efforts in this area.

B. HIL Validation

The proposed RL-based scheme is validated following the scheme shown in Fig. 8. As seen, the actor (FDIA attacker) obtained from the training process is used to generate an FDIA targeting the $SM^L_{a\_1}$ placed in the lower arm of phase "a" (see Fig. 9). The following five scenarios are considered:

Scenario 1 (benchmark case): It studies the performance of the step FDIA (i.e., $V^a_i(k) = constant$ in (1)) on the detector [10]. To this end, a step FDIA targeting the voltage measured by the $SM^L_{a\_1}$ is considered. As observed in Fig. 11(a), the step FDIA starts at 2s, and it persists during the rest of the test. Fig. 11(b) shows the SM voltages in the attacked arm seen by the M2C control system. As seen, the FDIA on $SM^L_{a\_1}$ is interpreted as a perturbation on the control system, and it is appropriately managed. However, those voltages are actually incorrect, i.e., they do not correspond to the real voltages measured directly across the capacitor of each SM. The actual voltages are shown in Fig. 11(c), from where it is possible to observe the effects of the FDIA on the M2C operation. Note that such effects may lead to sub-optimal or even unstable operation of the whole M2C. Fortunately, this is avoided in this case, as the detector [10] effectively detects the step FDIA and the attacked SM, as shown in Fig. 11(e). Finally, Fig. 11(d) shows the three-phase currents at the M2C AC side.

Scenario 2: The effectiveness of the FDIA attacker (see Fig. 8) in generating FDIAs able to deceive the FDIA detector [10] is tested considering the observation vector as $O_i(k) = \left\{e_i(k), \dot{e}_i(k), \int e_i(k), I_i(k), m_i(k)\right\}$. The attack sequence obtained from the FDIA attacker generated considering this observation vector is shown in Fig. 12(a). In Fig. 12(b), the voltages seen by the control system for this scenario are

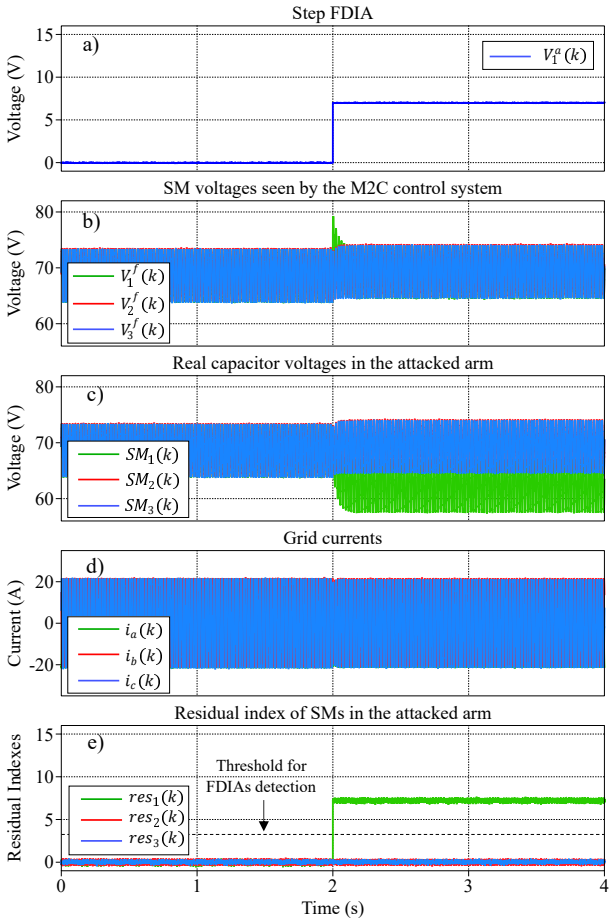


Fig. 11. Plots scenario 1: (a) Sequence attack targeting $SM^L_{a\_1}$ (step FDIA), (b) SM voltages seen by the M2C control system in the attacked arm, (c) Real SM voltages, (d) Grid currents, (e) Residual indexes generated by detector [10].

plotted. As observed in Fig. 12(c), the sophisticated attack generates a similar effect on the real SM voltages to the step FDIA considered in the previous example. However, differently from the latter, the attack sequence generated by the FDIA attacker is not detected by the FDIA detector [10], as shown in Fig. 12(e). These results show the effectiveness of the RL scheme in generating FDIAs able to deceive the detector [10]. The information generated by the proposed RL method can be used to improve the robustness of the detector, increasing its detection capability against more sophisticated FDIAs. Fig. 11(d) shows the three-phase currents at the M2C AC side, where it is possible to see a small imbalance in those currents when the FDIA is in operation.

Scenario 3: This case is similar to the previous one, but considering an observation vector of $O_i(k) = \left\{e_i(k), \dot{e}_i(k), \int e_i(k), res_i(k)\right\}$ for training the RL method (see Fig. 6). Fig. 13 shows the plots related to this scenario. In particular, Fig. 13(a) shows the attack sequence generated by the FDIA attacker (it starts at 2s). Fig. 13(c) shows the effects of this sophisticated FDIA on the real capacitor voltages in the attacked arm. The results shown in Fig. 13(e) demonstrate that the FDIA shown in Fig. 13(a) is not detected by the detector [10], showing the effectiveness of the proposal in

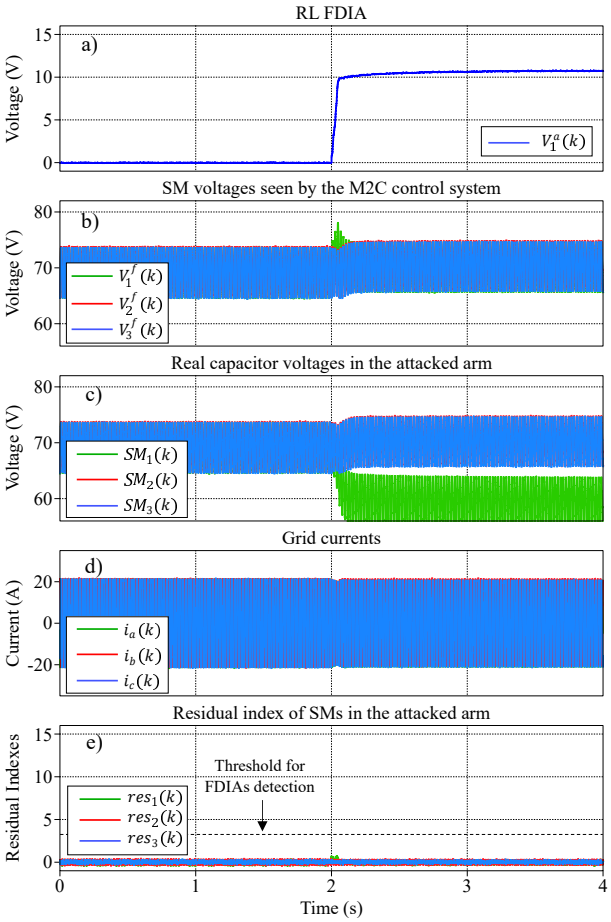


Fig. 12. Plots scenario 2: (a) Sequence attack targeting $SM_{a\_1}^L$ generated by the proposed RL method, (b) SM voltages seen by the M2C control system in the attacked arm, (c) Real SM voltages, (d) Grid currents, (e) Residual indexes generated by detector [10]. (The FDIA starts at 2s)

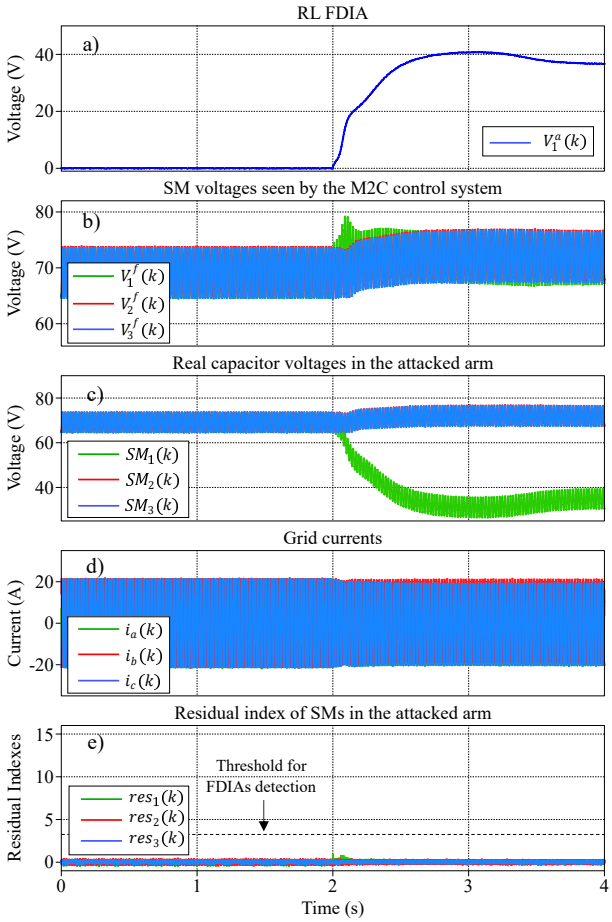


Fig. 13. Plots scenario 3: (a) Sequence attack targeting $SM_{a\_1}^L$ generated by the proposed RL method, (b) SM voltages seen by the M2C control system in the attacked arm, (c) Real SM voltages, (d) Grid currents, (e) Residual indexes generated by detector [10]. (The FDIA starts at 2s)

generating FDIAs which can deceive FDIA detectors used for M2Cs. In Fig. 13, the SM voltages seen by the M2C control system are shown as in Fig. 13(b). This means that from the M2C control system point of view, the FDIA is considered an external perturbation, and the control system adequately manages it. In this situation, everything looks fine for the control system; however, in reality, the normal operation of the M2C is affected, which might produce a critical failure of the M2C. Similarly to the previous scenarios, the FDIA produces unbalanced currents on the M2C AC side, as shown in Fig. 13(b).

Scenario 4: This test considers the observation vector $O_i(k) = \{e_i(k), \dot{e}_i(k), \int e_i(k)\}$ for training the RL-based method and obtaining the DFIA attacker. Note that for this case, the information (observation vector) of the M2C control system that the attacker knows is less than that considered in cases 2 and 3. Nonetheless, with this limited information, and similar to scenarios 2 and 3, the proposal can produce an FDIA attacker that generates attack sequences (see Fig. 14(a)) that deceive the studied FDIA attacker, as shown in see Fig. 14(d). This is evidenced in Fig. 14, which shows the main waveforms associated with this scenario. Fig. 14(a) displayed the attack sequence generated by the FDIA attacker obtained in this scenario, and Fig. 14(d) shows that such attack sequence is not detected by the FDIA detector [10]. Finally, Fig. 14(b) displayed the SM voltages seen by the M2C control system, whereas in Fig. 14(c), the real capacitor voltages are shown. These figures show the effects of the FDIAs, in the sense that everything looks fine for the control system; however, in reality, this is not true, as shown in Fig. 14(c).

Scenario 5 (simultaneous attacks): Note that the FDIA detector [10] is implemented in each local controller of the M2C (see Fig. 3). In this context, in this test, the sub-modules $SM_{a\_1}^L$ and $SM_{b\_2}^U$ shown in Fig. 9 are simultaneously attacked by a sophisticated FDIA (the one used in scenario 2) and a step FDIA, respectively. Fig. 15 shows the output of the FDIA detector [10] related to this test. As seen in Fig. 15(b), the FDIA detector used to validate our RL-based scheme can detect the simple attack given by a step FDIA. However, this detector does not detect the sophisticated attack sequence generated by our RL scheme for detecting vulnerabilities of FDIA detectors for M2Cs, as shown in Fig. 15(a). This latter result shows the proposal's effectiveness in generating attack sequences able to deceive the studied FDIA detector [10].

Finally, based on the scenarios discussed in this section, it is possible to get helpful information about the type of attack

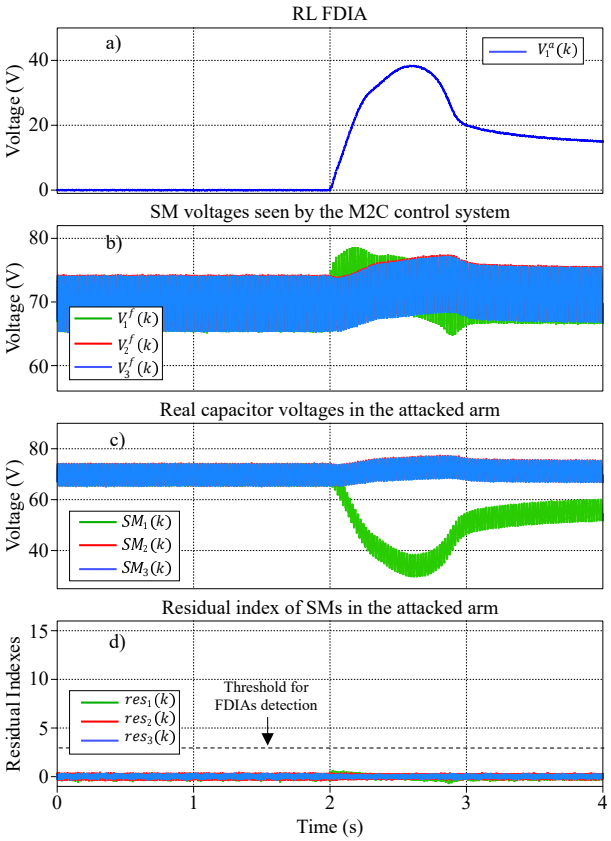


Fig. 14. Plots scenario 4: (a) Sequence attack targeting $SM^L_{a\_1}$ generated by the proposed RL method, (b) SM voltages seen by the M2C control system in the attacked arm, (c) Real SM voltages, (d) Residual indexes generated by detector [10]. (The FDIA starts at 2s)

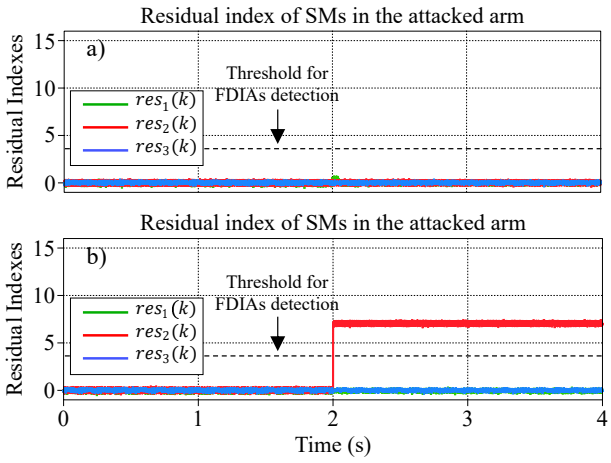


Fig. 15. Plots scenario 5: (a) Residual index generated by detector [10] when a sophisticated FDIA attacks $SM^L_{a\_1}$, (b) residual index generated by detector [10] when $SM^U_{b\_2}$ is attacked by a step FDIA.

sequences that deceive the studied FDIA detector. Using this information, an RL-based scheme can be developed to train an agent able to detect this type of FDIAs and, therefore, complement the FDIA detector studied here [10].

VI. Conclusions

This paper proposes a deep reinforcement learning-based method to exploit the vulnerabilities of the existing cyberattack detection methods. Based on the results of the implementation, the current FDIA detectors can be improved, increasing their detection effectiveness. To sum up, the proposed RL-based scheme effectively can generate an FDIA attacker (a NN) able to generate attack sequences (FDIAs) that deceive the studied FDIA detector [10].

It must be recalled that the attack sequences generated by the proposal and discussed in section V depend on the characteristics of the FDIA detector studied, i.e. the more sophisticated the FDIA detectors, the more complex attack sequences will be generated by the proposed RL-method.

Finally, note that this study serves as a foundation for establishing more reliable cybersecurity solutions for M2Cs. The proposal source code is available in [21] to promote future research. Future work will consider iterative training of detectors, i.e., train a new detector, exploit its vulnerabilities, train a second version of the new one, etc.

As future work, the following aspects can be studied further: (i) to study FDIAs on measurements associated with the arm currents of the M2C, (ii) the study of more sophisticated FDIAs and methods for their detection, and (iii) using information about whether the current operating conditions were considered in training to determine the best course of action during an attack.

References

[1] S. Du, A. Dekka, B. Wu, and N. Zargari, *Modular multilevel converters: analysis, control, and applications*. John Wiley & Sons, 2017.

[2] S. P. Teeuwsen, "Modeling the trans bay cable project as voltage-sourced converter with modular multilevel converter design," in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1–8.

[3] Y. Luo, Z. Li, Y. Li, and P. Wang, "A distributed control method for power module voltage balancing of modular multilevel converters," in *2016 IEEE Energy Conversion Congress and Exposition (ECCE)*. IEEE, 2016, pp. 1–5.

[4] B. Xia, Y. Li, Z. Li, F. Xu, and P. Wang, "A distributed voltage balancing method for modular multilevel converter," in *2017 IEEE 3rd International Future Energy Electronics Conference and ECCE Asia (IFEEC 2017-ECCE Asia)*. IEEE, 2017, pp. 1944–1948.

[5] S. Song and J. Liu, "Interpreting the individual capacitor voltage regulation control of psc-pwm mmc via consensus theory," *IEEE Access*, vol. 7, pp. 66 807–66 820, 2019.

[6] S. Yang, Y. Tang, and P. Wang, "Distributed control for a modular multilevel converter," *IEEE Transactions on power Electronics*, vol. 33, no. 7, pp. 5578–5591, 2017.

[7] H. Wang, S. Yang, H. Chen, X. Feng, and F. Blaabjerg, "Synchronization for an mmc distributed control system considering disturbances introduced by submodule asynchrony," *IEEE Transactions on Power Electronics*, vol. 35, no. 12, pp. 12 834–12 845, 2020.

[8] S. Yang, S. Liu, J. Huang, H. Su, and H. Wang, "Control conflict suppressing and stability improving for an mmc distributed control system," *IEEE Transactions on Power Electronics*, vol. 35, no. 12, pp. 13 735–13 747, 2020.

[9] L. Mathe, P. D. Burlacu, and R. Teodorescu, "Control of a modular multilevel converter with reduced internal data exchange," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 248–257, 2016.

[10] C. Burgos-Mellado, F. Donoso, T. Dragičević, R. Cardenas-Dobson, P. Wheeler, J. Clare, and A. Watson, "Cyber-attacks in modular multilevel converters," *IEEE Transactions on Power Electronics*, vol. 37, no. 7, pp. 8488–8501, 2022.

[11] S. Sahoo, J. C.-H. Peng, A. Devakumar, S. Mishra, and T. Dragičević, "On detection of false data in cooperative dc microgrids—a discordant element approach," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 8, pp. 6562–6571, 2019.

[12] M. R. Habibi, H. R. Baghaee, T. Dragičević, and F. Blaabjerg, "Detection of false data injection cyber-attacks in dc microgrids based on recurrent neural networks," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 9, no. 5, pp. 5294–5310, 2020.

[13] A. J. Abianeh, Y. Wan, F. Ferdowsi, N. Mijatovic, and T. Dragičević, "Vulnerability identification and remediation of fdi attacks in islanded dc microgrids using multiagent reinforcement learning," *IEEE Transactions on Power Electronics*, vol. 37, no. 6, pp. 6359–6370, 2021.

[14] K. Manandhar, X. Cao, F. Hu, and Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using kalman filter," *IEEE transactions on control of network systems*, vol. 1, no. 4, pp. 370–379, 2014.

[15] S. Sahoo, T. Dragičević, and F. Blaabjerg, "Cyber security in control of grid-tied power electronic converters—challenges and vulnerabilities," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 9, no. 5, pp. 5326–5340, 2019.

[16] Z. Li, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah, "Bilevel model for analyzing coordinated cyber-physical attacks on power systems," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2260–2272, 2015.

[17] A. Pinceti, L. Sankar, and O. Kosut, "Detection and localization of load redistribution attacks on large-scale systems," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 2, pp. 361–370, 2021.

[18] S. Mousavian, M. Erol-Kantarci, L. Wu, and T. Ortmeyer, "A risk-based optimization model for electric vehicle infrastructure response to cyber attacks," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6160–6169, 2017.

[19] S. Mousavian, M. Erol-Kantarci, and T. Ortmeyer, "Cyber attack protection for a resilient electric vehicle infrastructure," in *2015 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2015, pp. 1–6.

[20] T. Ding, Z. Zeng, B. Qin, J. Zhao, Y. Yang, F. Blaabjerg, and Z. Dong, "Quantifying cyber attacks on industrial mmc-hvdc control system using structured pseudospectrum," *IEEE Transactions on Power Electronics*, vol. 36, no. 5, pp. 4915–4920, 2020.

[21] [Online]. Available: https://tinyurl.com/Claudio-UOH

[22] H. Li, X. He, Y. Zhang, and W. Guan, "Attack detection in cyber-physical systems using particle filter: An illustration on three-tank system," in *2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*. IEEE, 2018, pp. 504–509.

[23] 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.

[24] W. Ao, Y. Song, and C. Wen, "Adaptive cyber-physical system attack detection and reconstruction with application to power systems," *IET Control Theory & Applications*, vol. 10, no. 12, pp. 1458–1468, 2016.

[25] A. N. Akpolat, M. R. Habibi, E. Dursun, A. E. Kuzucuoğlu, Y. Yang, T. Dragičević, and F. Blaabjerg, "Sensorless control of dc microgrid based on artificial intelligence," *IEEE Transactions on Energy Conversion*, vol. 36, no. 3, pp. 2319–2329, 2020.

[26] W.-H. Chen, S.-H. Hsu, and H.-P. Shen, "Application of svm and ann for intrusion detection," *Computers & Operations Research*, vol. 32, no. 10, pp. 2617–2634, 2005.

[27] Y. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2505–2516, 2017.

[28] F. L. Lewis, H. Zhang, K. Hengster-Movric, and A. Das, *Cooperative control of multi-agent systems: optimal and adaptive design approaches*. Springer Science & Business Media, 2013.

[29] H.-J. Knaak, "Modular multilevel converters and hvdc/facts: A success story," in *Proceedings of the 2011 14th European conference on power electronics and applications*. IEEE, 2011, pp. 1–6.

[30] H. Rao, "Architecture of nan'ao multi-terminal vsc-hvdc system and its multi-functional control," *CSEE Journal of Power and Energy Systems*, vol. 1, no. 1, pp. 9–18, 2015.

[31] L. Zhang, Y. Zou, J. Yu, J. Qin, V. Vittal, G. G. Karady, D. Shi, and Z. Wang, "Modeling, control, and protection of modular multilevel converter-based multi-terminal hvdc systems: A review," *CSEE Journal of Power and Energy Systems*, vol. 3, no. 4, pp. 340–352, 2017.

[32] R. Fan, J. Lian, K. Kalsi, and M. A. Elizondo, "Impact of cyber attacks on high voltage dc transmission damping control," *Energies*, vol. 11, no. 5, p. 1046, 2018.

[33] X. Zha, Y. Liu, and M. Huang, "Resilient power converter: A grid-connected converter with disturbance/attack resiliency via multi-timescale current limiting scheme," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 11, no. 1, pp. 59–68, 2021.

[34] U. D. of Energy. [Online]. Available: https://tinyurl.com/538fpep2

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

[36] B. Çiftçi, S. Schiessl, J. Gross, L. Harnefors, S. Norrga, and H.-P. Nee, "Wireless control of modular multilevel converter submodules," *IEEE Transactions on Power Electronics*, vol. 36, no. 7, pp. 8439–8453, 2020.

[37] S. Bera, S. Misra, and J. J. Rodrigues, "Cloud computing applications for smart grid: A survey," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 5, pp. 1477–1494, 2014.

[38] B. Xu, H. Tu, Y. Du, H. Yu, H. Liang, and S. Lukic, "A distributed control architecture for cascaded h-bridge converter with integrated battery energy storage," *IEEE Transactions on Industry Applications*, vol. 57, no. 1, pp. 845–856, 2020.

[39] B. Xu, H. tu, Y. Du, H. Yu, H. Liang, and S. Lukic, "A distributed control architecture for cascaded h-bridge converter," in *2019 IEEE Applied Power Electronics Conference and Exposition (APEC)*. IEEE, 2019, pp. 3032–3038.

[40] C. L. Toh and L. Norum, "A high speed control network synchronization jitter evaluation for embedded monitoring and control in modular multilevel converter," in *2013 IEEE Grenoble Conference*. IEEE, 2013, pp. 1–6.

[41] C. L. Toh and L. norum, "A performance analysis of three potential control network for monitoring and control in power electronics converter," in *2012 IEEE International Conference on Industrial Technology*. IEEE, 2012, pp. 224–229.

[42] T. P. Corrêa, L. Almeida, and F. J. Rodriguez, "Communication aspects in the distributed control architecture of a modular multilevel converter," in *2018 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2018, pp. 640–645.

[43] B. Xia, Y. Li, Z. Li, G. Konstantinou, F. Xu, F. Gao, and P. Wang, "Decentralized control method for modular multilevel converters," *IEEE Transactions on Power Electronics*, vol. 34, no. 6, pp. 5117–5130, 2018.

[44] P. D. Burlacu, L. Mathe, and R. Teodorescu, "Synchronization of the distributed pwm carrier waves for modular multilevel converters," in *2014 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*. IEEE, 2014, pp. 553–559.

[45] S. Huang, R. Teodorescu, and L. Mathe, "Analysis of communication based distributed control of mmc for hvdc," in *2013 15th European Conference on Power Electronics and Applications (EPE)*. IEEE, 2013, pp. 1–10.

[46] Y. Rong, J. Wang, Z. Shen, R. Burgos, D. Boroyevich, and S. Zhou, "Distributed control and communication system for pebb-based modular power converters," in *2019 IEEE Electric Ship Technologies Symposium (ESTS)*. IEEE, 2019, pp. 627–633.

[47] K. O. Akpinar and I. Ozcelik, "Analysis of machine learning methods in ethercat-based anomaly detection," *IEEE Access*, vol. 7, pp. 184 365–184 374, 2019.

[48] C.-W. Lin and A. Sangiovanni-Vincentelli, "Cyber-security for the controller area network (can) communication protocol," in *2012 International Conference on Cyber Security*. IEEE, 2012, pp. 1–7.

[49] Y. Musa, O. Tantawi, V. Bush, B. Johson, N. Dixon, W. Kirk, and K. Tantawi, "Low-cost remote supervisory control system for an industrial process using profibus and profinet," in *2019 SoutheastCon*. IEEE, 2019, pp. 1–4.

[50] E. E. Miciolino, G. Bernieri, F. Pascucci, and R. Setola, "Communications network analysis in a scada system testbed under cyber-attacks," in *2015 23rd Telecommunications Forum Telfor (TELFOR)*. IEEE, 2015, pp. 341–344.

[51] M. Bozdal, M. Samie, and I. Jennions, "A survey on can bus protocol: Attacks, challenges, and potential solutions," in *2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE)*. IEEE, 2018, pp. 201–205.

[52] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1630–1638, 2016.

[53] L. Che, X. Liu, Z. Li, and Y. Wen, "False data injection attacks induced sequential outages in power systems," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1513–1523, 2018.
[54] P.-H. Wu, Y.-C. Su, J.-L. Shie, and P.-T. Cheng, "A distributed control technique for the multilevel cascaded converter," *IEEE Transactions on Industry Applications*, vol. 55, no. 2, pp. 1649–1657, 2018.
[55] H. Geng, S. Li, C. Zhang, G. Yang, L. Dong, and B. Nahid-Mobarakeh, "Hybrid communication topology and protocol for distributed-controlled cascaded h-bridge multilevel statcom," *IEEE Transactions on Industry Applications*, vol. 53, no. 1, pp. 576–584, 2016.
[56] B. P. McGrath, D. G. Holmes, and W. Y. Kong, "A decentralized controller architecture for a cascaded h-bridge multilevel converter," *IEEE transactions on industrial electronics*, vol. 61, no. 3, pp. 1169–1178, 2013.
[57] C. Burgos-Mellado, J. Gutierrez, C. Pineda, F. Donoso, A. Watson, M. Sumner, R. Cardenas, and A. Mora, "Distributed control strategy based on a consensus algorithm for the inter-cell and inter-cluster voltage balancing of a cascaded h-bridge based statcom," in *2020 IEEE 21st Workshop on Control and Modeling for Power Electronics (COMPEL)*. IEEE, 2020, pp. 1–8.
[58] S. Neira, P. Poblete, J. Pereda, and F. Nuñez, "Consensus-based distributed control of a multilevel battery energy storage system," in *2020 IEEE 21st Workshop on Control and Modeling for Power Electronics (COMPEL)*. IEEE, 2020, pp. 1–7.
[59] W. Yao, J. Liu, and Z. Lu, "Distributed control for the modular multilevel matrix converter," *IEEE Transactions on Power Electronics*, vol. 34, no. 4, pp. 3775–3788, 2018.
[60] J. Liu, W. Yao, Z. Lu, and J. Ma, "Design and implementation of a distributed control structure for modular multilevel matrix converter," in *2018 IEEE Applied Power Electronics Conference and Exposition (APEC)*. IEEE, 2018, pp. 1934–1939.
[61] Y. Zhou, D. Jiang, P. Hu, J. Guo, Y. Liang, and Z. Lin, "A prototype of modular multilevel converters," *IEEE Transactions on Power Electronics*, vol. 29, no. 7, pp. 3267–3278, 2013.
[62] Y. Koyama and T. Isobe, "Current control of modular multilevel converters using a daisy-chained distributed control system with communication path redundancy," in *IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society*, vol. 1. IEEE, 2019, pp. 6108–6113.
[63] A. The, C. Bruening, and S. Dieckerhoff, "Can-based distributed control of a mmc optimized for low number of submodules," in *2015 IEEE Energy Conversion Congress and Exposition (ECCE)*. IEEE, 2015, pp. 1590–1594.
[64] S. Yang, Y. Tang, and P. Wang, "Seamless fault-tolerant operation of a modular multilevel converter with switch open-circuit fault diagnosis in a distributed control architecture," *IEEE Transactions on Power Electronics*, vol. 33, no. 8, pp. 7058–7070, 2017.
[65] S. Yang, Y. Tang, M. Zagrodnik, G. Amit, and P. Wang, "A novel distributed control strategy for modular multilevel converters," in *2017 IEEE Applied Power Electronics Conference and Exposition (APEC)*. IEEE, 2017, pp. 3234–3240.
[66] E. Espina, J. Llanos, C. Burgos-Mellado, R. Cardenas-Dobson, M. Martinez-Gomez, and D. Saez, "Distributed control strategies for microgrids: An overview," *IEEE Access*, 2020.
[67] C. Burgos-Mellado, F. Donoso, and T. Dragičević, "Ac battery: Modular layout and cyber-secure cell-level control for cost-effective transportation electrification," in *2022 IEEE Transportation Electrification Conference & Expo (ITEC)*. IEEE, 2022, pp. 1163–1167.
[68] F. Donoso, R. Cardenas, M. Espinoza, J. Clare, A. Mora, and A. Watson, "Experimental validation of a nested control system to balance the cell capacitor voltages in hybrid mmcs," *IEEE Access*, vol. 9, pp. 21 965–21 985, 2021.
[69] J. Llanos, D. E. Olivares, J. W. Simpson-Porco, M. Kazerani, and D. Sáez, "A novel distributed control strategy for optimal dispatch of isolated microgrids considering congestion," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6595–6606, 2019.
[70] X. Wu, Y. Xu, J. He, C. Shen, G. Chen, J. C. Vasquez, and J. M. Guerrero, "Delay-dependent small-signal stability analysis and compensation method for distributed secondary control of microgrids," *IEEE Access*, vol. 7, pp. 170 919–170 935, 2019.
[71] Y. Yan, D. Shi, D. Bian, B. Huang, Z. Yi, and Z. Wang, "Small-signal stability analysis and performance evaluation of microgrids under distributed control," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 4848–4858, 2018.
[72] H. Dong, Z. Ding, S. Zhang, H. Yuan, H. Zhang, J. Zhang, Y. Huang, T. Yu, H. Zhang, and R. Huang, *Deep Reinforcement Learning: Fundamentals, Research, and Applications*, H. Dong, Z. Ding, and S. Zhang, Eds. Springer Nature, 2020.
[73] S. Fujimoto, H. van Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," 2018. [Online]. Available: https://arxiv.org/abs/1802.09477
[74] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," 2015. [Online]. Available: https://arxiv.org/abs/1509.02971
[75] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, ser. Proceedings of Machine Learning Research, G. Gordon, D. Dunson, and M. Dudík, Eds., vol. 15. Fort Lauderdale, FL, USA: PMLR, 11–13 Apr 2011, pp. 315–323. [Online]. Available: https://proceedings.mlr.press/v15/glorot11a.html

Claudio Burgos-Mellado (Member, IEEE) was born in Cunco, Chile. He received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Chile, Santiago, Chile, in 2012 and 2013, respectively, and the dual Ph.D. degree in electrical and electronic engineering from the University of Nottingham, U.K., and in electrical engineering from the University of Chile, Santiago, Chile in 2019. From 2019 to 2021, he was a Research Fellow in the Power Electronics, Machines and Control Group (PEMC group) at the University of Nottingham, United Kingdom. Currently, he is an Assistant Professor with the Institute of Engineering Sciences, Universidad de O'Higgins, Rancagua, Chile. His current interests include battery energy storage systems, electrical vehicle technologies, power electronics, microgrids, power quality issues, modular multilevel converters, and cybersecurity in electrical systems. In 2021, he received the best PhD thesis award in the category of Exact Science from the Chilean Academy of Sciences.

Claudio Zuñiga-Bauerle was born in Parral, Chile. He received the B.Sc degree in electrical engineering from the University of Chile, Santiago, Chile, in 2023. Currently, he is working towards the Electrical Engineering professional Degree and is a staff member of the Automatic Laboratory both at the University of Chile, Santiago, Chile. In 2021 he was a research assistant at the Universidad de O'Higgins, Rancagua, Chile. He also coursed a semester in The Bologna Master Degree in Electrical and Computer Engineering at Instituto Superior Técnico Lisboa in 2022. His interests include predictive control, fuzzy models, and neural networks.

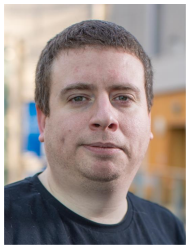
Diego Muñoz-Carpintero (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the Universidad de Chile, in 2009 and 2010, respectively, and the D.Phil. degree in control engineering from the University of Oxford, in 2014. From 2015 to 2016, he was a Research Fellow at Nanyang Technological University, and from 2017 to 2019, he was a Postdoctoral Researcher at the University of Chile. He is currently an Assistant Professor at the Institute of Engineering Science, Universidad de O'Higgins. His research interests include control theory in the areas of predictive, robust and stochastic control; intelligent control; fuzzy and neural modelling; decision making under uncertainty; and their application to energy efficient control of systems such as micro-grids, electric vehicle routing, battery management and agricultural irrigation.

Yeiner Arias-Esquivel (Student Member, IEEE) was born in Guápiles, Costa Rica. He received his B.Sc. degree in Electronic Engineering and M.Sc. degree in Embedded Systems from the Costa Rica Institute of Technology in 2014 and 2016, respectively. Currently, he is pursuing his Ph.D. degree at the University of Chile in Santiago, Chile. Since 2014, he has been with the Costa Rica Institute of Technology, where he currently serves as an Assistant Professor. His primary research interests encompass a range of topics, including Control of Modular Multilevel Converters, Model Predictive Control, Vibration Analysis, and Embedded Systems.

Roberto Cárdenas-Dobson (Senior Member, IEEE) was born in Punta Arenas, Chile. He received his B.S. degree from the University of Magallanes, Chile, in 1988 and his Msc. and Ph.D degrees from the University of Nottingham in 1992 and 1996 respectively. From 1989-1991 and 1996-2008 he was a lecturer in the University of Magallanes Chile. From 1991 to 1996 he was with the Power Electronics Machines and Control Group (PEMC group), University of Nottingham, United Kingdom. He is currently a full professor of power electronics and drives in the Electrical Engineering Department, University of Chile, Chile. Prof. Cardenas was a recipient of the 2019 Third Prize Paper Award from the IAS Industrial Power Converter Committee. He was also the recipient of the IEEE Transactions on Industrial Electronics Best Paper Awards in 2005 and 2019. From 2014 to 2021, Prof. Cardenas was an Associated Editor for the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS.

Tomislav Dragičević (Senior Member, IEEE) received the M.Sc. and the industrial Ph.D. degrees in Electrical Engineering from the Faculty of Electrical Engineering, University of Zagreb, Croatia, in 2009 and 2013, respectively. From 2013 until 2016 he has been a Postdoctoral researcher at Aalborg University, Denmark. From 2016 until 2020 he was an Associate Professor at Aalborg University, Denmark. Currently, he is a Professor at the Technical University of Denmark. He made a guest professor stay at Nottingham University, UK during spring/summer of 2018. His research interest is application of advanced control, optimization and artificial intelligence inspired techniques to provide innovative and effective solutions to emerging challenges in design, control and diagnostics of power electronics intensive electrical distributions systems and microgrids. He has authored and co-authored more than 330 technical publications (more than 150 of them are published in international journals, mostly in IEEE), 10 book chapters and a book in this field, as well as filed for several patents. He serves as an Associate Editor in the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, in IEEE TRANSACTIONS ON POWER ELECTRONICS, in IEEE Emerging and Selected Topics in Power Electronics and in IEEE Industrial Electronics Magazine. Prof. Dragičević is a recipient of the Končar prize for the best industrial PhD thesis in Croatia, a Robert Mayer Energy Conservation award, and he is a winner of an Alexander von Humboldt fellowship for experienced researchers.

Felipe Donoso (Graduate Student Member, IEEE) was born in Santiago, Chile. He received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Chile, Santiago, Chile, in 2014 and 2016, respectively, and the double Ph.D. degrees in power electronics from the University of Chile, Santiago, Chile, and the University of Nottingham, Nottingham, U.K., in 2021. From 2020 to 2022, he was a Research Associate with the Power Electronics, Machines and Control Group, University of Nottingham. Currently, he is a power electronics engineer with the Research and Development Department, Siemens Drives Technology unit in Congleton, UK. His research interests include control systems for power converters, resonant converters, DC-DC converters, modular multilevel converters, and renewable energy systems. Dr. Donoso was the recipient of the IEEE Transactions on Industrial Electronics Best Paper Awards, in 2019.



Alan Watson (Senior Member, IEEE) received the M.Eng. (Hons.) degree in electronic engineering from the University of Nottingham, UK in 2004, and a PhD, also from the University of Nottingham in 2008. In 2009, he became a Research Fellow with the Power Electronics Machines and Control Group, University of Nottingham. Since 2009, he has been involved in various projects in high-power electronics including resonant converters, high voltage power supplies, and multilevel converters for grid connected applications such as HVDC and Flexible AC Transmission Systems. In 2012, he was promoted to Senior Research Fellow before becoming an Assistant Professor in High Power Electronics in 2013. As of 2022 he is an Associate Professor in High Power Electronics. His current research interests include the development and control of advanced high-power conversion topologies for industrial applications, grid connected converters and HVDC Transmission.